

Linda Sutriani¹

Master of Information Technology,
Universitas Teknologi Digital Indonesia,
Yogyakarta, Indonesia
email: lindasutriani19@gmail.com

Indra Yatini Buryadi

Informatics,
Universitas Teknologi Digital Indonesia,
Yogyakarta, Indonesia
email: indrayatini@utdi.ac.id

Hera Wasiati

Retail Management,
Universitas Teknologi Digital Indonesia,
Yogyakarta, Indonesia
email: hera@utdi.ac.id

Sudarmanto

Application Soft. Engineering,
Universitas Teknologi Digital Indonesia,
Yogyakarta, Indonesia
email: darmanto@utdi.ac.id

ETL Data Warehouse On Performance Measurement

Measuring the organizational performance correctly will turn out accurate data which can be analyzed to provide important information for management in making decisions to improve company performance. This is because in the past, the performance appraisal was only based on financial measures. However, the editor-in-chief needs more than just financial indicators to improve performance. The Balanced Scorecard measures four organizational dimensions, such as customer, financial, internal business, and learning and growth. Although originally designed for the private sector, many public organizations apply it as modifications to suit their needs. The Balanced Scorecard is an effective method for measuring and achieving organizational performance. This research involves ETL in the Balanced Scorecard process to combine data from various sources into a large central repository called a data warehouse.

KeyWords: ETL, Data Warehouse, Performance, Balanced Scorecard

This Article was:

submitted: 19-06-23
accepted: 10-07-23
publish on: 14-07-23

How to Cite:

Sutriani, L., et al, "ETL Data Warehouse On Performance Measurement", Journal of Intelligent Software Systems, Vol.2, No.1, 2023, pp.12-17, [10.26798/jiss.v2i1.928](https://doi.org/10.26798/jiss.v2i1.928)

1 Introduction

ETL is a set of processes for collecting, filtering, processing, and combining data which have to be passed in forming a data warehouse. The ETL process consists of Extracting, Transforming, and Loading processes. Extracting is the process of selecting and retrieving data from a data set of a company. Transforming is the process of cleaning and changing the data structure from its original to a form that suits the needs of the data warehouse. Load Extracting is the last process that functions to enter data into the Data Warehouse[1] The use of the warehouse data in companies aims to assist the process of storing and presenting data so that the companies can record all structured transactions. The application of the data warehouse is a place for data synchronization in which data structure equalization occurs so that the transaction data can be received in the data warehouse. The synchronization process in the data warehouse is known as the ETL (extract, transform, and load) process which bridges the transaction data with the data warehouse storage media.[2]. Radar Sampit editors were born from a thorough selection process and went through a training process before plunging into the working world. Not only they were created to be able to process the raw news, they were also developed to be able to create new innovations that could increase high selling power. The balanced scorecard is a strategic management system that translates the organization's vision and strategy into operational objectives and measures. These operational objectives and measures are then stated in four objectives, specifically financial

objectives, customers, internal business processes, and learning and growth. The Balanced Scorecard is a collection of integrated performance measures derived from corporate strategy that supports the overall corporate strategy. The Balanced Scorecard is a strategic management system that translates the organization's vision and strategy into operational objectives and measures[3].

2 Literature Review

In the research[4] conducted a study on measuring company performance using the balanced scorecard method to determine the employees and listeners' level of satisfaction and provide suggestions for strategies to improve the company performance. It applied the Balanced Scorecard research methodology. The data used in this study were the company's financial report for 2009-2011 and questionnaires distributed to 100 Radio Gress listeners in Pekanbaru and 20 Radio Gress employees. The results of the study are as follows: the number determination of the listeners is based on the Slovin formula by the results of a financial perspective with a priority percentage of 50%. The internal business process perspective has a priority percentage of 25%. The listener perspective has a priority percentage of 15%. The growth and learning perspective has a priority percentage of 10%. In research[3] conducted a study on performance measurement using the balanced scorecard and integrated performance measurement system (IPMS). The company performance measurement is carried out using the Balanced Scorecard and the Integrated Performance Measurement System (IPMS). The performance measurement using the Balanced Scorecard is carried out by looking at four perspectives, namely the Financial, Customer, Internal Business Process, and Learning and Growth perspectives. Meanwhile the IPMS are carried out by observing the needs of the company's stakeholders. The data taken were in the form of company financial report data, income statements, balance sheets, questionnaire data and interview data. The results of the study are as follows. Based on the analysis results of the Balanced Scorecard, benchmarks which experience poor performance include Working Capital Turn Over (WTCO) with an average of -19.80; Total Debt to Equity Ratio (TDER): 175.13%; measurement of the ratio of growth rates and average demand of: 6.5%. The results of the analysis using

¹Corresponding Author.

IPMS obtained 30 KPIs, and the company's main priorities in the order: customer stakeholders, investment stakeholders, workforce stakeholders, supplier stakeholders, and community stakeholders. Thus, from the results of KPI weighting, the company's performance which has been good is customer stakeholders because it pays attention more to the customer satisfaction and comfort. In research[5] conducted a study on extract, transform, and load modules for Indonesian agricultural commodity warehouse data using talend. This study applies the ETL (Extract, Transform, Load) method. The data obtained from the Ministry of Agriculture's website is data which still requires a transformation process from raw data so that it becomes data which conforms to the data warehouse format. This is because the data that can be loaded into the data warehouse is only structured data and suitable to the format. Therefore, a data transformation process is needed to tidy it up into a format that matches the data format in the intended data warehouse. The results of this study thrived in building an ETL data warehouse module to transform Indonesian agricultural commodity yield data so that it can be integrated into the data warehouse. This research produces five job flow transformations to build four dimension tables and one fact table which are used for the needs of building a data warehouse. The transformation has been successfully carried out for the five jobs and the value generated by the transformation is in accordance with the initial value in the download file for Indonesian agricultural commodities.

3 Theoretical Review

3.1 ETL. ETL (Extract, Transform, Load) is the basic system of data warehouse. A good ETL design of a data source extraction system, prioritizing data quality and consistent standards, data from separate sources is appropriate, so that it can be integrated to provide a data format to be represented. The ETL system is a backbone activity which is not visible to the last users of the data warehouse. ETL fulfills 70 percent of the resources needed in the implementation and maintenance of the Data Warehouse. Extract, Transform and Load are also a collection of data preparation processes from OLTP (Online Transaction Process). ETL is the data processing phase from the data source to the data warehouse. The purpose of ETL is to collect, filter, process and combine relevant data from various sources to be stored in a data warehouse[6]. Rifaieh and Benharkat (2002) have defined a model that contains various types of mapping expressions. They apply this model to activate the ETL tool. In their approach, queries are used to achieve the warehousing process. Queries will be used to represent the mapping between source and target data; thus, enabling the DBMS to play an expanded role as a data transformation engine and also as a data store. This approach enables complete interaction between mapping metadata and warehousing tools. Additionally, it addresses the efficiency of query-based data warehousing ETL tools without suggesting any graphical models. It describes a query generator for reusable and more efficient processing of warehouse (DW) data[7].

3.2 Data Warehouse. Data warehouse is a database which has special structure for querying and analyzing. A data warehouse typically contains data which represents the business history of a company. The data is collected from various existing applications, and then restructured to be stored in a Relational Database Management System (RDBMS). The data warehouse is the heart and foundation of all EIS processes because it has one integrated data source with the right level of granularity. Data warehouses allow users to examine and analyze historical data in some form, but warehouse data cannot make decisions. The advantages obtained by using warehouse data are[8]:

- Data is well-organized for analysis queries and as material for transaction processing.

- Differences between homogeneous data structures on several separate sources can be overcome.
- Rules for data transformation are applied to validate and consolidate data when the data is moved from OLTP database to data warehouse.
- Security and performance issues can be solved without changing production systems.

3.3 Balanced Scorecard. The definition of performance appraisal (performance measurement) according is as a periodic determinant of the operational effectiveness of an organization, parts of the organization, and employees based on predetermined goals, standards and criteria[2].

The Balanced Scorecard is a strategic management system or to be more precise called a "Strategic based responsibility accounting system" which describes the mission and strategy of an organization into operational objectives and benchmarks for the company's performance. The Balanced Scorecard consists of two words, namely balanced and scorecard. Scorecard means the card score; it means the score card which will be used to plan the score which is recognized in the future. While balanced means unbiased, its meaning is to measure the performance of a person or organization measured in a balanced way from two perspectives, financial and non-financial, short term and long term, internal and external. Balanced Scorecard is a strategic management system which describes the vision and strategy of a company into operational objectives and benchmarks are developed for each of 4 (four) perspectives, namely: financial perspective, customer perspective, business process perspective and learning and growth perspective[9].

Through the Balanced Scorecard, company managers will be able to measure how their business units create value recently while taking into account future interests. The Balanced Scorecard makes it possible to measure what has been invested in the development of human resources, systems and procedures, for future performance improvement. Throughout the same method, it can also be assessed what has been fostered in intangible assets such as brands and customer loyalty[10].

4 Research Methodology

The literature study is carried out by studying previous research related to library data collection methods, reading and taking notes, and managing source research materials such as books, journal publications, papers, theses, dissertations. The apparatus used in this study are:

- Laptop with Intel i5-1135G7 2.40GHz processor specifications, with 8 Gb of RAM, Windows 10 64bit operating system.
- Microsoft Excel 2010 as storage media for feature extraction results.

While the material used in this study is the data of the Radar Sampit editor. The data collection was taken from the Radar Sampit Editor, until the research period.

5 The Purposed ETL Schema

5.1 Analysis and Design.

- (1) ELT DATA WAREHOUSE Design Analysis. The initial steps in system design are the extracting. The extract process is the first stage of the ETL system. Extract is the process of selecting and retrieving data from one or several sources (e.g. a database), then accessing the retrieved data. The second step is transform, one of the data has been taken through the extract process, then the data is cleaned by removing unnecessary data (e.g. anomaly data). Then change



Fig. 1 ETL Data Warehouse System Design

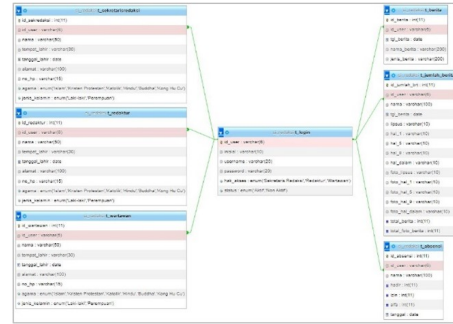


Fig. 2 MySQL Editor Database

the data from its original form into a form that suits the needs. The final step of the ETL system is load. During the load process, data storage occurs in the data warehouse and displays data to the application.

- (2) The Balanced Scorecard Design Analysis. This analysis is carried out by looking at the perspectives of Finance, Customers, Internal Business Processes, and also Learning and Growth Processes.

- (c) Load Then on Load to MySql Database, the first thing to do is to establish a connection to the database server. After creating a new connection identify the target table to apply.

5.2 Implementation. The implementation process consists of 2 stages, i.e.:

- (1) The Implementation of Warehouse ETL data
 - (a) Extract data retrieval from editorial performance in the form of a file in the form of Microsoft Excel shown in 1



Fig. 3 Weekly Report Output

Table 1 Imported Data Format of News Journalists/Educators

Date	TOTAL		POIN		ALPA (PP)	INDEKS TP
	NEWS	PHOTO	NEWS	PHOTO		
1	1	1	6	5		11
2	2	1	10	5		15
3	0	0	0	0		0
4	2	0	4	0		4
...
25	1	0	4	0		4
26	1	0	4	0		4
27	1	0	4	0		4
28	1	0	4	0		4
29	1	0	4	0		4
30	0	0	0	0		0
31	1	0	4	0		4
SUM	28	2	110	10	0	120

The weekly reports contain multidimensional data, explaining data journalists get a huge number of news per week. If the news is below the GPA, journalists do not get 1 week off.

- (b) Transform. One of the data has been taken through the extract process, then cleaning the data is carried out by removing unnecessary data which consists of 7 tables as follows: table_login, table_secretary, table_journalist, table_editor, table_absence, table_news, table_number_news.

- (2) The Implementation of the Balanced Scorecard data The following is an analysis of the performance of the Radar Sampit Editor Performance based on the Balanced Scorecard method.

- (a) Financial Perspective
Financial performance is measuring the company's performance in obtaining profits and market value. Financial size is manifested in profitability, growth and sales value of newspapers. The average annual newspaper sales are as shown in table 2:

Table 2 Financial Perspective

Year	Achievement
2018	80%
2019	87%
2020	91%

Table 3 Newspaper Sales in One Year

Month	City		Newspaper Sales Total
	Within the City	Out of the City	
January 2020	43%	44%	87%
February 2020	42%	43%	85%
March 2020	44%	45%	89%
April 2020	43%	42%	85%
May 2020	47%	43%	90%
June 2020	47%	43%	90%
July 2020	49%	43%	92%
August 2020	49%	46%	95%
September 2020	49%	46%	95%
October 2020	48%	45%	93%
November 2020	48%	47%	95%
December 2020	49%	46%	95%
Average			91%

(b) Customer Perspective

There are two measurement groups in this perspective; one of them are the core measurement group which measures the market share, the rate of acquiring new customers, the ability to retain long-time customers, the level of customer profitability and the level of customer satisfaction. The second is the customer value proposition which describes the performance drivers (work triggers), it measures the value in which the company can deliver to its consumers. The value proposition describes the attributes presented by the company in the products or services sold to create customers' loyalty and satisfaction.

The Table 3 illustrates that monthly newspaper sales have increased and decreased. However, for newspaper sales data in 2020 there has been a raise in average newspaper sales.

(c) Internal Business Perspective

To increase the high selling power of newspapers, so viral news, political news, regional news and so on is presented on newspaper news. Moreover, a caricature presentation, real photos or high HD photos are added. Therefore the reading enthusiasts are very significant for newspaper sales within and out of the city.

Table 4 Newspaper Sales Each Year

Year	Newspaper Sales	City	
		Within the City	Out of the City
2018	80%	40%	40%
2019	87%	44%	44%
2020	91%	46%	46%

It is illustrated from the Table 4 that the newspapers sales within and out of the city have increased every year.

(d) Learning and Growth Perspective

Table 5 Editor Absence per Week

Date	Editor	Presence			Point
		H	I	A	
15/05/2018	IGN	7	0	0	14
15/05/2018	YIN	7	0	0	14
15/05/2018	GUS	7	0	0	14

The Table 5 shown the absence or presence of the editor significantly affects the reason for getting points.

Table 6 Journalist Absence Per Week

Date	Journalist	Presence			Point
		H	I	A	
15/05/2018	ANG	7	0	0	14
15/05/2018	DC	7	0	0	14
15/05/2018	RON	7	0	0	14

The table 6 shown the absence or presence of the journalists deeply influences the points.

Table 7 The Number of Editor/Journalist News

Editor/Journalist	News Per Day	News Photo	News Total Per Day
IGN	2	1	3
ANG	4	2	6

Table 7 shown that the more news they write per day, the more news they get. It is because the maximum news obtained by journalists and editors is 5 and at least 1 news per day.

Table 8 Editor/Journalist Total Point

Editor/Journalist	News Point	Photo Point	Index (TP)
IGN	2	1	3
ANG	7	2	6

The more news they write per day, the more points they get, as seen in Table 8.

Table 9 Performance in A Week

Journalist/Editor	The Number of News	GPA
IGN	3	0.42
ANG	14	1.28

The number of news received per week will determine the weekly holidays, if the GPA is medium or high, as shown in Table 9.

Table 10 Performance in A Month

Name/Initial	The Number of News	Absence Point	GPA	Performance Allowance
Journalist				
IGN	12	7	0.63	0
Editor				
ANG	43	7	1.66	0

The description from the Table 10 is to create the quality news points and photo points, taken from the results of journalists and editors' points. Consequently,

the Index (TP) for the GPA itself is taken from the Index (TP) / 30 = GPA. The performance allowance achieved if the journalists and editors get interesting news, it will be multiplied with 2500.

The results of learning and growth perspective evaluation in Radar Sampit Editorial are divided into three aspects, namely for the Absence aspect showing good results. This is supported by the activeness of employees, and assessment of the Number of News that falls into the good category with the points they get. The factors that supporting growth and learning perspective are the budget, employee competence, and human resources.

5.3 Tests and Evaluations. The result of the test is shown at Table 11.

6 Conclusions

The conclusions from this research as follows:

- Based on the warehouse data testing process described in Chapter 4, the implementation of the data warehouse which consists of making ETL modeling, the ETL process, has been successfully carried out.
- The implementation report results applied the diagrams to transform data into practical information forms can be successfully understood by users.
- From the implementation results of the ETL data warehouse of editorial performance and implementation of queries for analysis, there are seven indicators for the login_table, secretary_table, reporter_table, editor_table, absence_table, news_table, news_number_table.
- This study generated evaluation results of the Four Balanced Scorecards. The performance of the financial perspective is measured using two indicators, namely, the number of monthly newspaper sales and the average newspaper sales in one year. It was increased. It was due to the large number of distributors who contributed to the sales of newspapers. This data is taken from the average annual sales.

Customer Perspective Number of Monthly Sales within the City + Number of Monthly Sales Out of the City = Total Newspaper Sales in One Year. Monthly newspaper sales have increased and decreased every year.

The Internal Business Process Perspective of Annual Newspaper Sales divided by the number of cities then multiplied by 100%. To increase the high selling power of newspapers, recent newspaper news should print news which is currently viral, political news, regional news and so on as well as they have excellent layout design. Therefore, there has been an increase in sales within the city and out of the city.

The learning and growth perspective of the Radar Sampit Editor is divided into three aspects, namely the absence aspect illustrates good results. This is supported by the activeness of employees, and the assessment of the news number which belong to the good category with the points they get. The factors supporting the growth and learning perspective are the budget, employee competence and human resources.

References

- [1] Ma, Y., Gao, X., and Gu, S., 2015, "Subject Hierarchy Structure Modeling in Data Warehouse," *International Journal of Database Theory and Application*, 8(4), pp. 256–272.
- [2] Hadiyati, N., 2014, "PENGUKURAN KINERJA DENGAN METODE BALANCED SCORECARD (Studi Empiris Pada Rumah Sakit PKU Muhammadiyah Delanggu Klaten)," 4, p. 4.

Table 11 Test Result

No	Scenario	Test Case	Expected Result	Test Result
1	Clear all login data fields, then click the login button	username: (blank) password: (blank)	Display a validation pop-up that "your username or password is incorrect"	Succeed
2	Enter all login data fields correctly, then click the login button.	username: sekred1 (correct) password: Click the menu tabel.123123 (correct)	Display a validation pop-up that "you have successfully logged in" and displays the main dashboard page.	Succeed
3	Look at the menu	Menu button	Display the drop down contained in the menu	Succeed
4	Click on the 'table' button on the table data to see the details of the data	Click the 'table' button	Display detailed data from the selected table	Succeed
5	Input the absence data	Press the add button	Fill the news data, save, attendance data then see the validation "are you sure do you want to save) and the attendance is successfully saved	Succeed
6	Input the news data	Press the add button	Fill the news data, save, then see the validation "are you sure do you want to save) and the number of news data has been successfully saved	Succeed
7	Input the news number data	Press the add button	Fill the news data, save, then see the validation "are you sure do you want to save) and the number of news data has been successfully saved	Succeed
8	Look at the filters/diagrams on the dashboard, to see the development of the editorial performance per year	Press menu: select report, weekly report	Display the results of the weekly report diagram	Succeed
9	Look at the filters/diagrams on the dashboard, to see the development of the editor's monthly performance	Press menu: select report, monthly report	Display the results of the monthly report diagram	Succeed
10	Look at the filters/diagrams on the dashboard, to see the development of the editor's absence performance	Press menu: select report, editor performance	Display the results of the editor performance diagram	Succeed
11	Print the detailed data from the selected table. Generate reports.	Select print on month/week/editor performance	Present the display for detail data print settings.	Succeed
12	Exit the Editorial Performance system or dashboard	Choose the exit button	Appear the validation "are you sure do you want to leave" if OK then display the editorial performance dashboard login form.	Succeed

- [3] Susetyo, J. and Sabakula, A., 2014, "Pengukuran Kinerja Dengan Menggunakan Balanced Scorecard Dan Integrated Performance Measurement System (IPMS)," *Jurnal Teknologi*, **7**(1), pp. 56–63.
- [4] Nurdin, R. H., 2019, "Pengukuran Kinerja Perusahaan Pada Pt. Yyy Dengan Menggunakan Metode Balanced Scorecard," *Jurnal Manajemen Bisnis dan Kewirausahaan*, **3**(3), pp. 83–90.
- [5] Astriani, W. and Trisminingsih, R., 2016, "Extraction, Transformation, and Loading (ETL) Module for Hotspot Spatial Data Warehouse Using Geokettle," doi: [10.1016/j.proenv.2016.03.117](https://doi.org/10.1016/j.proenv.2016.03.117).
- [6] Goar, V., Rai, J., Rajasthan, N., Universitas, V., and Tanwar, G., 2010, "Meningkatkan Kinerja Extract , Transform and Load (ETL) di Data Warehouse," **02**(03), pp. 786–789.
- [7] Kakish, K. and Kraft, T. a., 2012, "ETL Evolution for Real-Time Data Warehousing," *Proceedings of the Conference on Information Systems Applied Research*, pp. 1–12, <http://proc.conisar.org/2012/pdf/2214.pdf>
- [8] Christian, J., 2010, "Model Data Warehouse Dengan Service Oriented Architecture Untuk Menunjang Sistem Informasi Eksekutif," *Jurnal TELEMATIKA MKOM*, **2**(2), pp. 103–115.
- [9] Sari, M., 2015, "Analisis Balanced Scorecard Sebagai Alat Pengukuran Kinerja Perusahaan PT. Jamsostek Cabang Belawan." *E-Jurnal Akuntansi*, **15**(1), pp. 52–64.
- [10] Malgwi, D. A. A. and Dahiru, H., 2014, "Balanced Scorecard Financial Measurement of Organizational Performance: A Review," *IOSR Journal of Economics and Finance*, **4**(6), pp. 1–10.