Business Intelligence Dashboard Implementation on a Travel Agency in Jakarta

Euis Nina Saparina Yuliani, Heru Subawanto, Anggi Oktaviani

Magister of Industrial Engineering,, Mercu Buana University, Jakarta, Indonesia

Abstract— *Information is growing at an alarming rate. As the development of information, organizations need to manage them and make them can be processed are growing as well. So this makes the problem to get the right information at the right place for the right people. And this fact is important for the company to be successful. This is what causes the Business Intelligence (BI) in the preferences of today's technology. BI is a process from raw data to be read. BI solutions help transform raw data into actionable information that can help support business decision making. This can help companies develop new opportunities. By identifying new opportunities and implement effective strategies, it will result in a competitive market advantage and stable in the long term. In this study, analysis and visualization of large amounts of data from a travel agency in Jakarta to help make the right business decisions using BI tools*

Keywords— *Business Intelligence, Dashboard, Data Mining, Strategy, Travel Agency.*

I. INTRODUCTION

the introduction of the paper should explain the nature of the problem, previous work, purpose, and the contribution of the paper. The contents of each section may be provided to understand easily about the paper.

In current era, information developed with very high speed. Along with the development of information, the need for organizations to manage and to make them can be processed grow as well. This makes getting the right information at the right place to be a problem for the right person. And this is the important thing for the company to be successful. According [15], Business Intelligence (BI) is a technology for managing and displaying raw data in useful ways and have meaning that will help organizations to make the right business decisions. BI relies on data collected from other systems, so that the quality of data is very important for BI.

The purpose of BI is to help decision makers get the right information and make the right decisions to manage the business [10]. BI makes the organization get a more accurate and detailed information about what is happening both on business and on consumers.

There are several advantages of using BI. A well implemented and managed BI will: first, eliminate work based on assumption [11] as BI can provide a more accurate historical data. Second, help identify business opportunities by making the trend of market conditions [4]. Third, help the understanding of consumer behavior [2]. Fourth, help to create a realistic goal [7]. Fifth, help identify opportunities for cross-selling and up selling [1]. And the sixth increase efficiency [14].

II. THEORITICAL REVIEW

BI is a tool to strengthen, analyzing, and visualizing of huge data to help consumers in making decisions. BI systems mainly focus on reporting, fetching data (querying), and analysis of data contained in company data warehouse. According to [16], BI consists of 5 components: Data Sources, ETL (Extract, Transform, and Load), Data Warehouse, Interactive Dashboard and Reports. Fig 1 shows the components of the BI system.

Fig. 1: BI Components

Each of the components are described below. First, data sources. Data can be either fact or statistics collected for reference or analysis. Thus, data such as documents, products, activities, and events can be called as data sources. Second, ETL (Extract, Transfer and Load). ETL is the process of extracting data from source system and bring it to a data warehouse. ETL is a three-function database to pull data from one or more sources of data and put it into another database or data warehouse, both with the same or a different format. Extract is the process of extracting or reading data from a data source of the database. Transform is the process of converting extracted data into a desired format that can be placed in a new data warehouse. Load is the process of placing data into the new data warehouse. Third, the Data Warehouse. Data warehouse is to store data or database that can be used for reporting, query and analysis of data. Fourth, Data Marts. Data Mart is a database that has the characteristics of a data warehouse, but smaller and focus on data on a division of the company (such as the finance or sales).

Fifth Report and Dashboard. Report and Dashboard can be made with various shapes using BI. Dashboard can be used for data analysis and decision-making. The dashboard can be very interactive and very user-friendly.

There are four characteristics of a data warehouse. First: Subject Oriented. Data warehouses are designed to help analyze the data that needs to be constructed for each analysis, for example for sales. Second, integrated. Data can come from various sources and formats, so it must be converted into one format which will reduce inconsistencies and is very useful for further interpretation. Third, non-volatile. Once data is entered into the data warehouse, should there be no corrections or changes again. It is better to perform validation before the data is entered into the data warehouse because data modification in the data warehouse will require time and effort. Fourth, characterized by time or time variant. This means data can change over time and changed data can be traced.

Meanwhile, the Data Mart can be considered as part of the Data Warehouse which is commonly used for specific business team. The main use of a data mart is for BI applications that can be used to collect, store, access, and analyze data. Data Mart require less time and is cheaper than a data warehouse. There are two basic types of data marts, which form the main categories based on data source as source to data marts. First, Dependent. Dependent Data Mart pulls data from an established central data warehouse. While Independent is a stand-alone data marts that are created by pulling data directly from external data sources.

There are several functions that can be performed using BI. The first is reporting [2].The second is analysis to find patterns and deliver significant of data [2]. There are three types of data analysis, forecasts or predictions of the future based on the data at this time, a simple descriptive summary of the observations, and prescriptive to find appropriate measures to the present circumstances. According [13], the third is business performance management (BPM). BPM is one of BI form that is used to monitor and manage the performance of the company by gathering data from multiple sources, analyze and use the knowledge gained to improve the company's performance. Fourth is the benchmark [6]. BI can be used to benchmark and evaluate or compare with standard. Company performs benchmarking to identify areas where performance is less or more.

Data analysis includes identification of important patterns. An analyst is able to see the patterns of grouped small data. While special tools of data mining (DM) have the ability to see patterns in large data sets. These tools are able to analyze a significant relationship that appears only if a number of dimensions visible at the same time.

The purpose of data mining is to understand a large amount of data without the need of supervision on any domain (eg Finance, Marketing, Sales). Data Mining is used to analyze large amounts of data, because if the data set is small, the analysis can be performed by standard techniques or even manually.

Data mining is also called knowledge discovery. Knowledge Discovery Process (KDP) is defined as the process of identifying valid, potential usefulness, and patterns that can be understood on the data [8]. Fig. 2 shows the sequential structure of the KDP model consisting of many steps that every step of trying to solve particular tasks.
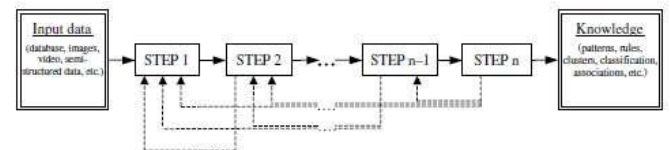

Fig. 2: Structure of KDP sequential model [3]

KDP includes how data is stored and accessed, how the algorithm is effective and can be developed, how to interpret and describe the results, and how to model and support the interaction between man and machine. In KDP, the steps are performed in sequence, the next steps begin after the completion of the previous steps, covering the activities of the understanding of the domain and project data through the preparation and analysis of data until the evaluation and application of the results generated, and the KDP is iterative; there the feedback which is triggered due to the revision.

There are three models of KDP that are quite popular. One of them is 9 steps model [5] where the process consists of nine steps. First, develop and understand the scope of application. It includes the study and understanding of the relevant goal specified by the user. The second is to make the target group data. Selected data based on the attributes that will be used in the search. This includes queries of data that is required. The third is cleaning and initial processing of the data. It includes eliminating data out of scope, associated with the noise and the missing values and the time sequence information. Fourth is the projection and data reduction. This includes the search for useful properties using dimension reduction methods and transformation. Fifth, choose data mining tasks. In this stage, adjustments were made with the goal defined in the first step of data mining methods such as classification, regression, clustering, and others. Sixth, the selection of data mining algorithms to find patterns of data. The seventh is the data mining that generate patterns in the form of a representative such as classification rules, decision tree, regression model and trend. Eight is to interpret patterns in data that it includes visualization patterns and models produced. Ninth is

consolidating Discovered Knowledge which consist of combining the discovered knowledge in system performance, including documenting and reporting to the user.

Characteristics of 9 steps model are iterative process, technical information described in detail but less about the business aspect, and its main application is in commercial Knowledge Discovery and facilitate projects such as engineering, pharmaceutical, manufacturing, and e-business.

Another model that is quite popular is the CRISP-DM (Cross-Industry Standard Process for Data Mining). According [12] in general, this model consists of six steps. The first is the understanding of the business. The first phase focuses on understanding the purpose and requirements of the business perspective which is then converted into a definition of the problem and plan to start the project plan to achieve the goal. The first step is subdivided into smaller, namely: the determination of business objectives, estimate the situation, the goal of determining the DM, making the project plan. The second step is understanding the data. This step begins with the initial data collection and familiarization with the data. These include the identification of data quality problems, the invention of a preliminary understanding of the data, and detection of interesting data parts. This step is fragmented into: initial data collection, data description, exploration data, and data quality checks. The third step is the preparation of data which includes all activities to establish the final data set which requires data to be included in the means DM in the next step. This includes the selection of tables, records, properties, data cleaning, construction of new atributtes and data transformation. This step is prepared in: data selection, data cleaning, data creation, data integration, and the establishment of the data parts. The fourth step is to choose the model and perform the modeling using a variety of modeling tools. It encompasses several methods for the same DM problem and calibration parameters to optimal values. Because some of the methods require a special format for data entry, often looping back to the previous step is needed. This step is sorted in selecting modeling techniques, the establishment of test design, modeling, and estimation of models will be made. The fifth step is the assessment from the perspective of business objectives and to review the steps taken to form the model. The goal is to determine if there are businesses issues that are important that have not been considered. This step is sorted in: evaluation of results of the assessment, reviews the process and determine the next steps. Step six is the implementation which includes the organization and presentation of knowledge discovery with a user friendly way. Depending on your needs, it can be in the form of as

easy as a report or as complex as repeated KDP. This step consists of: plan implementation, monitoring and maintenance plan, final report and review the previous steps.

CRISP-DM model is characterized by: using vocabulary that is easily understood and well documented, understand the iterative nature of the process of the relationship between steps, extensively using model and applied in the fields of pharmaceuticals, marketing and sales.

Another model is the 6 Steps model [3]. This model is inspired by the CRISP-DM model and applicable to academic research. Fig. 3 shows of 6 Steps model. The difference is: to provide steps of the more common and evidence-oriented research, it has DM step (not modeling), introduces some clear feedback mechanism. CRISP-DM only has three major feedback, while this model has a detailed feedback, and modification of the final step; knowledge discovery for some domain may be applied in other domains.
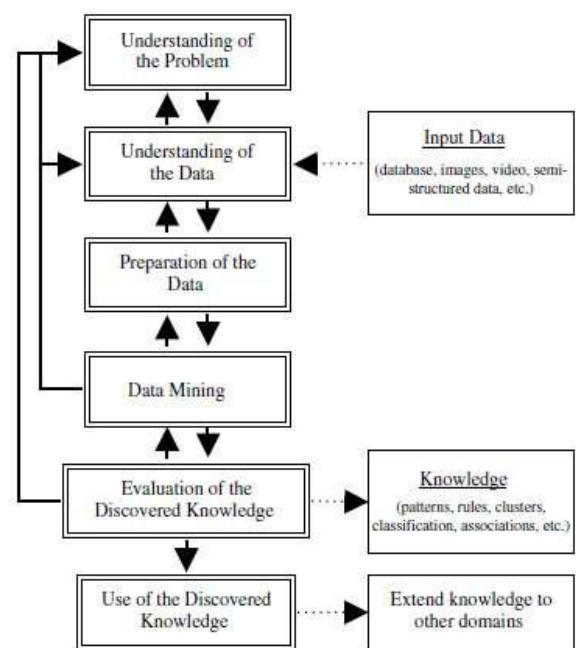

Fig. 3: KDP 6 Steps Model [9]

Description of this 6 steps are: first, understanding the coverage of issues including working closely with experts to define problems and goals of the project, the introduction of key personnel, learn about the current solution to the problem. Goal of the project is translated into DM objective and early selections of DM tools that will be used. The second is the understanding of data including collecting samples data and determine which data -including form and size- that will be needed. Basic understanding is used to assist this effort. Data completeness is examined, redundancy, missing values, and others. The third step is the preparation of data.

Attention to specify what data will be used as input to the DM tool in the next step. This step includes sampling, test correlation and importance, data cleansing and data completeness check, remove or correct the interference and missing values. Data that has been cleaned then processed further by selecting features and extraction algorithms to reduce the dimensionality by lowering properties and to new data summarization. The result is data that meets the specific needs of DM tools. Fourth is Data Mining which includes the use of various DM methods to get understanding or to get new information from the data to be processed. Fifth is the assessment of Knowledge Discovered which include understanding the results, check whether the discovered knowledge is quite interesting, interpretation of the results by the experts, and examining the possible effects of discovered knowledge. Only approved models are used and the whole process is reexamined to identify any alternative actions that can be done to improve results. Step six is the use of Discovered Knowledge consisting of planning where and how the discovered knowledge will be used. Applications in this domain may be used in other domains. Plan to monitor the implementation of discovered knowledge is created and the entire project is documented. Finally, knowledge discovered is used officially.

III. IMPLEMENTATION OF BI AND DASHBOARD DESIGN

This research will use free version of QlikView that can be downloaded via the Internet. Because the test can not be performed directly in conjunction with ERP to run, then this study will use annual data that has been exported. QlikView will be used to create dashboards and reports.

QlikView is a software to plan and execute BI. QlikView allows users to create an interactive dashboard and can be operated without the need for specialized expertise to run it. QlikView is highly interactive and also works very quickly.

QlikView can use and easily filter data from multiple sources such as ERP, document, even data from the internet. Design in QlikView also can be designed according to the needs and views. The advantages of QlikView is that QlikView is highly scalable to follow the requirements of the company, not vice versa.

Unlike other BI softwares, QlikView is based on businesses and consumers so it is so easy to use. QlikView turns data into an information that can help in the decision making process. Any changes will be dynamically calculated.

Fig. 4: Database Sample

The sample data used in this dashboard implementation design is using transaction data from a travel agency in Jakarta, which has branches in many big cities in Indonesia. The amount of data in the sample reached a record 18,948 with column consists of ID, Transaction Date, Tour Name, Customer Name, Branch, and Remark. Fig. 4 shows the columns in the data source.

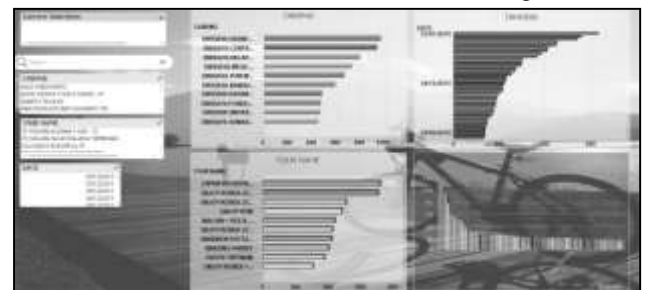The initial view of the dashboard is shown in Fig. 5.

Fig. 5: Initial view of a Travel Agency Dashboard

In the view of Fig. 5, there are four graphs. The first graph is a graph that displays the branch of the selected branch, which is named Cabang in Bahasa Indonesia. The second chart is a graph that displays the code Tour Name tourism, such as Japan Vacation. The third graph is a graph that displays the date, which is titled Tanggal in Bahasa Indonesia. And the fourth graph is a graph that shows Remark record of a transaction.

The selection of what will appear in the four graphs, is determined by what is selected in the three list boxes on the left side. The first list box is a list box that shows the branch can be selected. The second box is the list box Tour name list that displays the names of the tour code can be selected. And list the last box is the list box data showing the dates of transactions that can be selected. What we choose, look at the Current Selection box at the top left.
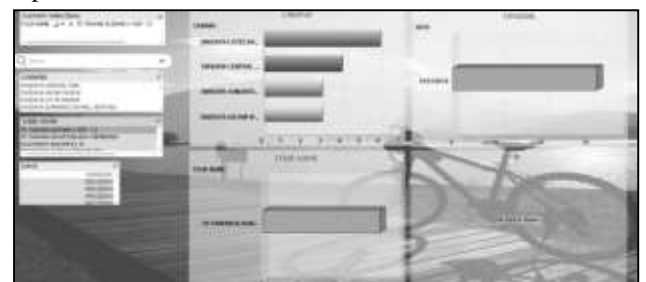

Fig. 6: Tour Name Selection 7D Taiwan Alishan

For example, we choose only the tour name of 7D Taiwan Alishan. The results, shown in the Fig. 6. In this figure, it seems that the graph only shows Tour Name 7D Taiwan Alishan with a total of 16 transactions. In the graphic branch, there are 4 branch transactions to 7D Taiwan Alishan. In the chart Date, seen that 15 transactions 7D Taiwan Alishan only occur on March 16, 2016. And on the chart Remark, nothing is displayed, which means that no records are required when making transactions.
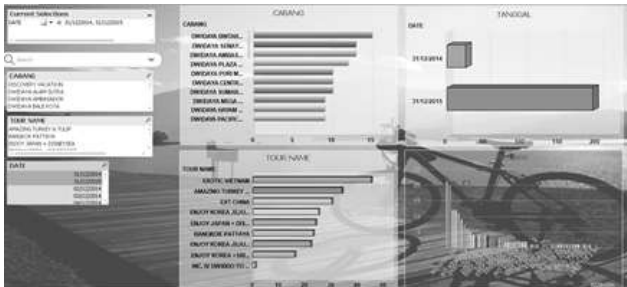

Fig. 7: Comparison of transaction date December 31, 2014 and December 31, 2015

In addition, we can compare the number of transactions at the end of 2014 and the end of 2015. Fig. 7 shows this. The graph shows that the end of 2015 there were more transactions than at the end of 2014. Trade Branch of the largest seen in the chart, which is done by the Bintaro branch and the biggest Tour is a visit to Vietnam and then to Turkey.

Fig. 8: Selection of Branch and Tour Name

We can also make more than one choice, like choosing a Bintaro branch and travel to Korea. The results are shown in Fig. 8. In Fig. 8, seen in the graphic branch, just Bintaro branch is displayed. While the graphics Tour Name, just Korea is displayed. While the chart date, several dates of the transactions are displayed.

IV. CONCLUSION

From the above, we can transform raw data into information that is meaningful and useful by using BI tools. BI is a great tool for reporting, benchmarking and analyzing data to make the right business decisions. BI is used by decision makers throughout the organization. At the senior level, is used to create a strategy and on a lower level, will assist each individual in his daily job.

REFERENCES

[1] Bose, R. : Advanced analytics: Opportunities and challenges. Industrial Management & Data Systems, 109(2), 155-172. (2009)

[2] Chen, H., Chiang, R. H., & Storey, V. C. : Business Intelligence and Analytics: From Big Data to Big Impact. MIS quarterly, 36(4), 1165-1188. (2012)

[3] Cios, K. J., Pedrycz, W., & Swiniarski, R. W. : Data Mining and Knowledge Discovery. In Data Mining Methods for Knowledge Discovery (pp. 1-26). Springer USA. (2007)

[4] El Deen, M. A., & Solayman, M. M. : Maximizing strategic performance results: Adopting balanced scorecards and BI tools. International Journal of Computer Applications, 117(10) (2015).

[5] Fayyad, U. M., Smyth, P., & R. Uthurusamy Piatetsky-Shapiro, G. : Advances in knowledge discovery and data mining (Vol. 21). (Eds.). Menlo Park: AAAI press. (1996)

[6] Govindarajan, M., & Srinivasan, B. : Ensembles of classification methods for data mining applications. International Journal of Information Engineering and Electronic Business, 5(6), 6-21. (2013).

[7] Kasem, M., & Hassanein, E. E. : Cloud business intelligence survey. International Journal of Computer Applications, 90(1) (2014).

[8] Mariscal, G., Marbán, Ó., & Fernández, C. : A survey of data mining and knowledge discovery process models and methodologies. The Knowledge Engineering Review, 25(2), 137-166. (2010).

[9] Pal, N. : Advanced techniques in knowledge discovery and data mining. Springer Science & Business Media. (2007).

[10] Pourshahid, A., Johari, I., Richards, G., Amyot, D., & Akhigbe, O. S. : A goal-oriented, business intelligence-supported decision-making methodology. Decision Analytics, 1(1), 1. (2014).

[11] Ranjan, J. : Business intelligence: Concepts, components, techniques and benefits. Journal of Theoretical and Applied Information Technology, 9(1), 60-70. (2009).

[12] Shearer, C. : The CRISP-DM model: the new blueprint for data mining. Journal of data warehousing, 5(4), 13-22. (2000)

[13] Tank, D. M. : Enable better and timelier decision-making using real-time business intelligence system. International Journal of Information Engineering and Electronic Business, 7(1), 43-48. (2015)

[14] Trkman, P., McCormack, K., De Oliveira, M. P. V., & Ladeira, M. B. : The impact of business analytics on supply chain performance. Decision Support Systems, 49(3), 318-327. (2010)

[15] Uçaktürk, A., Uçaktürk, T., & Yavuz, H. : Possibilities of Usage of Strategic Business Intelligence Systems Based on Databases in Agile Manufacturing. Procedia-Social and Behavioral Sciences, 207, 234-241. (2015)

[16] Wajong, A. M. R. : Business intelligent system to make management decision in project management. International Information Institute (Tokyo).Information, 18(8), 3353-3360. (2015)