

Data Warehouse Implementation To Support Batik Sales Information Using MOLAP

I Made Adi Bhaskara^{1*}, Luh Gede Putri Suardani², Made Sudarma³

^{1,2}Department of Electrical and Computer Engineering, Post Graduate Program, Udayana University

³Department of Electrical and Computer Engineering, Udayana University

*Email: adibhaskara39@gmail.com

Abstract— Batik product has spread in various regions in Indonesia. Batik companies usually have data that accumulates and accumulates without any follow-up to the data. It is also not supported by well-executed final reports. Therefore it is necessary to build a data warehouse that can be used as a source of information for Batik management related to the trend of the type of batik motif based on the category of goods and marketing area from time to time. One important process that must be done in the operation of data warehouse is the process of copying data from the operational database. Before the operational data goes into the data warehouse, ETL process (extract, transform, load) to the data is done. The process is intended to standardize the data used in the data warehouse. Meanwhile, the scheme designed for the development of data warehouse using Snowflake Schema model. The results showed that batik warehouse data has four dimension tables (Product dimension, Region dimension, Time dimension, and Customer dimension), four sub dimension tables (Category dimension, Sub_Category dimension, Pattern dimension and Gender dimension) and one Fact table, Fact Sales. The extraction process produces dimension tables (Product dimension, Region dimension, Time dimension and Customer dimension) and sub-dimensional tables (Category dimensions, Sub_Category dimensions, Pattern dimensions and Gender dimensions). All monitoring of sales data of batik products is done using cube browser. The information displayed by each dimension can be viewed in more detail with the drill down or roll up process in accordance with the hierarchy rules of each dimension field.

Keyword—ETL, Snowflake scheme, Cube browser, drill down, drill up

I. INTRODUCTION

The advancement of information technology has caused many people can get data easily and even tend to overdo it. The data is getting more and more and accumulated, consequently the utilization of the accumulated data is not optimal.

On the other hand, one of the major assets owned by a company is the huge amount of business data. This gives rise to the need for technology that can leverage it to generate new knowledge, which can help in the setting of business strategy. Prediction of consumer interest is very important for a company, where with the prediction that a company can take a decision or strategy that is right and right for the

consumer. Data warehouse and data mining technologies exist as solutions.

Batik retail companies that have spread in various regions in Indonesia. Many companies usually have accumulated and accumulated data without any follow-up to the data [9]. It is also not supported by well-executed final reports. The final report required by top management is reporting that can provide comprehensive information, so that based on the report the leader can determine the optimal policy in determining the production capacity so as to achieve efficiency in terms of cost and effective in time.

The problem occurs because of lack of knowledge about the utilization of such abundant data. Therefore, the availability of abundant sales data will be utilized for the development of a data warehouse which can then be used as a business solution for strategic marketing planning in the future.

Based on this background, the researcher intends to build a data warehouse that can be used as a source of information for the management of batik company associated with the company's strategic plan in the future related to the trend of the type of batik motif based on the category of goods and its marketing area.

II. Literature Review

A. Data Warehouse

Data warehouse is a place of data storage company or institution arranged in such a way that it contains meaning and for analysis and reporting. There are three or more leading approaches to storing data in a data warehouse — the most important approaches are the dimensional approach and the normalized approach.[11] So a data warehouse is a source of information whose data is obtained from Online Transaction Processing (OLTP) [5]. Usually this data warehouse store data that is historical. As Turban says, the Aronson and Ting data warehouse is a comprehensive database that supports all decision analysis required by an organization by providing a summary and detailed information. [1]

Meanwhile, according to Connolly and Begg, the data warehouse is a data set that is subject-oriented, integrated, time-variant, and non-volatile in supporting the decision making process [2].

The process of designing a data warehouse architecture is a very complicated process to begin with. We must define the needs of the most needed users and the data that should be prioritized. [8]

B. Data Warehouse Structure

Metadata is a form of information that contains data to be used, data type, data length and data source to be used in data warehouse. As Inmon says, metadata is data about data or descriptions of structures, contents, keys, indexes, etc. concerning data [3]. While Mallach, said metadata is data about data useful as a storage center information to explain to the user about whether the data warehouse, where it came from and who is responsible for it and so on. [4]

According to Connolly, Metadata is used for a variety of purposes including:

- Extraction and loading process. Metadata is used to map the data source into the general view of the data in the warehouse.
- Warehouse management process. Metadata is used to automate the creation of summary tables.
- As part of the query management process. Metadata is used to link a query with the right data source.

C. ETL (Extract, Transform, Load)

ETL is a very important process in the data warehouse, with ETL here the data from the operational can be entered into the data warehouse. ETL can also be used to integrate data with pre-existing systems [7]

The purpose of ETL is to collect, filter, process, and combine relevant data from multiple sources for storage into the data warehouse. The result of the ETL process is the generation of data that meets the criteria of data warehousing such as historical, integrated, summarized, static, and has a structure designed for the purposes of the analysis process.

- Extract

The first step in the ETL process is to extract data from the data sources. Most data warehouse projects combine data from different sources. Separate systems are very likely to use different data formats. Extraction is converting data into a format that is useful for the transformation process.

- Transform

Stages of transformation use a set of rules or functions to extract data from the source and then will be entered into the data warehouse. Here are the things that can be done in the transformation stage:

- Only select specific columns to put in the data warehouse.
- Translating values in the form of code.
- Codes values into free form (Example mapping "Men" with "P" and "Woman" into "W").
- Calculating new values (eg value = qty * harga_satuan).

- Combine data collectively from multiple sources.
- Create a summary of a set of rows of data.
- Generating surrogate key value.
- Transposing or pivoting (Changing a set of columns into a set of rows or vice versa).
- Separating a column into various columns.
- Using various forms of data validation either simple or complex.

- Load

The load phase is a step that serves to enter data into the final target, which is usually into a data warehouse. The duration of this process depends on the needs of the organization. Some data warehouses can weekly fill in the entire information cumulatively, data is changed, while other data warehouses (or other parts of the same data warehouse) can add new data in a historical form, for example every hour. The time and range to replace or add data depends on the design of the data warehouse when analyzing the information requirements.

The load phase interacts with a database, the constraint is defined in the database schema as a trigger that is activated during data load (eg uniqueness, referential integrity, mandatory fields), which also contribute to the overall look and quality of data from the ETL process.

D. Facts Table

- Concatenated key. The rows in the fact table are identified by using the primary key of the dimension table table, then the primary key of the fact table is the combined primary key of all dimension tables.
- Data grain, is the level of detail for measurement. For example, the number of bookings relates to the number of specific products on an order, a specific date, to a specific customer and is obtained by a specific sales representative. If the number of orders is seen as the amount for a product per month, then the data grain is different and at a higher level.
- Fully additive measures. Aggregation of fully additive measures is accomplished by a simple summation of the attribute values.
- Semi additive measures. Semi additive measures are values that can not be directly summed, for example the percentage of profit.
- Large table, not wide. The fact tables generally have fewer attributes than the dimension table, but have a larger number of records.
- Sparse data. The fact table does not need to store records whose values are null. Then the fact table can have a gap.
- Degenerate dimensions. There are data elements of the operational system that are not facts or dimensions, such as order number, bill number, and so on. However, these attributes can be useful in certain types of analysis. For example, look for an average number of products.

E. Multidimensional Online Analytical Processing (MOLAP)

MOLAP stands for Multidimensional OLAP which is an OLAP type that has its own storage. The advantage of MOLAP is the performance of excellent data access speed. But the disadvantage is that if the combination of data aggregation is generated for all levels, then the storage size will be larger than the source itself. [10].

F. Snowflake Scheme

Snowflake Scheme are another variation of the star schema model, where several dimensional tables have been normalized, and there are some separate data into several new tables. The advantage of multidimensional formats, snowflakes schema is a normalized table to reduce data redundancy, but the time required for query processing becomes longer. [6].

III. RESEARCH METHODOLOGY

This research includes the type of research and development because it will research and develop a data warehouse to produce production information, batik trends in a particular region to support the strategic plan of product marketing in the future.

Architectural planning in the development of data warehouse includes logical architecture and physical architecture. Logical architecture is the design of the data stream stage from the data source to the data in the data warehouse, while the physical architecture is a description of the technical configuration of the data warehouse. Architectural planning is designed in accordance with existing conditions in the company. The logical architecture design of a data warehouse can be seen in Figure 1.

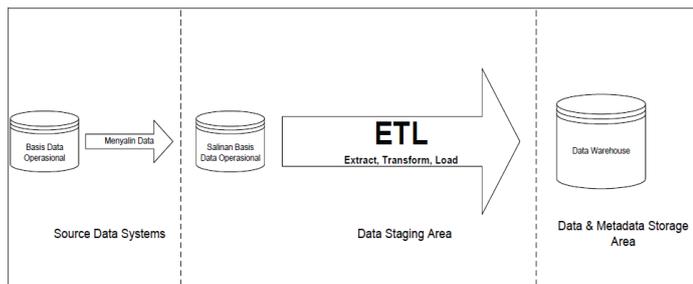


Figure 1. Architectural Logical Data Warehouse Product Marketing

Data needed to build the product marketing data warehouse is sourced from the company's operational database. Based on the results of data retrieval in XYZ company, data recorded in the database comes from daily sales transaction records stored in Microsoft Excel application.

Table 1 is a sales transaction data conducted by PT. XYZ in 2010.

Table 1. Sales Transaction Data Batik

	A	B	C	D	E	F
1	WAKTU	NAMA BARANG	HARGA	JUMLAH	PEMBELI	DAERAH
2	2010-01-02	sarimbit Standar Tulis Lukis	400000	11	Ibu Hatamah	gresik
3	2010-01-04	longdress Standar Tulis Gradasi	170000	17	Ibu niken	jakarta pusat
4	2010-01-05	kaos anak Standar Print Gradasi	55000	4	Bapak heru	malang
5	2010-01-05	kaos anak Standar Print Kebumen	55000	4	Bapak heru	malang
6	2010-01-05	kaos anak Standar Print Lukis	55000	4	Bapak heru	malang
7	2010-01-06	bolero Standar Cap Kontemporer	225000	1	Bapak ardik	batam
8	2010-01-06	bolero Standar Cap Mahkota	225000	1	Bapak ardik	batam
9	2010-01-06	sarimbit Standar Tulis Semen rama	400000	11	Ibu niken	pekalongan
10	2010-01-14	selendang Standar Cap Sogan	90000	2	Ibu siti arya	ciamis
11	2010-01-14	selendang Standar Cap Tumpal	90000	2	Ibu siti arya	ciamis
12	2010-01-18	bahan Lurik Print Abstrak	90000	1	Ibu atik	semarang
13	2010-01-18	bahan Lurik Print Kelengan	90000	1	Ibu atik	semarang
14	2010-01-18	bahan Lurik Print Sogan	90000	1	Ibu atik	semarang
15	2010-01-19	celana pendek Standar Cap Abstrak	55000	17	Ibu hadi sukarni	bogor
16	2010-01-19	celana pendek Standar Cap Sogan	55000	17	Ibu hadi sukarni	bogor
17	2010-01-19	celana pendek Standar Cap Truntum	55000	17	Ibu hadi sukarni	bogor
18	2010-01-28	hem Standar Cap Gradasi	125000	1	Ibu tyas	kudus
19	2010-01-28	hem Standar Cap Jumput	125000	1	Ibu tyas	kudus
20	2010-02-06	setelan Santung Print Mahkota	60000	17	Ibu niken	pekalongan
21	2010-02-06	setelan Santung Print Wahyu tumurun	60000	17	Ibu niken	pekalongan
22	2010-02-07	hem Standar Tulis Mahkota	300000	10	Ibu tyas	kudus
23	2010-02-07	hem Standar Tulis Mahkota	300000	10	Ibu tyas	kudus
24	2010-02-07	kaos batik Standar Cap Lahar merapi	100000	1	Ibu tyas	kudus
25	2010-02-07	kaos batik Standar Cap Sogan	100000	1	Ibu tyas	kudus

Data Warehouse developed in batik company will be used to answer questions that are strategic to support decisions to be taken by the company. In accordance with the research objectives, the information to be extracted from the data warehouse developed at PT.XYZ is "the most sold batik product based on its category and marketing area".

One important process that must be done in the operation of data warehouse is the process of copying data from the operational database. The process of copying the data is done every period of time, especially after a transaction that resulted in the addition or change of data in the operational database.

Only a few data in the operational database is used for the purpose of building a product marketing data warehouse. Only data are related and needed to be presented in the product marketing data warehouse that is used. Place of execution of the process of selection and preparation of data that is loaded into the data warehouse is called staging area. Data residing in the staging area for further processing into the data warehouse is known as staging data.

Before data goes into the data warehouse, ETL process (extract, transform, load) to the data is done. The process is intended to standardize the data used in the data warehouse.

Meanwhile, a scheme designed for the development of data warehouses in batik companies using the Snowflake Schema model. This model is chosen with the consideration of being able to provide information sought easily by the company that if using transactional database such information is very difficult to obtain.

IV. ANALYSIS OF RESULT AND DISCUSSION

In 2010 there has been recorded sales transaction data conducted by PT.XYZ as much as 1201 times transaction, which is stored in Microsoft Excel application. Furthermore, the transaction data will be used as a data source for the data warehouse, so that will be the process of extraction, transformation and load on the data.

1. Snowflake Design Results

The main information to be extracted in the data warehouse is "the most sold product by category and marketing area for a certain time". Based on the question asked then from the source database transactional records owned by the company of batik can be raised 4 (four) dimensional table and 1 (one) fact table (fact table), as follows:

1. Product dimension, connected with 3 (three) sub dimension that is category, subcategory and pattern.
2. Region dimension
3. Time dimension
4. Customer dimension, connected with 1 (one) sub dimension that is gender
5. Sales Facts Penjualan

Based on the results of determining the tables, then the relationships between dimensions and fact tables are designed in the form of Snowflake Schema shown in Figure 2.

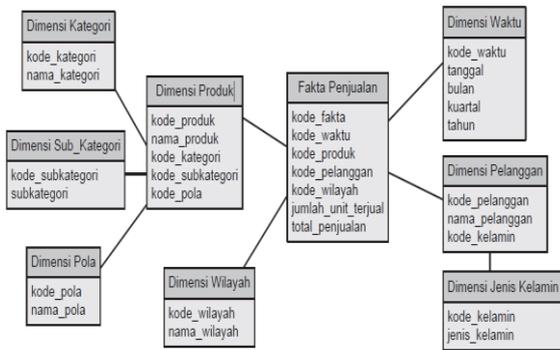


Figure 2. Snowflake Scheme applied

2. ETL Process (Extract, Transform and Load)

ETL process is performed to move the transactional data from the data source table to the data warehouse. However, since the data retrieved from the source table is inconsistent with the format and standard of the data warehouse, an ETL process is required so that data entered into the data warehouse conforms to the designed snowflake schema standard. The overall ETL process diagram can be seen in Figure 3..

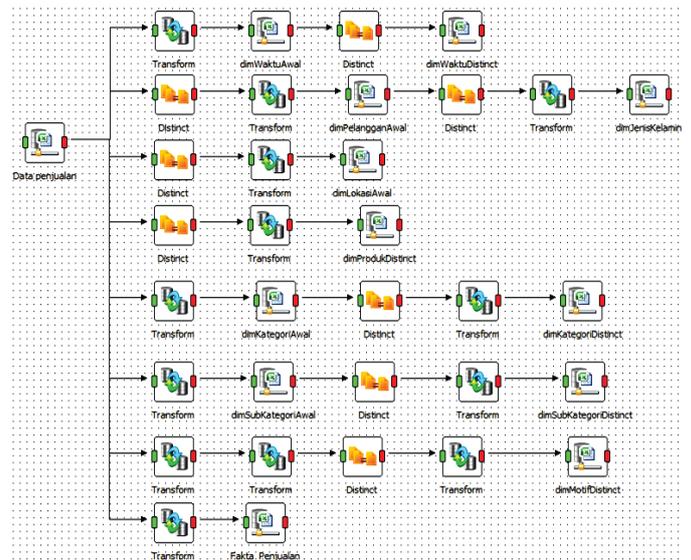


Figure 3. Overall ETL Process Diagram

a) ETL process product dimension.

Product dimensions are tables that contain specific data about the identity of the product. This dimension consists of 2 (two) columns namely product_code and product_name. Separation of product data from sales data required special ETL process for product data, as shown in figure 4..

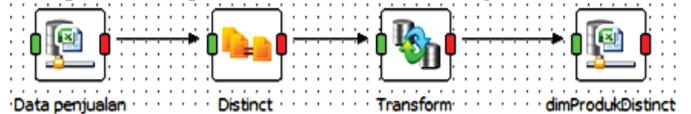


Figure 4. ETL Product Dimension Process Diagram

The result of the ETL process is the Product Dimension table shown in table 2. In this table, in addition to extracting the data of product names, also created an index structure as the primary key on the product dimension.

Table 2. Product Dimension Table

KODE_PRODUK	NAMA_PRODUK
PRO-8B20	BAHAN STANDAR CAP PARANG KUSUMO
PRO-90CC	BOLERO STANDAR CAP GRADASI
PRO-7EF7	KAOS BATIK STANDAR CAP SIDO ASIH
PRO-B447	HEM STANDAR PRINT KELENGAN

b) ETL Process Category Dimension

The category dimension is a table that contains specifically about the data categories of a product. The result of the ETL process, this dimension consists of 2 (two) columns namely code_category, and nama_kategori. Separation of product category data from sales data required special ETL process for product category data, as shown in figure 5..

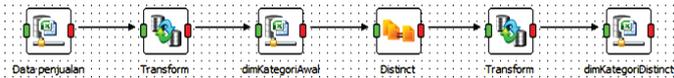


Figure 5. ETL Process Dimension Category Diagram

The result of the ETL process is a table of Category Dimensions shown in the example in Table 3. In this table, in addition to extracting data category names, also created an index structure as the primary key in the product category dimension.

Table 3. Category Dimension Table

KODE_KATEGORI	NAMA_KATEGORI
K-3236	Blus
K-BFC5	Celana
K-5292	Bolero
K-EEE0	Kaos

c) ETL Process Dimension Sub Category

Dimension sub category is a table that contains special about data sub category of a product so that information contained in more detail. The result of the ETL process, this dimension has 2 (two) columns namely code_subkategori, and subcategories. Separation of data sub category of product from sales data required special ETL process for data sub category, as shown in picture 6.

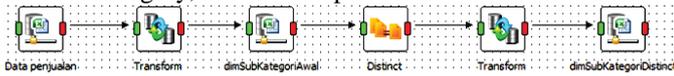


Figure 6. ETL Process Dimension Diagram Sub_Category

The result of the ETL process is a table of Dimensions Sub Category shown in Table 4. In this table, in addition to extracting the data of sub-category names, also created an index structure as the primary key in the dimension of the product sub category.

Table 4. Dimension Table Sub Category

KODE_SUBKATEGORI	SUBKATEGORI
SUB-F437	Polar
SUB-28D6	Sby

The same way that can be done to the other dimensions so obtained the tables as follows:

The result of the ETL process from the Sales Data source table is the Regional Dimension table shown in Table 5. In this table, extracted the data of the city names, and also created an index structure as the primary key in the region dimension with the name of the kode_lokasi column.

Table 5. Region Dimension Table

KODE_LOKASI	KOTA
W-DE13	CILACAP
W-1F1C	MALANG
.....
W-01E1	SEMARANG
W-A15A	JAKARTA SELATAN

The result of the ETL process from the Sales Data source table is the Customer Dimension table shown in the example in Table 6. In this table, has extracted the data of the customer names and the gender code. This dimension also created an index structure as the primary key on the customer dimension with the name of the kode_pelanggan column.

Table 6. Customer Dimension Table

KODE_PELANGGAN	NAMA_PELANGGAN	JENIS_KELAMIN
P-0FF9	BAPAK SUGENG	1
P-31FD	BAPAK IMRON	1
.....
.....
P-F063	IBU ENDAH SARI	0
P-4509	IBU HERUYONO	0
P-2520	BAPAK HERU	1

The result of the ETL process from the Customer Data source table is the Sex Dimension table shown in the example in Table 7. In this table, it has extracted the data of the customer's gender in order to know the trend of the customer by sex. This dimension also created an index structure as the primary key in the gender dimension with the column name kode_jk.

Table 7. Sex Dimension Table

KODE_JK	JENIS_KELAMIN
1	PRIA
0	WANITA

The result of the ETL process from the Sales Data source table is the Time Dimension table shown in the example in Table 8. In this table, extracted the time data to be able to know the trend of transactions conducted based on time. This dimension also created an index structure as the primary key in the gender dimension the column name kode_waktu .

Table 8. Time Dimension Table

KODE_WAKTU	TANGGAL	BULAN	TAHUN	KUARTAL
40180	02	1	2010	1
40182	04	1	2010	1
40183	05	1	2010	1
40184	06	1	2010	1
40192	14	1	2010	1

The result of the ETL process from the Sales Data source table is the Dimension of the Pattern shown in Table 9. In this table, it has extracted the name_pola data in order to know the product trend based on the pattern name. This

dimension also created an index structure as the primary key in the pattern dimension with the kode_pola column name.

Table 9 Table Dimension Pattern

KODE_POLA	NAMA_POLA
P-01	CAP
P-02	PRINT
P-03	TULIS
P-04	LAINNYA

C. Facts Table Creation

Sales (Sales Fact) From the results of ETL process four of the eight dimensions of time dimensions, products, regions, and customers, in each dimension has formed a column as an index (primary key). This index is then used as a reference key to be included in the Facts Sale table. Thus, the sales fact table contains only 1 primary key column, the foreign key columns of the 4 main dimensions and 2 columns as the measure.

Table 10 shows the data snippet from the sales transaction table after being converted into a Sales Fact table consisting of 7 columns..

Table 10 Sales Facts Table

KODE FAKTA	KODE WAKTU	KODE PRODUK	KODE PELANGGAN	KODE DAERAH	JUMLAH	TOTAL PENJUALAN
1	40180	PRO-68FD	P-1EE1	W-0ACE	11	4400000
2	40182	PRO-6BB3	P-958B	W-4HE4	17	2890000
3	40183	PRO-C895	P-2520	W-1F1C	4	220000
4	40183	PRO-D97F	P-2520	W-1F1C	4	220000
5	40183	PRO-062F	P-2520	W-1F1C	4	220000
.....
.....
.....
.....
1197	41266	PRO-7866	P-9440	W-1190	23	575000
1198	41268	PRO-1760	P-C3B9	W-58A1	12	2160000
1199	41270	PRO-0A6A	P-C3B9	W-58A1	21	1890000
1200	41270	PRO-7697	P-5B88	W-5AE1	25	5000000
1201	41273	PRO-5918	P-B26D	W-0D39	1	125000

Furthermore, the process of exporting data to PostgreSQL DBMS to be processed using Analysis Manager Microsoft SQL Server as data warehouse in the form of cube

D. Design and Making of Cube.

The data warehouse used in 2010 is ready to be built using data in the PostgreSQL DBMS. The data will be displayed in the browser cube provided in the Microsoft SQL Server application. However, the cube creation stage takes a long time because there are some items that need to be set and configured.

Below are the steps of making multidimensional sales data cube:

1. Data Source Configuration

Data source needs to be configured to specify the PostgreSQL database that is used as the new cube creation data source.

2 Design of Snowflake Model in Cube

The next stage is the process of determining the tables in the data source that will be displayed multidimensi in cube browser and field rules in accordance with the hierarchy for the process of roll up and drill down. All tables contained in the PostgreSQL DBMS are used as the source of the snowflake model which is then displayed multidimensionally.

Sales Facts Table is used as a data analysis center in snowflakes. The columns in the form of foreign keys refer to the connected dimensions. While the sum and total_ sales columns are used as a measure. This sales fact is then displayed multidimensional using cube browser with a named cube "Sales".

The dimensioning rules as well as the field hierarchy used in the "Sales" cube are :

- a. Dimensions consisting of 4 main dimensions: time, customer, region, and product.
- b. The measure rule is intended to determine the columns that are the size of the data analysis of sales transaction data that occurs. In Sales Facts, the field set as the measure is the Amount to see the number of units of products sold and Total Sales to know the value of the sale of the product.

3. Storage Design

Storage Design used MOLAP type (Multidimensional Online Analytical Processing) in accordance with the purpose of making data warehouse used to view data multidimensi. Aggregate function setting is performed using the Performance Gain rule of 40%.

4. Cube Sales Results

The built-in sales data warehouse is ready to use and is displayed multidimensional by using Cube Browser. The initial view of Sales cube as shown in figure 6.

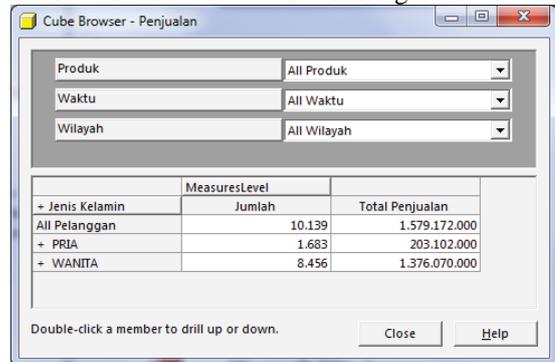


Figure 6. Cube Sales

Figure 6 shows in the cube there are 4 (four) dimensions that is Product, Time, Region and Customers with Customer dimension contained in table, and 2 (two) field of measure that is Total and Total Sales.

Each dimension can be inserted or removed from the table to view the data according to the needs of the analysis by drag and drop. In addition, the information displayed by each dimension can be viewed in more detail with the drill

down or roll up process in accordance with the hierarchy rules of each dimension field.

An example of a cube display to find detailed information regarding the number of products sold and total sales by Product and Time for all customers and regions can be seen in figure 7.

Nama Pota	Subkategor	- Tahun		+ 2010		+ 2011	
		Jumlah	Total Penjualan	Jumlah	Total Penjualan	Jumlah	Total Penjualan
All Produk	All Wilayah Total	10.139	1.579.172.000	3.096	545.645.000	2.083	371.497.000
	CAP Total	2.229	299.982.000	1.091	73.130.000	393	42.372.000
	- AMAR	5	275.000			5	275.000
	- BATAK	176	37.434.000	399	27.063.000		
	- BELUKU	4	2.000.000	2	1.000.000		
	- KARTUN	229	15.568.000	96	7.110.000	38	1.760.000
	- LUPIS	65	5.980.000	22	2.640.000		
	- PAKSI	2	300.000				
	- SANAPORISE	31	6.200.000			17	3.400.000
	- SAMPUNG	6	420.000	2	140.000	2	140.000
	- SELINDANG	6	540.000	4	360.000		
	- SMOKE	1	130.000			1	130.000
	- STANJAR	1.305	137.560.000	476	52.840.000	239	35.950.000
	- SURIA	1	1.150.000				
	- VSC-COSE	9	925.000			3	675.000
	LADANG Total	280	56.122.000	96	15.035.000	106	15.114.000
	- BATAK	2	250.000				
	- BELUKU	1	500.000				
	- PALAMEN	1	90.000	1	500.000		
	- KARTUN	3	240.000			3	240.000
	- LARANGAN	1	175.000				
	- POLAR	57	16.728.000			25	5.235.000
	- SANUNG	9	1.839.000			5	1.119.000
	- STANJAR	211	36.004.000	95	14.535.000	73	12.500.000
	- SURIA	2	1.495.000				
	- TREST	1	799.000				
	Unlabeled Total	4.881	487.134.000	1.114	107.167.000	1.198	117.103.000

Figure 7. Sample Cube Sales by Product and Time

Cube Browser can be used to answer strategic questions related to sales data. Examples of strategic questions that are often raised include:

1) What is the total and total sales of print pattern products that occurred in Bogor City in Q1 of 2010?

- Propinsi	Kota	- Tahun		+ Kuartal	
		All Waktu Total	2010 Total	- 2010	+ 1
All Wilayah	All Wilayah Total	4.982	1.574		292
+ BALI	BALI Total	198	7		5
+ DKI JAKARTA	DKI JAKARTA Total	236			
	JAWA BARAT Total	1.567	388		51
	BEKASI	4			
	BOGOR	943	298		51
	CIAMIS	598	90		

Figure 8. Information on the number and total sales of print pattern products that occurred in Bogor City in the first quarter of 2010

2) How is the comparison of number of users by gender for print pattern products in 2010?

MeasuresLevel	+ Jenis Kelamin		
	All Pelanggan	+ PRIA	+ WANITA
Jumlah	4.982	1.214	3.768
Total Penjualan	487.735.000	87.330.000	400.405.000

Figure 9. Comparison of number of users by sex information for print product patterns in 2010

II. CONCLUSION

Based on the results of research conducted it can be concluded that:

1. Data warehouse PT.XYZ has four dimension tables (Product dimension, Region dimension, Time dimension and Customer dimension), four sub dimensions table (Category dimension, Sub_Category dimension, Pattern dimension and Sex dimension) and one Fact table, Fact Sales.
2. The extraction process is done so that the transaction data in accordance with the data warehouse format. The extraction process produces dimension tables (Product dimension, Region dimension, Time dimension and Customer dimension) and sub-dimensional tables (Category dimensions, Sub_Category dimensions, Pattern dimensions and Gender dimensions).
3. All monitoring of sales data of PT.XYZ batik products is done using cube browser, so that each dimension can be inserted or removed from the table to see the data according to the needs of analysis by drag and drop. Information displayed by each dimension can be viewed in more detail with the drill down or roll up process in accordance with the hierarchy rules of each dimension field.

ACKNOWLEDGMENT

Thank you to Lecturer and to all parties who have helped and supported the author, as well as the support of fellow writers in the environment until the research is completed.

REFERENCES

- [1] Turban, Efraim, Rainer, R. Kelly Jr., Potter, Richard E.(2005).Introduction to Information Technology. 3rd Edition, John Wiley & Sons,Canada.
- [2] Connolly, Thomas dan Carolyn Begg.(2005). Database Systems: A Practical Approach to Design, Implementation, And Management, 4th Edition. Addison Wesley Publishing Company Inc., California.
- [3] Inmon, W.H. (2005). Building Data warehouse, 4th Edition. John Wiley& Sons, Canada.
- [4] Mallach, Efrem G. (2000). Decision support and data warehouse systems. McGraw-Hill.
- [5] Djoni Darmawikarta, Mengenal Data Warehouse, Ilmu Komputer, 2003
- [6] Haryanto, D., Oslan, Y., Dwiyan, D., "Implementasi Analisis Keranjang Belanja Dengan Aturan Asosiasi Menggunakan Algoritma Apriori Pada Penjualan Suku Cadang Motor", Jurnal Buana Informatika Volume 2, Nomor 2, Juli 2011.
- [7] Darudiat, S., "Perancangan Data Warehouse Penjualan untuk Mendukung Kebutuhan Informasi Eksekutif Cemrlang Skin Care", Seminar Nasional Informatika 2010 (SemnasIF 2010), May 2010