DEVELOPMENT OF DATA WAREHOUSE ON INSTITUTIONS FOR AGRICULTURAL TECHNOLOGY (BPTP) WEST JAVA

Hengky Saputra

Teknik Informatika – Universitas Komputer Indonesia
Jl. Dipatiukur 112-114 Bandung
Email : sputrahengky@gmail.comm

ABSTRACT

Institutions for agricultural technology on West Java are technical execution unit (UPT) research agency and as agricultural development that structurally are under supervision Big hall for the assessment and the development of agricultural technology (BBP2TP) in Bogor With the working areas covering West Java province. At this time BPTP West Java a tools for supporting as yet there can be used for making the report where this report in making use of data consists various types of the type of data and data there are arrayed in accordance periodic or historical . This might impact the work of the officers in BPTP In terms of search data needed to make a report or an analysis of information there to be used to the needs of the identification of agricultural productivity, plantations or livestock. Because there is not yet management of data that both this is risky against loss of data and safety data.

Because of that required a data storage to designing complex data and could make it suitable for making analysis of data and decision-making is to build data warehouse. Data warehouse has a great many advantages. Data stored only in the form of metadata so that the data cannot be modified. The nature of data warehouse integrated data and other data is historical. The Construction of data warehouse is equipped with tools such as OLAP and reporting tools. This Tools provides convenience for the user to be able to design complex data, access data quickly, data analysis and reporting and decision making can be done quickly and precisely.

Based on the analysis and testing that has been done, Data Warehouse can give ease in access to information and combines data can be in the data warehouse with various ways. So that it can display the most appropriate information and fast. An integrated information and contains relevant decision makers bits on information to a decision and maximize the quality of the decision made by top management at BPTP.

Keywords: Data Warehouse, Warehousing, ETL, OLAP, reporting tools.

1. INTRODUCTION

Institutions for agricultural technology on West Java are technical execution unit (UPT) research agency and as agricultural development that structurally are under supervision Big hall for the assessment and the development of agricultural technology (BBP2TP) in Bogor With the working areas covering West Java province. In order for the development of the agriculture is based on the concept of efficiency to win the comparative excellences and competitiveness in facing globalization era trade, then BPTP West Java It needs to be to identify seed commodities in west java province.

Based on observations , not the availability of the design of the data can be used to the needs of the report .The data for this type of data consisting of various types of data and there was not arrayed according to periodic or historical . This affect the working time because it requires a longer time in data processing to be used to analyze existing information. Besides, for there is no good management to its data and will run the risk of losing the data. Other causes of support lies in the absence of tools that can be used for making the report.

Because of it , the solution given to BPTP to build applications that help in changing the archives of data consisting of a variety of the resources to be knowledge with information that is integrated with one another. Easily accessible source of data and fast it could help to improve the performance in making the data and analysis on products bptp for research of west java. The data storage of data that can support the application of the warehouse and speed the process of etl in the process of integrating the data there to be new information , and can also become one of a solution that can accelerate the process of data collection and multidimensional and concise presentation of information that can maximize the quality of decisions being made .

Data warehouse Is a collection of the data have the nature of oriented subject, integrated time-variant, and tend to stay off a collection of data in favor of the decision-making process management [1]. In making a data warehouse there is a phase where the transformation of data this phase it aims to integrate data from source data into data warehouse. The

resulting in phase transforming it will be used OLAP tools to analyze data as well as reporting tools to produce reports required by worker in BPTP.

The purpose of which will be achieved from this research is as follows

1. Provide facilities for access to information in bptp as required by the and it can also combines data warehouse in various ways. So we can put it right and fast.
2. To provide information integration. Data Wareshouse Contains relevant to the needs of the party look bptp in a decision makers.
3. Help speed up the process in data collection and presentation of information that multidimensional and concise so as to maximize the quality of the decision made by top management .

1.1 Data Warehouse

Data warehouse Is a place of data storage is complete and consistent where data stored berelasi each other and designed based on the query and analysis rather than process their transactions [6]. The data in the transformation of right into a information that can be accessed at any time and always up to date. This information will then used to analyzed to produce new information

Data warehouse Can also spoken of as a collection of data that is oriented the subject , integrated , could not updates, having dimensions time used to support the process management decision-making and business intelligence. Based on that definition, maka data warehouse Having the characteristics of [6] such as :

a. Subject oriented
b. Integrated
c. Time-variant
d. Non-volatile

1.2 The Purpose of the Data Warehouse

The following are the purpose of a data warehouse and an explanation [9] :

1. Provide easy access to existing information.
2. Provide consistent information.
3. Be able to adapt and resistant to the environment.
4. Ability to secure information.
5. Able to give support in decision making.
6. User friendly.

1.3 Data Warehouse Architecture

Architecture on data the warehouse grouped into five parts such as Figure 1.1 [10] :

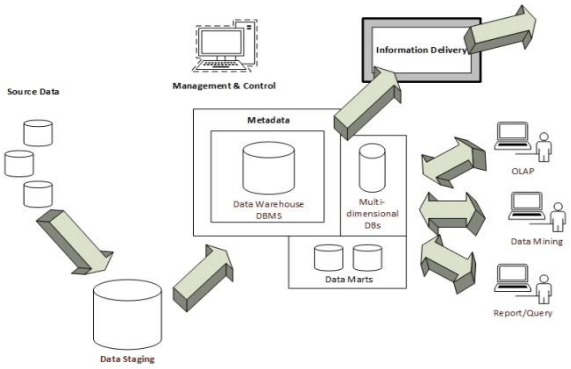


Figure 1.1 Architecture on data the warehouse [10]

1.4 Dimensional Model Data Warehouse

User requirements and realities of data that becomes the deciding factor for designing dimensional data warehouse, such as what is the most necessary business, such as what the details and dimensions as well as facts of what should be included. [6].

Then dimensional must be in the model consistent with the needs from the user. The model also must be designed in such a way so it can stand and can adapt all of the changes will occur. Design for model produced formed into a relational database that supports OLAP cubes to provide the results in instant query to analysts.

1. Dimension Tables
   Dimension table describes the business entities of an enterprise [6]. Dimension tables generally contain data information, where data rarely changes.
2. Fact Tables
   Fact table is a table that describes the business transactions of an enterprise is usually called detail tables [6]. General fact table contains data that is directly related to its business processes.
3. Scheme Dimensional Model
   Scheme dimensional model consist of:

a) Scheme Star
A scheme called star scheme if all dimension tables linked directly to the fact table and the fact table is required to have a relationship with at least one dimension table [6]. Figure 1.2 is an example of star schema.

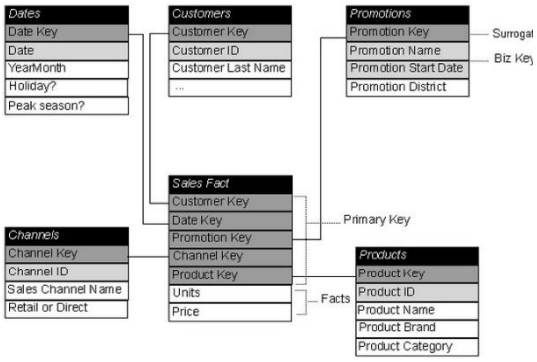


Figure 1.2 Star Schema [7]

b) Snowflake Schema

A scheme called snowflake schema if one or more dimension tables are not directly related to the fact table but must be related through another dimension tables [6]. Figure 1.3 is an example image of the snowflake schema.

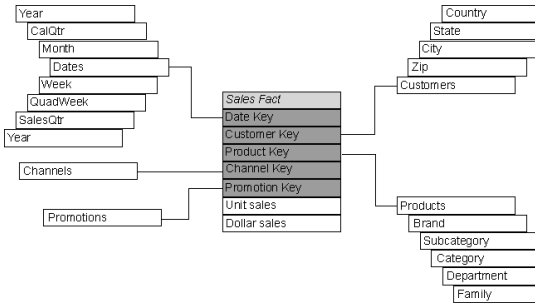


Figure 1.3 Snowflake Schema [7]

c) Constellation Schema

A scheme is said as a constellation scheme if there is one dimension tables that are used simultaneously by one or more fact tables [11]. Figure 1.4 is an example image of the constellation scheme.

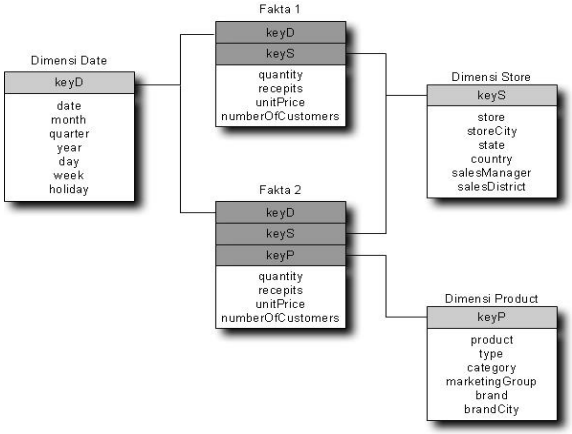


Figure 1.4 Constellation Schema [11]

1.5 ETL Process Data Warehouse

ETL process or so-called Extract, Transform, and Load is the process of converting data from the OLTP database into a data warehouse. When viewed from architecture data warehouse, ETL process is a process that is in the data staging.

ETL process is a process to convert, reformat and integrate data coming from one or several OLTP systems. [12].

1. Extraction
   Extraction is a process where the process of searching for the source of the data and then using some of the criteria that have been given to sort the data and also to look for high-quality data, then the data is transported to another file or database. [12].
2. Transformation
   Data transformation is a phase that occurs when the data has become raw data (the results of extraction) converted into a form that has been set whereby the form must be used in data warehouse [10]. The following are some process of the base that there must be in the data transformation: `
   a. Selection
   b. Splitting/Joining
   c. Conversion
   d. Summarization
   e. Enrichment
3. Loading
   Loading is a process of data transfer from OLTP systems physically into the data warehouse. Loading operation consists of inserting records into various dimensions and fact tables of a data warehouse.

1.6 Data Warehouse tools

Here is a tool that is used by user after the data warehouse is formed with a different purpose: [7] :

1. OLAP (On-Line Analytical Processing)

OLAP warehouse is one of the tools to perform data analysis. OLAP itself is a technology designed to deliver superior performance for ad hoc business intelligence queries [7]. is designed to operate efficiently with the data that is organized in accordance with the general dimensions of the model used in the data warehouse.

Function of OLAP [10] such as:

a) Increase productivity from business manager, executives and analysts.
b) Olap using with good users can make can with confidence to make their own analysis without the assistance IT assistance.
c) An advantage to IT developers namely the use of olap could be most helpful own speed up the performance of his application.
d) Increase the efficiency of work.

Olap can be used to do things like [9]:

a) Consolidation (roll-up)
Consolidated data involving grouping to see data is global and a summary (summary).
b) Drill-down

A form that is the opposite of consolidation , to get more details about a dimension and it can be as a navigation from the more common to the more specific. For more explained roll-up and drill-down can be seen in Figure 1.5

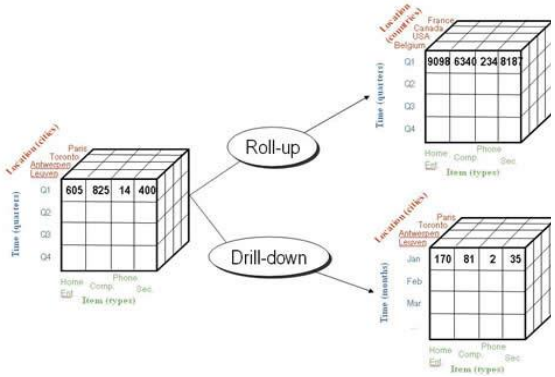


Figure 1.5 Roll-up dan Drill-down

c) Slicing and dicing

Slicing and dicing operations are to see data as visualizing cube. By slicing and dicing of users can see the data from some perspective. Users can extract part of the data agregrated and can check with detail based on the dimensions of desired. Data agregrated is precalculated data in the form of a summary of data so that query on cube cube more quickly . Slicing cut cube so that it can be focused on specific perspective (at a dimensions). While dicing give the ability to see the election of data in two dimensions or more .Namely by swap perspective cube on another so users could see more specific against that analyzed data. To further explained slicing and dicing can be seen in Figure 1.6

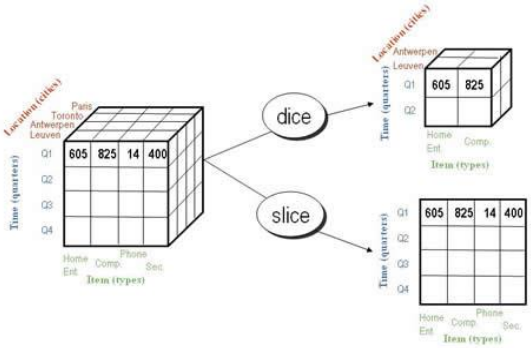


Figure 1.6 Slicing and Dicing

d) Pivot

Showing values size in the layout table different dimensions and can also set back in olap cube. To further explained slicing and dicing can be seen in Figure 1.7

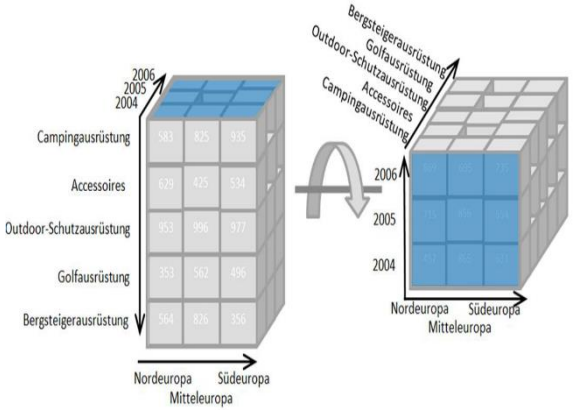


Figure 1.7 Pivot

2. Reporting

Reporting tools are tools that are used to facilitate users obtain the data that is old or present the data and perform some standard statistical analysis [10]. The data generated from the reporting tools can be a form of regular reports and graphs can also be.

3. Data mining

Data mining is a technology that applies sophisticated and complex algorithms to analyze the data and look for interesting information from the data set. The fundamental difference between OLAP and data mining that is located in what would be analyzed. In OLAP, the model is analyzed, but the analysis of data mining is the data (must be a large amount). [7].

3. CONTENTS RESEARCH

2.1 Method Of Data Warehouse Development

Method of development on data warehouse at an Institution for agricultural technology (BPTP) west java consisting of several stages namely [3]:

1. Business Requirement Definition
This stage is the analysis of business process and all the needs of the agricultural sector that is at an Institution for agricultural technology (BPTP) west java in building data warehouse.

2. Dimensional Modeling
This stage is the stage of modeling the data into multidimensional data based on the results of the business requirements defintion.

3. Physical Design
Tahap ini merupakan tahap perancangan fisik data warehouse. Seperti hardware dan software yang dibutuhkan, banyaknya memory yang diperlukan, pembentukan partisi jika diperlukan, dan lain-lain.

4. Data Staging Design
This stage is the stage of the physical design of the data warehouse. As the hardware and software needed, the amount of memory required, the formation of the partition if necessary, and others.

5. OLAP & Reporting Tools
   This stage is the stage to set the existing data in the data warehouse into a multidimensional cube based dimensional model that will be displayed to the user for decision-making.
6. Deployment
   This stage is the stage of the operation of data warehouse also reporting tools has been finished.

3.2 Data Source

An analysis of a source of data is the process of analyzing a source of existing data at an Institution for agricultural technology (BPTP) west java. Data sources it consists of several flat file.

3.3 Strategic Information Needs

According to the interviews at an Institution for agricultural technology (BPTP) west java. Obtain information for strategic business as follows:

1. Extensive information on crop yields of cultivated land area for any commodity in each district every month every year.
2. Extensive information crop yields compared to an area of land of various products in every region each year every month.
3. The number of damaged information on the basis of an attack an some pest in total damage in every sub-district every month every year.
4. Information the amount in damages based on the type of attack an some pest in total damage in each district every month each year.
5. The number of food information most excellent in each district any his month of the year.
6. The number of vegetables information most excellent in each district any his month of the year.
7. The number of fruits information most excellent in each district any his month of the year.
8. The number of livestock information most excellent in each district any his month of the year..
9. The number of food information most excellent in each region any his month of the year.
10. The number of vegetables information most excellent in each region any his month of the year.
11. The number of fruits information most excellent in each region any his month of the year.
12. The number of livestock information most excellent in each region any his month of the year.
13. The results of the information the amount of food in each district every month each year.
14. The results of the information the amount of fruits in each district every month each year.
15. The results of the information the amount of vegetables in each district every month each year.
16. The results of the information the amount of livestock in each district every month each year.
17. Information lq food in each district every year.
18. Information lq fruits in each district every year.
19. Information lq vegetables in each district every year.
20. Information lq livestock in each district every year.
21. The total number of land information in each district every month every year .
22. The total number of land information in each region every month every year

3.4 Dimension and Fact Business

Seen from the strategic information needs of Institution for agricultural technology (BPTP) west java, we can make the dimensions of its business model.

The following is based on dimension fact needs its strategic information:

a. Table dim_waktu
b. Table dim_wilayah
c. Table dim_komoditas
d. Table dim_hama
e. Table dim_tanam_panen
f. Table dim_lahan

The following is based on fact table needs its strategic information:

a. Table fact_luas_hasil_panen
b. Table fact_jumlah_kerusakan
c. Table fact_ prod_unggul
d. Table fact_ lq
e. Table fact_jumlah_jenis_lahan
f. Table fact_jumlah_jenis_ternak

3.5 Data Staging

At this stage processes will be implemented etl or commonly called Extract, Transform, and Load. ETL process for each dimension table and fact table at an Institution for agricultural technology (BPTP) west java. For details on the process of etl can be seen on Figure 2.1

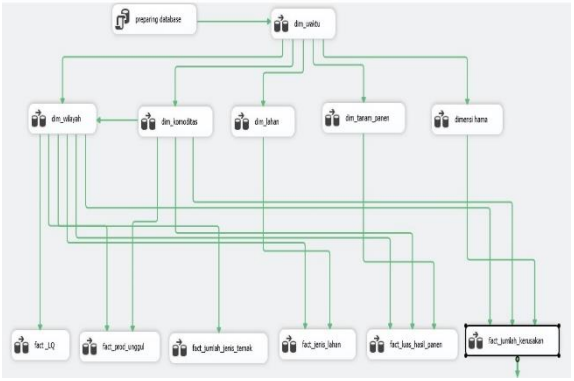


Figure 2.1 Proses ETL

At Figure 2.1 Is control flow of any process etl his. Any figure above icons on having the function of own. Are presented in the name of an icon Figure 2.2

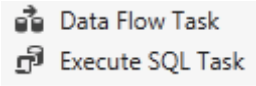


Figure 2.2 Ikon proses etl

On the data flow of a task serves to mengfigurekan a groove of existing dataWhile execute sql task serves to do execution query sql his usual laid the first time before call data flow of a task. Here is the explanation for some task existing etl of every process existing in Figure 2.1

1. Preparing SQL

This is a task execute sql task. An order that in use in this task is a drop and create all the table, view, and function, procedure which will be in use in database data warehouse

2. Dim Waktu

This task flow represents data task. Any flow data to details that were on this task can be in look at Figure 2.3

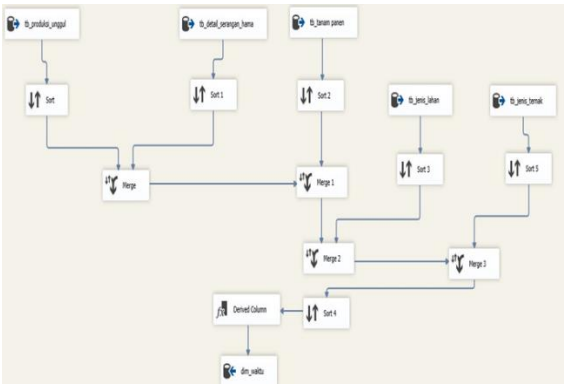


Figure 2.3 data flow dim_waktu

3. Fact_jumlah_kerusakan

This task flow represents data task. Any flow data to details that were on this task can be in look at Figure 2.4

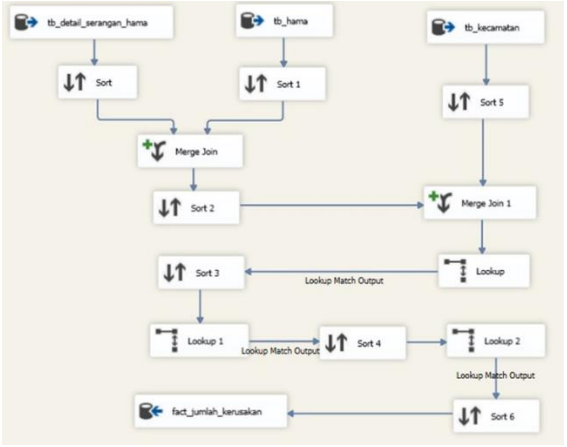
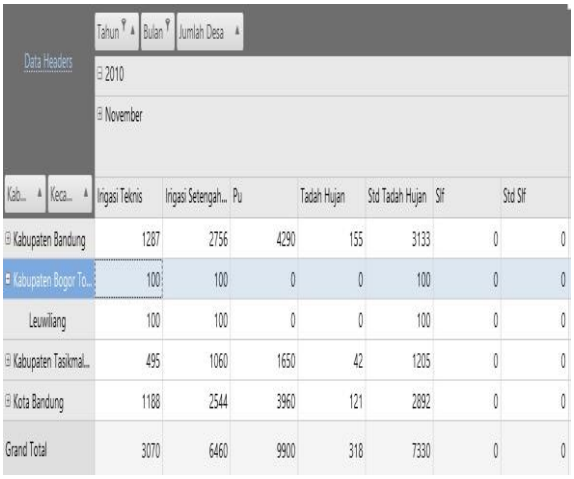


Figure 2.4 Fact Jumlah kerusakan

3.6 OLAP & Reporting Tools

OLAP represent the data with a cube multi-dimensional that easily understandable. OLAP also provide an answer to queries analytic to the data multidimensional. In this research, OLAP method used is a pivoting method. This method allows users to rotate the cube into a form of new or the size of values can say that it can display different in the layout table. It is very needed parties bptp west java to help existing data analysis .

OLAP following the result of application of data warehouse made:



Kab...	Keca...	Irigasi Teknis	Irigasi Setengah...	Pu	Tadah Hujan	Std Tadah Hujan	Slf	Std Slf
⊟ Kabupaten Bandung		1287	2756	4290	155	3133	0	0
⊟ Kabupaten Bogor To...		100	100	0	0	100	0	0
	Leuwiliang	100	100	0	0	100	0	0
⊟ Kabupaten Tasikmal...		495	1060	1650	42	1205	0	0
⊟ Kota Bandung		1188	2544	3960	121	2892	0	0
Grand Total		3070	6460	9900	318	7330	0	0

Figure 2.5 Olap luas hasil panen

Data Headers	Tahun	Kabupaten	Bulan
	2009		
	Kabupaten Bogor		
Ke...	Ringan	Sedang	Berat	Parah	Total Kerusakan
WBC	506	33	5	0	544
Walang Sangit	12	0	0	0	12
Tungro	138	50	26	0	214
Tikus	1371	639	10	0	2020
Siput Murbai	324	48	6	0	378
PBPP	60	60	0	0	120
HPP	125	47	40	0	212
BLB	94	32	0	0	126
BLAB	1184	659	45	0	1888
Grand Total	3814	1568	132	0	5514

Figure 2.6 Olap Jumlah Kerusakan berdasarkan serangan hama

Jumlah	Tahun	Nama Bulan			
	2009				2009 Total
N... Na... N...	Januari	Februari	Oktober	Desember	
Ayam Kampung Total			2141		2141
Kabupaten Subang...			2141		2141
Cisalak			2141		2141
Subang					
Bawang Daun					
Bawang Merah				2178	2178

Figure 2.7 Olap Jumlah produksi unggul pangan, ternak, buah sayur

Luas Panen	Tahun	Bulan	Luas Tanam		
	2009	2010	2011	2015	Grand Total
K... Ka... Ke...					
Bawang Daun Total		14651		2424	17075
Kabupaten Kuninga...		14651		2424	17075
Babakan		1684		2424	4108
Gebang		2859			2859
Karangsembung		2014			2014
Karangwareng		3307			3307
Lemahabang		4787			4787
Bawang Merah		15816			15816
Bayam		6121			6121

Figure 2.8 Olap Jumlah Hasil pangan, buah sayur

Data Headers	Tahun	Bulan						
	2009							
	Januari			Februari				
... ...	Ruminansia Besar	Ruminansia Kecil	Unggas	Ruminansia Besar	Ruminansia Kecil	Unggas		
Ayam Kampung Total	0	0	5000	0	0	5000		
Kabupaten Ba...								
Kabupaten G...	0	0	1000					
Kabupaten Su...	0	0	4000	0	0	5000		
Kabupaten Ta...								
Kota Cirebon								
Kota Depok								
Bebek				0	0	3000		
Domba	0	8000	0					
Itik				0	0	10000		
Kambing	0	8000	0					
Sapi Pedaging								
Sapi Perah	600	0	0					
Grand Total	600	16000	5000	0	0	18000		

Figure 2.9 Olap Jumlah Hasil Ternak

Lq	Tahun	Komoditas							
	2009				2010	2011	2015	Grand Total
N... ...	Durian	Duku	Buncis	2009 Total				
Kabupaten B...	43	209		648				648
Kabupaten Ci...					976			976
Kabupaten Ci...					48	780		828
Kabupaten In...				71	879			950
Kabupaten Ka...					591			591
Kabupaten K...					827		19	846
Kabupaten M...			17	17	961			978
Kabupaten P...				105	482			587
Kabupaten Su...					19			19
Kabupaten Su...					285			285
Kabupaten Su...					270			270
Grand Total	43	209	17	841	5338	780	19	6978

Figure 2.10 Olap LQ pangan, ternak, buah sayur

Data Headers	Tahun	Bulan	Jumlah Desa					
	2010							
	November							
Kab... Kec...	Irigasi Teknis	Irigasi Setengah...	Pu	Tadah Hujan	Std Tadah Hujan	Slf	Std Slf	
Kabupaten Bandung	1287	2756	4290	155	3133	0	0	
Kabupaten Bogor To...	100	100	0	0	100	0	0	
Leuwiliang	100	100	0	0	100	0	0	
Kabupaten Tasikmal...	495	1060	1650	42	1205	0	0	
Kota Bandung	1188	2544	3960	121	2892	0	0	
Grand Total	3070	6460	9900	318	7330	0	0	

Figure 2.11 Olap Jumlah Jenis lahan

4. CONCLUSION

Based on the analysis and testing of the data warehouse, it can be concluded that:

1. The application of Data Warehouse provides ease of access to the information needed and can

compress the data in the data warehouse in different ways so that it displays the appropriate information and fast.

2. Provides information integrated and contains bits on information relevant to the needs of parties bptp in a decision with decision makers.

3. The application of data warehouse help speed up the process that data collection and multi dimensional and concise presentation of information so as to maximize the quality of the decision made by top management.

Advice given to the development of a system of next, such as;

1. Need the addition of some of his menu items that can help to cultivate other data that is still in need

2. Added a unit of each in their own items of information that there are.

REFERENCES

[1] W. Inmon, Building the Data Warehouse, Canada: John Wiley & Sons, Inc., 2002.

[2] I. Sommerville, Software Engineering 9th Edition, Addison Wesley, 2010.

[3] R. Kimball, The Data Warehouse Toolkit Third Edition, John Wiley & Sons, Inc., 2013.

[4] M. S. Joseph M. Hellersterin, Architecture of a Database System, Washington: NOW the essence of knowledge, 2007.

[5] Ralph Kimball and Margy Ross, The Data Warehouse Toolkit, Redmond: Willey, 2002.

[6] P. Lane, Oracle9i Data Warehousing Guide Release 2 (9.2), Redwood: Oracle Corporation, 2002.

[7] D. Browning and J. Mundy, "Data Warehouse Design Considerations," MicrosoftSQL2000 TechnicalArticles - MSDN.com, 2001.

[8] A. Kadir, Pengenalan Sistem Informasi, Yogyakarta: Penerbit Andi Yogyakarta, 2003.

[9] M. Y. Pusadan, Rancang Bangun Data Warehouse, Yogyakarta: Graha Ilmu, 2013.

[10] P. Ponniah, Data Warehousing Fundamentals. A Comprehensif Guide For IT Professionals, John Wiley & Sons, Inc., 2001.

[11] M. Golfarelli and S. Rizzi, Data Warehouse Design: Modern Principles and Methodologies, The McGraw-Hill Companies, 2009.

[12] W. H. Inmon, Building the Data Warehouse Third Edition, Canada: John Wiley & Sons, Inc., 2002.

[13] S. Benett, S. McRobb and R. Farmer, Object-Oriented Systems Analysis And Design Using UML, McGraw Hill Higher Education, 2005.