

Business Intelligence for Construction Company Acknowledgement Reporting System

Abba Suganda Girsang
Computer Science Department, BINUS
Graduate Program-Master of Computer
Science
Bina Nusantara University
Jakarta, Indonesia 11480
agirsang@binus.edu

Sani Muhamad Isa
Computer Science Department, BINUS
Graduate Program-Master of Computer
Science
Bina Nusantara University
Jakarta, Indonesia 11480
sani.m.isa@binus.ac.id

Herry Saputra
Computer Science Department, BINUS
Graduate Program-Master of Computer
Science
Bina Nusantara University
Jakarta, Indonesia 11480
herry.saputra@binus.ac.id

M. Apriadin Nuriawan
Computer Science Department, BINUS
Graduate Program-Master of Computer
Science
Bina Nusantara University
Jakarta, Indonesia 11480
m.nuriawan@binus.ac.id

Reginald Putra Ghozali
Computer Science Department, BINUS
Graduate Program-Master of Computer
Science
Bina Nusantara University
Jakarta, Indonesia 11480
reginald.ghozali@binus.ac.id

Emil Robert Kaburuan
Information Systems Management
Department, BINUS Graduate
Program-Master of Information
Systems Management
Bina Nusantara University
Jakarta, Indonesia 11480
emil.kaburuan@binus.edu

Abstract—As the one of Indonesian association of construction service companies, the association ensure to recommend and record every construction service company (their members) to works according to valid Indonesian law and constitution. The problem in this association, as the job is to give detailed reports to government, investor either domestic or international, and every institutions that need information such as market needs, resource developments, and improvement of construction technologies, is to create queries for every data from tables in MySQL to create reports quickly, efficiently, and automatic. Example of reporting needs are: In what month construction project done at the most within 1 year; Project duration, project cost, and experience at the most for each construction company; or amount of projects based on source of funds (APBN, APBD, LOAN PEMERINTAH, BUMN, or SWASTA) for each construction companies. Recommended solution is using Data Warehouse where it has structured storage and can analyze data with OLAP that has capability to process and visualize data at high speed. To construct the Data Warehouse, the design used Kimball Method so this association company will have capability to generate report according to every instances' needs.

Keywords—Data Warehouse, kimball, pentaho, report, Qlik Sense

I. INTRODUCTION

Construction service companies in Indonesia have lots of potentials, where the investments, either governmental or private (service clients) increases every years. This also connected to growth rate of construction service companies in Indonesia from small, medium, and big scale companies. By observe market needs, resource developments, and improvement of construction technologies, they are expected to expand the company's capability building. by expanding the company, it will open chance of employments, improvement of public infrastructure, and chance for local constructions to compete with domestic and international markets. Therefore, construction service companies are part

of factors that affect Indonesia's development that equally important with education, economy, and culture developments.

These information are valuable assets to every companies, because knowledge has important role in guiding business analysts to take decisions[1]. And that is why, these assets needs to be secured and easy to be access in times they needed even in case of security breach[2]. But data available are so many that affect extraction of information to be difficult. Data Warehouse is a phenomenon of large collections of data processed into information from certain period of years and from that, to use data for reporting needs.

In this study case of association for construction service companies, the association work to recommend and register construction company in Indonesia, where this association already legitimated by government's ministry of public works and public housing since 2002, this association currently did not have reporting system that capable of quick, efficient, and automatic way to generate reports from database. Because of that limit, time needed for generate report is relatively too long. Currently, every reports have been done by export database from MySQL manually to Microsoft Excel file, then analyze the Excel file directly. Which is why Data Warehouse is needed to accommodate appropriate, quick, and easy to understand reporting.

By implement Data Warehouse and OLAP Analysis based on association needs, it is to be expected that analysis result capable to represent data from different reports where association might be need to take decision based from the data. The decision making need few historical data to be analyzed to support or object hypothesis and to take decision from the analysis [3].

Methodology used in this solution is using Kimball Method. Kimball Method is a system where transactional data collected from every sources using queries to analyze data. Few terms used for technologies for data transaction

and analysis are OLTP (Online Transactional Processing) and OLAP (Online Analytical Processing). OLTP and OLAP have important role in data transaction and analysis. This ideas from Ralph Kimball still used in current implementation of Data Warehouse [4].

II. RELATED WORKS

Business Intelligence (BI) is a capability or knowledge to process data gathered and stored in order to provide business with information that gives advantages in making decisions. BI requires data and tools to store historical data and process them for decision making. This will help business to expand itself by following decisions based from business company's data[5].

Data Warehouse is relation between databases constructed for query processes that aim for ease data reporting and analysis. Data Warehouse capable to help organization to analyze trend based on repository of organization's data in a period of time. Main functions of Data Warehouse are to facilitate organization in strategic planning for long term from data saved and to make quick decision that appropriate to the problem. Modeling dimension is used to implement Data Warehouse [6]. Modeling dimension is a data structure technique optimized and created to read, summary, and analyze numeric information such as value, sum, weight, etc. Few limitations to do Data Warehouse is required storage space and memory capacity to process these data into simplified views [7].

Star Schema is a model that still in use until now to represent multidimensional data that include 2 kinds of table, such as Fact table and Multidimensional table. these tables will undergo denormalization, to form a diagram of Entity Relationship Diagram (ERD) shaped like a star. this schema is drawn in relationship where fact table is placed in the middle of several Dimension tables [8]. Fact table is the table that store the facts, in this case, all historical data that will be used in analysis process. Then the dimension table is the table contains attributes that supports the facts[9].

Data Warehouse or usually called Online Analytical Processing (OLAP) is oriented to market compared to conventional OLTP. OLAP is a basic to create a Decision Support System (DSS) where this tool is used for create complex decision. OLAP can be integrated from many sources of data and capable to generate multidimensional reports that used for data analysis and decision making [10]. There are 3 variances from extension of OLAP, they are: Relational OLAP (ROLAP), Multidimensional OLAP (MOLAP), and Hybrid OLAP (HOLAP). ROLAP act as interface between Relational Database Management System with OLAP users. MOLAP store data from OLTP database in the cube. HOLAP combines both of them [11].

OLAP Cube is a feature from OLAP that connect every dimensions into fact table. This will help decision making and extraction of data in multidimensional way based on hierarchy of each dimensions for gaining information relevant based on selected level of details. Data Cube stored in cell with structure resemble a 3D spreadsheet [12]. Operations used in this method is Aggregation operations used to aggregate cell within the cube. Aggregate operations used are SUM, AVG, COUNT, MIN, and MAX. Example of

syntax used for aggregate operation is : AggFunction (CubeName, Measure) [By Dimension*].

ETL (Extraction, Transformation, and Loading) is an order of steps in data processing from database that involve process of extraction data. by maintain data quality, preserve data standards, and reform data in certain shape (Transformation) later to be presented (Loading) to Data Warehouse [13].

Pentaho is open source application to execute business analysis tasks, ETL, dashboard, and reporting. Pentaho ease these tasks of business analysis with solutions in application to create and develop business analysis project [14]. ETL process in Pentaho is using application called Pentaho Data Integration to make data extraction and transformation process easy. For design of OLAP Cube is using Mondrian with Pentaho Workbench Schema. Then to design report from ETL process is using Pentaho Report Designer.

Qlik Sense is an interactive application for analyze and visualize large amount of data to create presentation, dashboard, and reporting. This application is easy to use because of drag and drop components feature. Qlik Sense is available in shape of desktop application (Qlik Sense Desktop) and cloud hosting (Qlik Sense Cloud) that free to use. For company usage, they provide Qlik Sense Enterprise that removes limitations in free version [15].

III. PROPOSED METHOD

In this part, there are steps that explained later in this paper to solve problems in association company. Starting from data collection and analysis, designing star schema, ETL process, create a OLAP cube, then finally, design reports and dashboards.

Start with data collection, every data that support in the case study is analyzed to produce information about calculations and measurements that will be shown in OLAP analysis, report design, and indicators to be shown inside dashboard. These data will become reference to develop the Data Warehouse. Data sources that used to develop Data Warehouse are from operational database taken from Relational Database Management System (RDBMS) MySQL and tables shown in ERD of Figure 1.

With these data sources, a step of extraction data will be done to extract data from OLTP database into OLAP database. these data will be reform into filtered data that used for analysis. Data will be divided into fact and dimensions. Dimensions are used for making categories that group the data from fact. The fact in other side will have role as main source of information to be analyzed. The design of data is represented by star schema in Figure 2. the star schema is composed with 1 fact table and 5 dimension tables.

Based on fact and dimensions in star schema data design, a step of ETL will be done using Pentaho Data Integration application. The ETL Process start from extracting data from OLTP database by selecting tables that possess required data. Then, select columns of data that needed for OLAP analysis. Data will be given with surrogate key that represent the identity for each data. Last, these data will be stored in table inside OLAP database. Figure 3,4,5,6,7 in order are ETL steps to create dimension tables. Start from construction company, category, source of funds, place, and time.

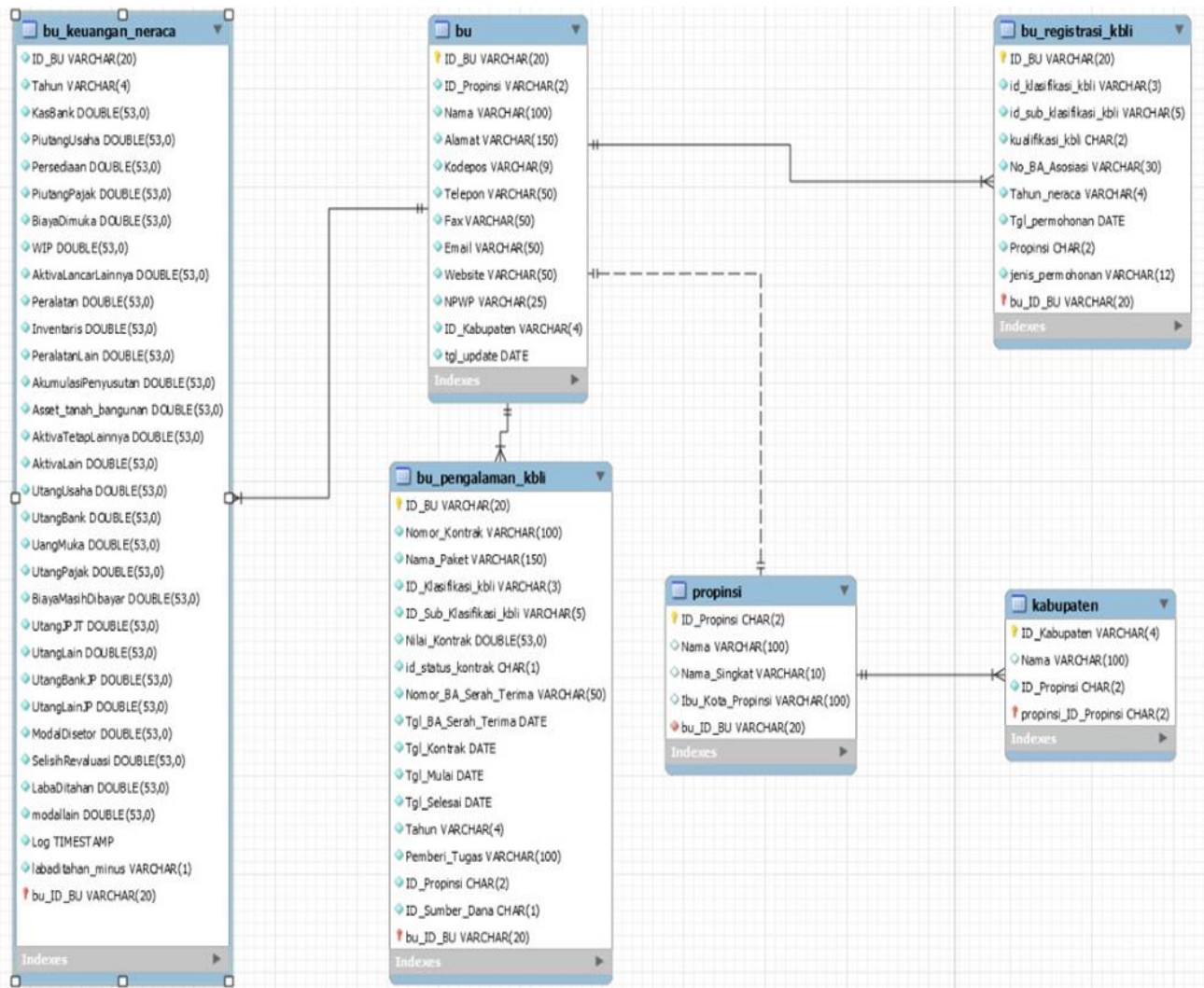


Fig. 1. ERD (Entity Relationship Diagram)

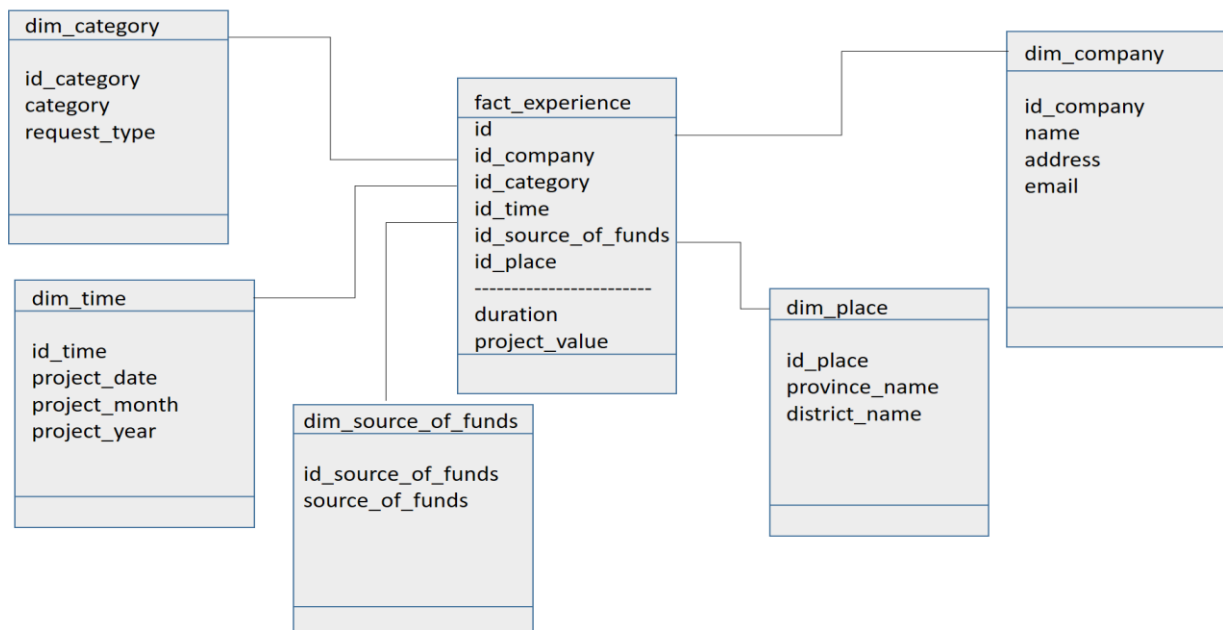


Fig. 2. Star Schema

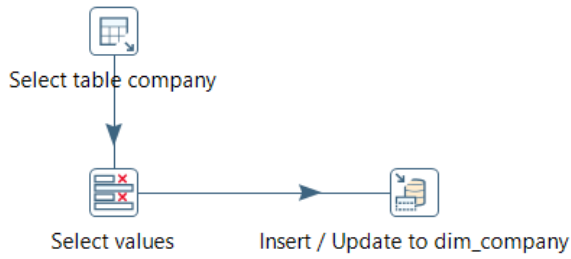


Fig. 3. Load Company Dimension

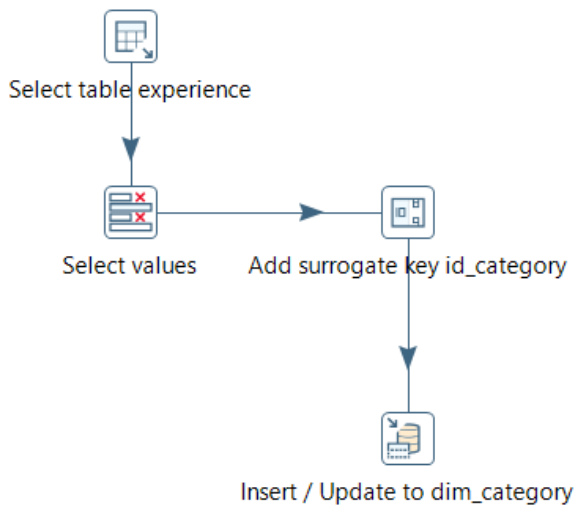


Fig. 4. Load Category Dimension

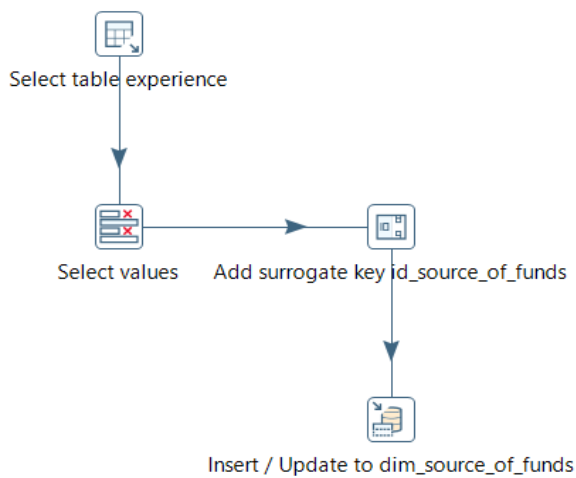


Fig. 5. Load Source of Funds Dimension

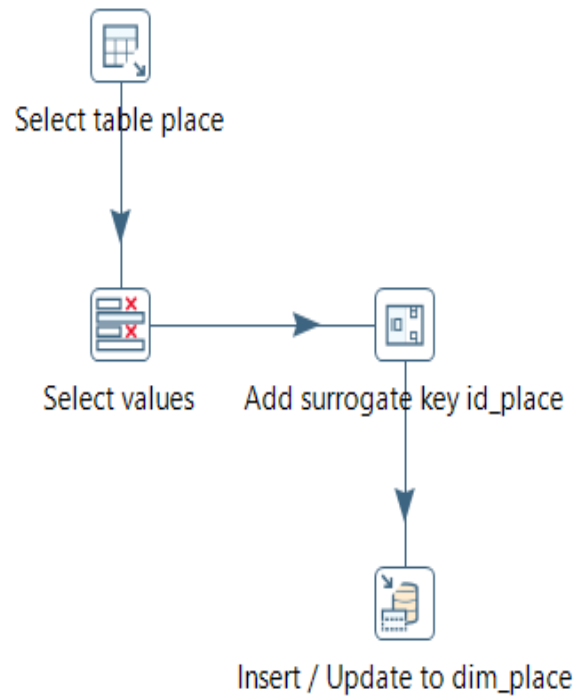


Fig. 6. Load Place Dimension

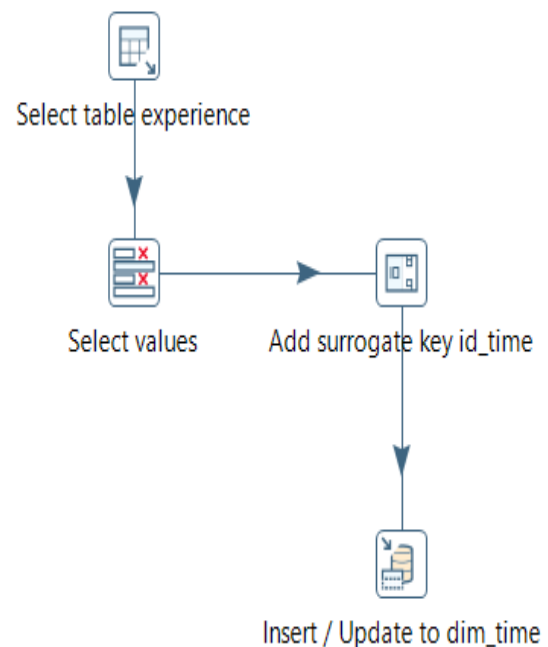


Fig. 7. Load Time Dimension

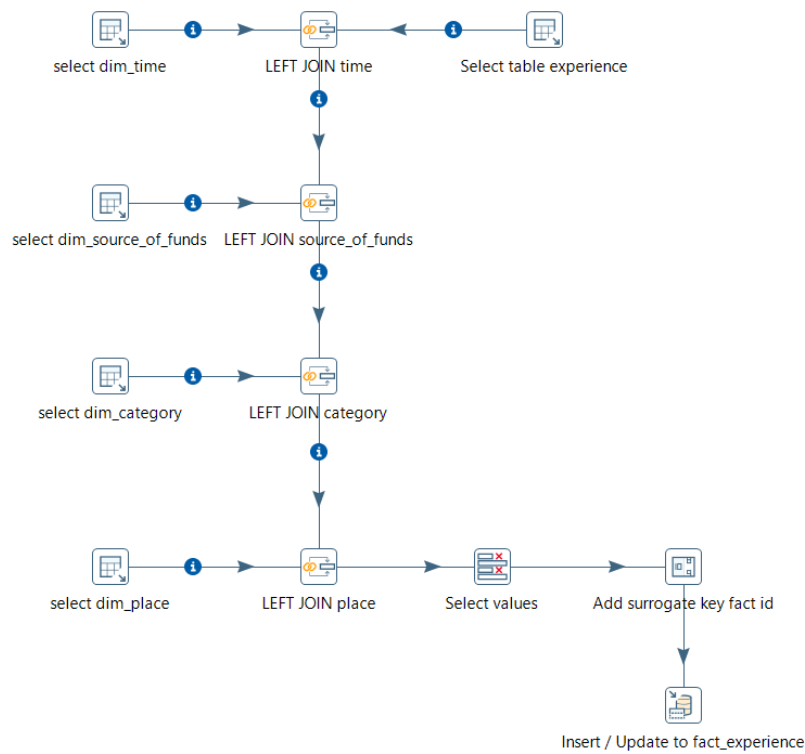


Fig. 8. Load Experience Fact Table

For fact table, each surrogate key from dimension tables will be taken for connecting the relevant data. These data relevant to surrogate key will be selected to become sample of fact table's data. Similar with dimension table, columns need to be chosen for taking core values needed in analysis. In this step, OLAP database is ready to use for dashboard and reporting. Figure 8 is an example of extraction process of experience table for construction companies' project (records of project done by each of company).

Next step is to form an OLAP Cube. the design for this case study is shown in Figure 9 that used Pentaho Workbench Schema application. Result from this OLAP Cube design is stored in form of xml file that also called as Mondrian file.

IV. DISCUSSION

We present an application of Data Warehouse to generate reports and dashboards to support decision making for construction company association management. The application intended to change manual and difficult way to generate report into more orderly way. This method used ETL process to manipulate data from operational database into star schema that easier to analyzed. The data will be represented with reports and dashboards to help management to review and take decision.

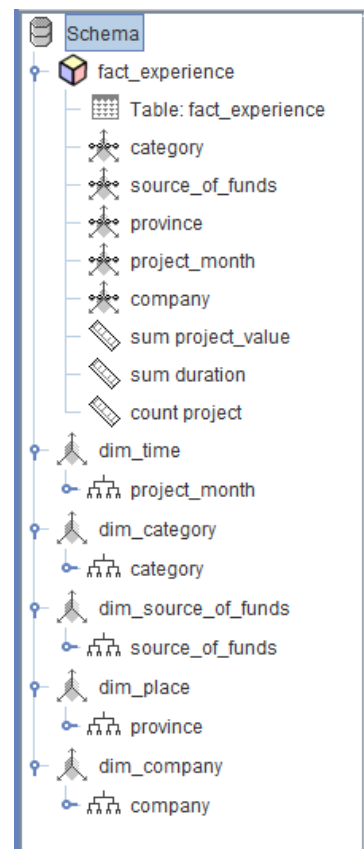


Fig. 9. OLAP Cube

V. RESULT

By using reporting and dashboard application, data from OLAP database can be processed into diagram or table that easy to understand. Reports are easy to understand by using diagrams that represent summary of data based on conditions given. If details of data are needed to be analyzed, dashboard will give more complete visualization.

In this case study, analysis will be done to find out summary of data with conditions:

- In what month construction projects have been done in a year.
- Accumulation of project execution duration for each construction companies.

- Accumulation of contract's value from contracts done by a construction company.
- Sum of projects based on source of funds for each construction companies.

Each of these conditions are represented by reports and dashboards. Figure 10 and 11 represent in what month construction projects have been done in a year. Figure 12 and 13 represent accumulation of project execution duration for each construction companies. Figure 14 and 15 represent accumulation of contract's value from contracts done by a construction company. Figure 16 and 17 represent sum of projects based on source of funds for each construction companies.

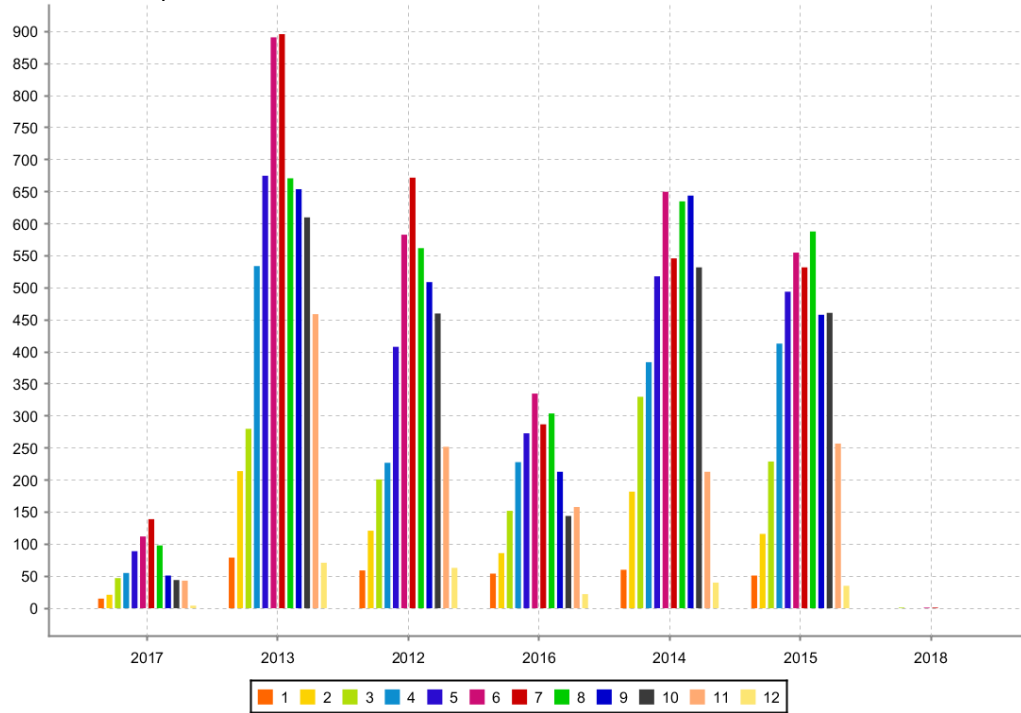


Fig. 10. Report of In What Month Construction Projects Have Been Done In a Year

1. In What Month Construction Projects Have Been Done In a Year

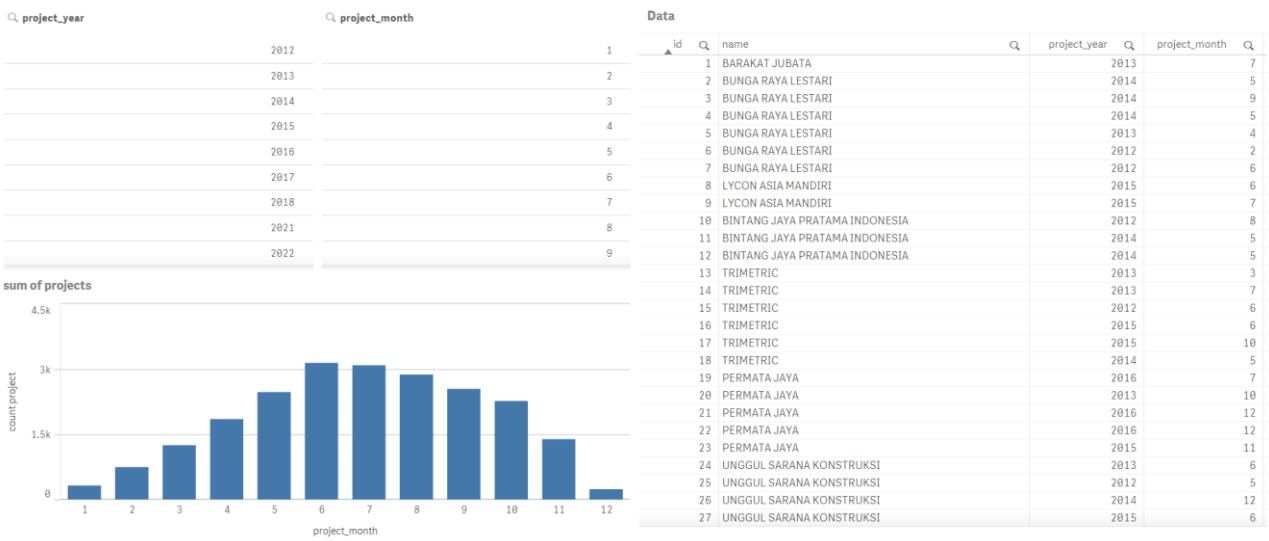
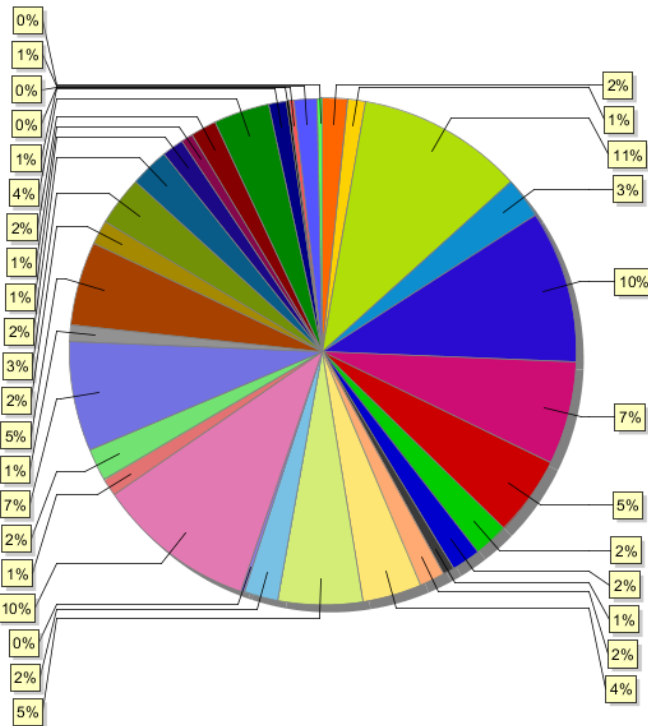


Fig. 11. Dashboard of In What Month Construction Projects Have Been Done In a Year



- Kalimantan Barat ● Nusa Tenggara Barat ● Jawa Timur ● Jawa Barat ● DKI Jakarta ● Sumatera Barat ● Jawa Tengah
- Sulawesi Tengah ● Lampung ● Nanggroe Aceh Darussalam ● Bengkulu ● Sulawesi Selatan ● Sulawesi Tenggara ● Riau ● Bali
- Papua ● Banten ● Kalimantan Timur ● Nusa Tenggara Timur ● DI Yogyakarta ● Sumatera Selatan ● Jambi ● Maluku Utara
- Kepulauan Riau ● Kepulauan Bangka Belitung ● Sumatera Utara ● Papua Barat ● Kalimantan Tengah ● Maluku ● Kalimantan Selatan
- Gorontalo ● Sulawesi Utara ● Sulawesi Barat

Fig. 12. Report of Accumulation of Project Execution Duration For Each Construction Companies

2. Accumulation of Project Execution Duration For Each Construction Companies

Data

id	Q	name	Q	project_date	Q	project_month	Q	project_year	Q	category	Q	duration (days)	Q
1		BARAKAT JUBATA		26		7		2013		K2		59	
2		BUNGA RAYA LESTARI		16		5		2014		B2		178	
3		BUNGA RAYA LESTARI		17		9		2014		B2		212	
4		BUNGA RAYA LESTARI		5		5		2014		B2		207	
5		BUNGA RAYA LESTARI		16		4		2013		B2		216	
6		BUNGA RAYA LESTARI		6		2		2012		B2		182	
7		BUNGA RAYA LESTARI		13		6		2012		B2		153	
8		LYCON ASIA MANDIRI		1		6		2015		M1		107	
9		LYCON ASIA MANDIRI		29		7		2015		M1		113	
10		BINTANG JAYA PRATAMA INDONESIA		30		8		2012		M2		123	
11		BINTANG JAYA PRATAMA INDONESIA		19		5		2014		M2		145	
12		BINTANG JAYA PRATAMA INDONESIA		19		5		2014		M2		153	

Q	name	Q	province_name	Q	project_year	Q	sum of duration
	YO FAMILY		Bali		2012		
	@LITE PERKASA		Banten		2013		
	2 PUTRA ABADI		Bengkulu		2014		
	3A GROUP		DI Yogyakarta		2015		
	3RAS ENGINEERING		DKI Jakarta		2016		
	4 PUTRA		Gorontalo		2017		
	4 SAUDARA		Jambi		2018		
	5 PUTRI		Jawa Barat		2021		
	7 BIKKU		Jawa Tengah		2022		

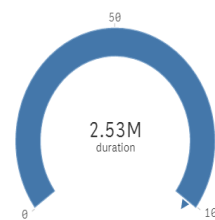


Fig. 13. Dashboard of Accumulation of Project Execution Duration For Each Construction Companies

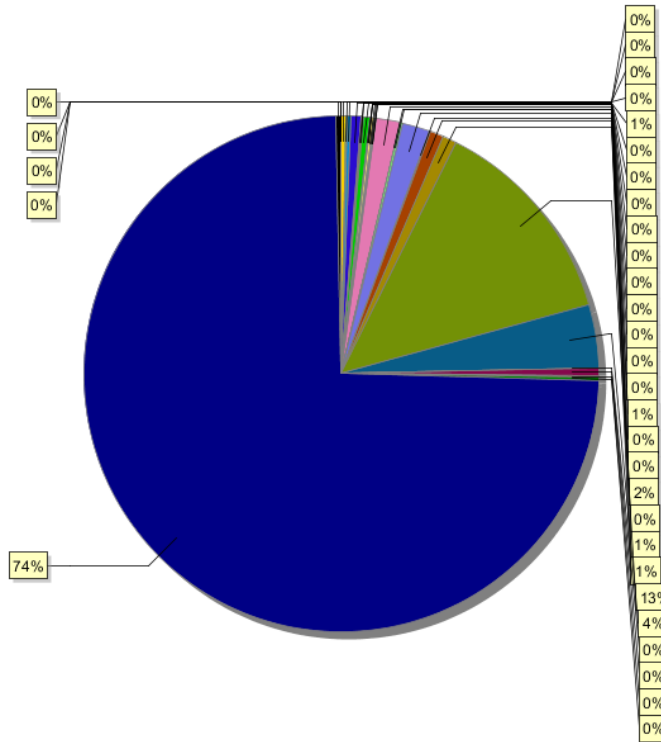


Fig. 14. Report of Accumulation of Contract's Value From Contracts Done By a Construction Company

3. Accumulation of Contract's Value From Contracts Done By a Construction Company

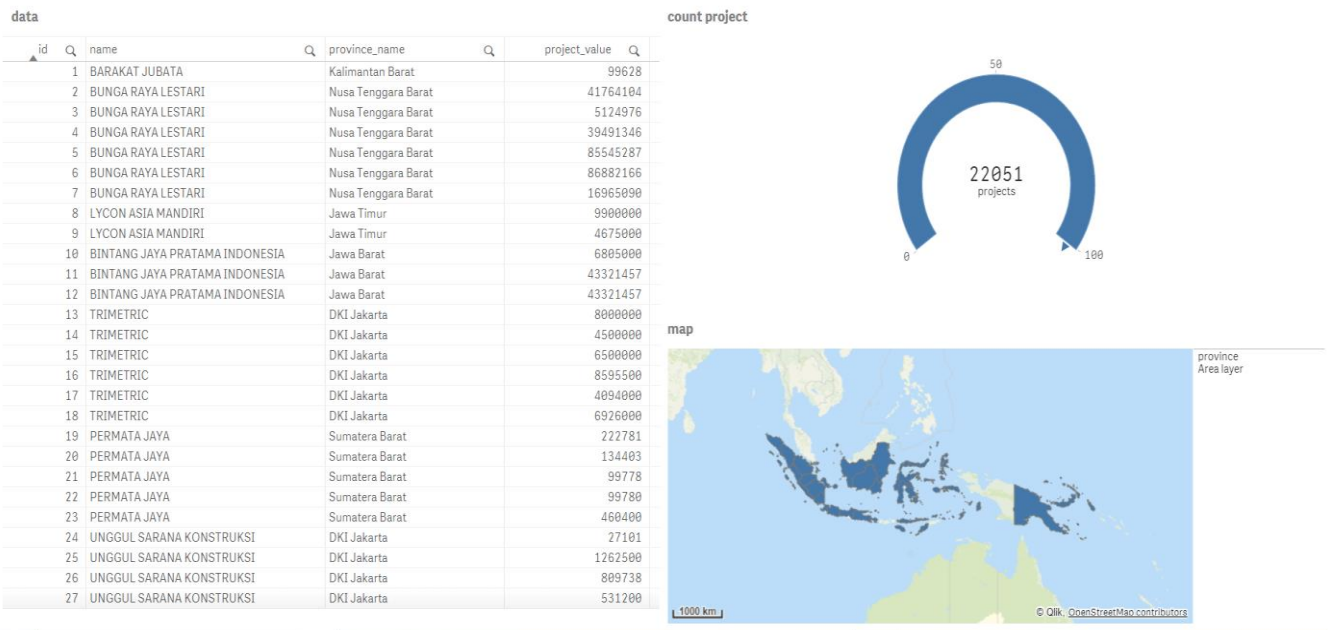


Fig. 15. Dashboard of Accumulation of Contract's Value From Contracts Done By a Construction Company

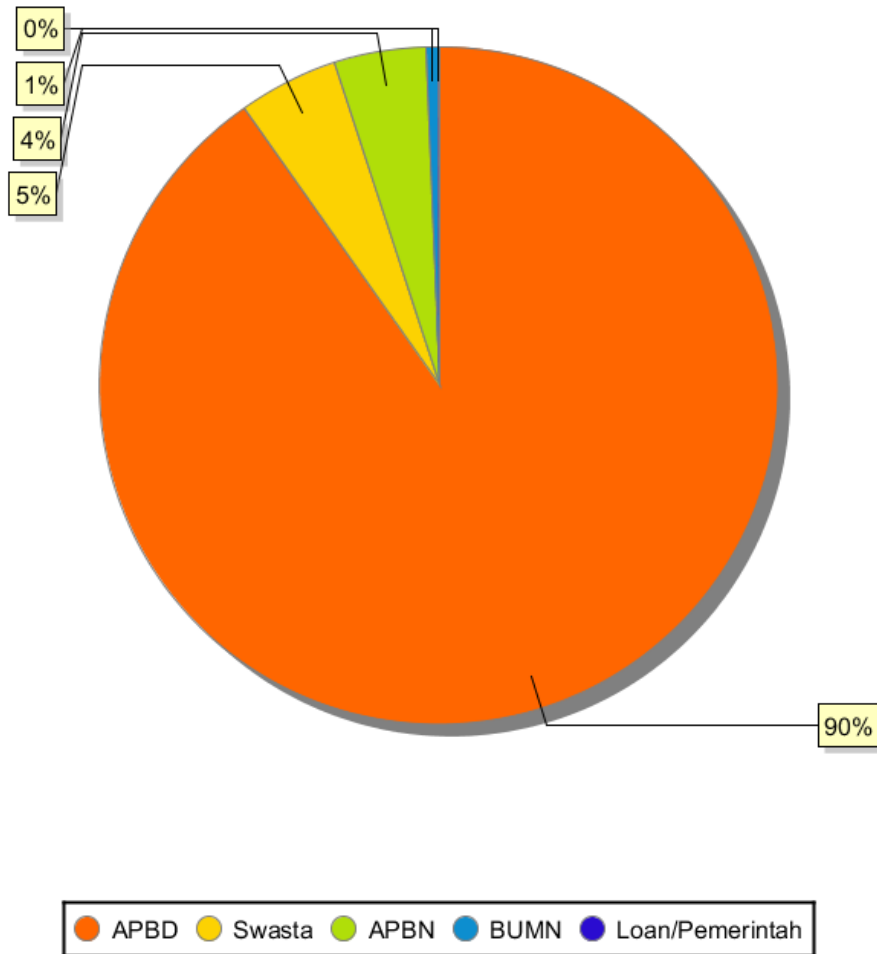


Fig. 16. Report of Sum of Projects Based On Source of Funds For Each Construction Companies

4. Sum of Projects Based On Source of Funds For Each Construction Companies

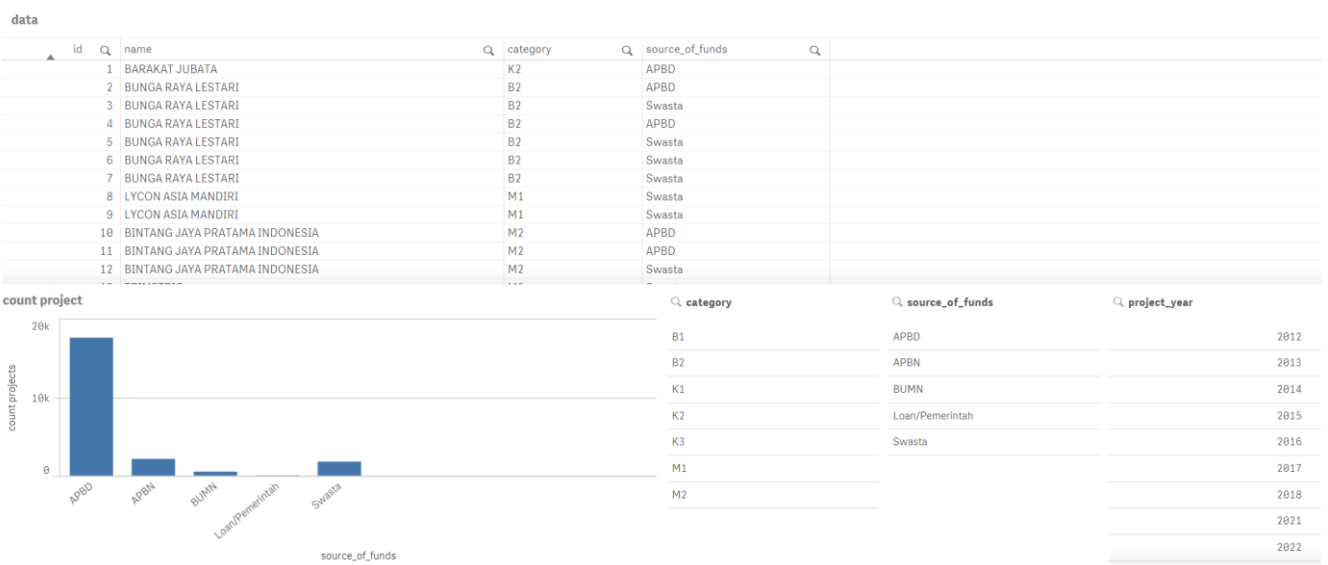


Fig. 17. Dashboard of Sum of Projects Based On Source of Funds For Each Construction Companies

VI. CONCLUSIONS

By using dashboard and report, data analysis process become much easier to done. Data Warehouse is able to give capability to process fact data with multiple dimensions in OLAP database. Result from data processing can be presented with many kinds of diagrams, tables, and dashboards. Combination of reports and dashboards made a easy way for business analysts to take appropriate and quick decision to solve the problems.

We noted some limitations while doing the study case. Dashboard used in this case is not comply with real-time operational data. Qlik Sense Desktop needs to receive data from operational database source then create a data schema based on data at the time it received. This means every new data added in operational database needs to be loaded periodically in Qlik Sense Desktop, adding time costs and efforts for creating real-time reports and analysis.

Report designer used in this case can only use single query to generate a report. Query used in the report that contains join function will takes more time to process the query for generate report. There also a moment that extra queries are needed in order to create different diagrams or tables that relevant with the case in a single report.

For future works, we plan to create decision support system that helps the managers and high-level managements to take decisions from these reports and dashboards. Also, we plan to investigate different application that supports real-time operational data to create real-time business intelligence.

REFERENCES

- [1] D. J. Pauleen and W. Y. Wang, "Does big data mean big knowledge? KM perspectives on big data and analytics," *Journal of Knowledge Management*, vol. 21, no. 1, pp. 1-6, 2017.
- [2] I. A. T. Hashem, I. Yaqoob, N. B. Anuar, S. Mokhtar, A. Gani and S. U. Khan, "The rise of "big data" on cloud computing: Review and open research issues," *Information Systems*, vol. 47, pp. 98-115, 2015.
- [3] J. Kamki, *Digital Analytics: Data Driven Decision Making in Digital World*, Notion Press, 2017.
- [4] R. Kimball and M. Ross, *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling 2002*, US: John Wiley & Sons, 2002.
- [5] D. Larson and V. Chang, "A review and future direction of agile, business intelligence, analytics and data science," *International Journal of Information Management*, vol. 36, no. 5, pp. 700-710, 2016.
- [6] A. S. Girsang, D. Satya, S. M. Isa, S. Al Fariz, B. Susilo and D. Ramdani, "Decision support system using data warehouse for hotel reservation system," in *Sustainable Information Engineering and Technology (SIET), 2017 International Conference*, 2017.
- [7] R. Goswami, D. K. Bhattacharyya, M. Dutta and J. K. Kalita, "Approaches and issues in view selection for materialising in data warehouse," *International Journal of Business Information Systems*, vol. 21, no. 1, pp. 17-47, 2016.
- [8] R. Hart and A. M. H. Kuo, "Meeting Health Care Research Needs in a Kimball Integrated Data Warehouse," in *Data Science and Advanced Analytics (DSAA), 2016 IEEE International Conference*, 2016.

- [9] E. Sidi, M. El Merouani and A. A. El Amin, "The impact of partitioned fact tables and bitmap index on data warehouse performance," *International Journal of Computer Applications (973-93-80891-16-1)*, 2016.
- [10] A. S. Girsang, E. Andita, S. M. Isa, A. Purnama and F. Samuel, "Business intelligence for physical examination platform service reporting system," in *Sustainable Information Engineering and Technology (SIET), 2017 International Conference*, 2017.
- [11] V. N. Gudivada, M. T. Irfan, E. Fathi and D. L. Rao, "Cognitive analytics: Going beyond big data analytics and machine learning," in *Handbook of Statistics*, vol. 35, Elsevier, 2016, pp. 169-205.
- [12] Z. Wang, Y. Chu, K. L. Tan, D. Agrawal and A. E. Abbadi, "HaCube: extending MapReduce for efficient OLAP cube materialization and view maintenance," in *International Conference on Database Systems for Advanced Applications*, 2016.
- [13] R. Kimball and J. Caserta, *The data warehouse ETL toolkit: practical techniques for extracting, Cleaning, Conforming, and Delivering Data*, 2004.
- [14] V. Vargas, A. Syed, A. Mohammad and M. N. Halgamuge, "Pentaho and Jaspersoft: a comparative study of business intelligence open source tools processing big data to evaluate performances," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 14569, pp. 1-10, 2016.
- [15] M. S. Gounder, V. V. Iyer and A. Al Mazyad, "A survey on business intelligence tools for university dashboard development," in *Big Data and Smart City (ICBDSC), 2016 3rd MEC International Conference*, 2016.