

Analysis and Design Data Warehouse For E-Travel Business Optimization

I Nyoman Angga Prabawa^{1*}, Dewa Agung Krishna Arimbawa², and I Gusti Ngurah Janardana³

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

³Department of Electrical and Computer Engineering, Udayana University

*prabawa_angga@yahoo.co.id

Abstract Tourism, accommodation, and transportation service industries are business with high prospects, especially in Bali. Development of digital technology into the e-travel industry has business processes that can be optimized through a data warehouse with a variety of data as the initial start of analyst business data. A problem reality is less and almost no travel company has developed a data warehouse to support business needs. Building a data warehouse is a technology to increase business value through data analysis on a variety of stored data. Using the fourth steps dimensional modeling methodology, the data warehouse is expected to be in accordance with the company's analysis to increase business needs. The design results consist of determining business processes, grain declaration on dimensions, identifying dimensions, and determining facts. The design results using snowflake scheme and become a solution for companies to predict business trends, maintain quality, improve competitiveness, and exist in the long run.

Index Terms—E-Travel, Data Warehouse, Fourth Step Dimensional Modelling, Snowflake Schema, Business Optimization

I. INTRODUCTION

Information technology has developed until now. Many jobs are completed quickly, less cost, efficiently, but produce good output with the presence of information technology. Technology is able to provide good information through a variety of data related. Data relating and having connectivity is the present day challenges to be processed, treated, and utilized for a good purpose. In the business industry, data is a valuable asset for companies. Data consists of various types, such as employee, product, customer, and transaction data. The amount of data that is increasing every time can help a company's in optimizing its business. This amount of data will require a technology known as Data Warehouse. The data warehouse is a place to store all important data to be processed, observed, and understood by the company. This technology has benefits especially in business optimization by conducting analyst data. In this study, the design of the data warehouse is expected to solve cases in the tourism industry or e-travel companies in Bali.

E-Travel is one type of online-based tourism industry that has developed in various regions, especially Bali. Bali is one of the best tourist destinations in the world and is always visited by various foreign tourists. One of the biggest revenues for Bali is the tourism sector. Thus, to maintain the viability of this sector is very important. Various E-Travel companies in Bali are very competitive to serve tourists on vacation. E-Travel is being considered for the Tourism Office to be able to maintain quality and be able to have long-term sustainability, so that tourism in Bali

remains good in the world.

E-Travel in Bali has used information technology to support business processes. Booking via website, promotion of tour packages, up-to-date news about tourism, and easy and fast customer service. This service will make consumers interested in using the services of an e-travel. With the increase in services, it will bring in many new consumers and increase the intensity of old consumers to vacation in Bali through e-travel services. In a certain period of time, transaction data from an e-travel will certainly increase rapidly. The data growth will increase the amount of memory capacity in the daily database. Daily databases are called Online Transactional Process or OLTP. Data on OLTP will be useless if it is not used as a collection of unique information to be useful for the development of e-travel business. How are business trends, the pattern of tourist arrivals as service users, the frequency of favorite tour packages in a certain period of time, the intensity of the vehicle running, and the frequency of drivers serving each transaction within a certain period of time. This must be known by e-travel companies to monitor and serve as a knowledge base in developing the e-travel business to increase. Data Warehouse technology is needed as a solution for storing all data storage on OLTP. OLTP data is sent to the Online Analytical Process or OLAP as a data warehouse manifest, where a business analyst is able to identify all the needs to improve e-travel company's business optimization.

Many study literature is a variety of references in designing a data warehouse. Based on [1], in the study conducted a data warehouse design at the radiology department. The design was carried out to meet analysis and reporting needs on recording patient history and assessing

accuracy level of the diagnosis process. In [2], it provides a design overview of the data warehouse that aims to assist decision-making process system for stakeholders in case of sales and inventory systems. Research by [3] explains the development of a data warehouse as a decision support system in the tourism sector. The research aims to answer the needs for reporting and analysis in the Egyptian state tourism. Discussion by [4] identified brief and detailed reporting needs regarding amount complaints in a government agency. The study designed a data warehouse that showed good results and could facilitate users in reconciling concise and detailed complaints reports.

The process of designing a data warehouse can refer to several model references. One of the data warehouse modelings is the fourth step dimensional modeling proposed by Kimball. The fourth step dimensional modeling process consists of selecting a business process, choosing a grain, identifying dimensions, and identifying a fact table. The scheme used in this study is the snowflake schema. This scheme describes dimension-to-dimension which has the advantage of storing the memory capacity used. The dimension table will help the fact table to display the data requirements that the analyst wants to improve e-travel business optimization.

In this study, the design of a data warehouse with a snowflake schema was produced to optimize e-travel business. This design is illustrated and expected to answer the needs of e-travel companies in business optimization. Then, they can maintain quality and exist for a long time.

Discussion in this study is arranged as follows, Section II literature review focused on important points that support research through literature, Section III research methodology explains the sequence and stages of research in designing a data warehouse, Section IV result and analysis contains a discussion of research along with results. The final results of the study convey conclusions from the summary of the research stages in Section V.

II. LITERATURE REVIEW

A. E-Travel

Electronic Travel is a modern term from the development of conventional travel services in the digital era. E-Travel utilizes technology, such as the internet to simplify its business processes. Marketing of promotions, looking for customers, and all products or services offered can be distributed through the internet network, whether a website or mobile application. Travel service is one of the fields of business that is currently starting to change from conventional to online which is felt very easy and practical for service owners and users.

B. E-Travel Business in Bali

The travel business in Bali has been going on for a long time. Along with the development of the era, the business culture that used to be conventional systems has now begun to expand online or in the internet network. The e-travel

business continues to grow and Bali is one of the provinces that has the highest income through the tourism sector. The number of tourist arrivals in Bali is recorded from year to year as high and continues to increase in a certain period. By the Central Statistics Agency [5] stated tourism in Bali regarding the high number of foreign tourist arrivals every month. Travel services in Bali are also regulated in the applicable laws, namely the Regional Regulation of Bali Province Number 1 of 2010 concerning Tourism Travel Services. It was written that the business of tourist travel services is one part of the service business in the tourism sector needs to be regulated and managed by fulfilling legal protection, security, and safety for tourists.

C. Data Warehouse

Data Warehouse is a very large collection of data, subject-oriented, formed naturally, integrated, time variants, and fixed on the stored data management decision-making processes, which can be improvised from analysis and quality to produce information that is easily accessible to users [6]. According to [1], the data warehouse has four characteristics, including subject-oriented, integrated, time variant, and non-volatile.

D. Fourth Step Dimensional Modelling

Kimball recommends a very unique development methodology for the data warehouse, including a bottom-up that develops a data mart in every business process [7]. The methodology for dimensional modeling consists of four processes:

1. Select The Business Process

This stage identifies the business processes needed in using the data warehouse to determine the dimensions and facts at the end of the modeling.

2. Declare The Grain

The contents of the fact table can be identified through declared grain. Grain will map and match the information needs to be obtained from the design of the data warehouse.

3. Identify The Dimensions

Dimension tables are built to meet the needs provided in the fact table. Development the dimension table will make it easier to understand and interpret targets to be achieved in the fact table.

4. Identify The Fact

The fact table is the embodiment of the data warehouse and is the main target of the business needs. The fact table will display the results of business processes that are expected to meet all needs. By [2], the fact table is a table that generally contains numbers and historical data where the key is very unique because it becomes the primary and foreign keys relationship of the existing dimensional relations.

E. Snowflake Schema

The snowflake schema is one method of modeling a data warehouse with a normalization process for existing dimension tables [8]. Data dimensions in this scheme are the subject of information to become the basis for decision-making.

III. RESEARCH METHODOLOGY

This research is through various processes needed to achieve the expected results. The methodology carried out is illustrated in Figure 1 below.

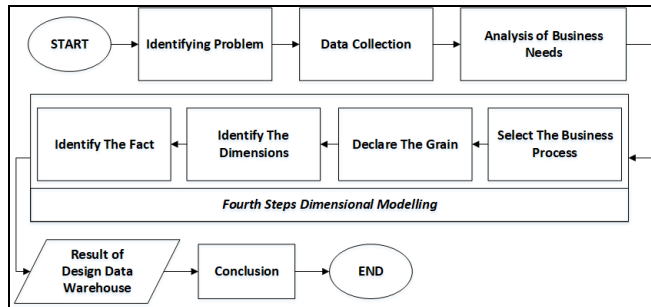


Figure 1 Research Methodology

Figure 1 shows the methodology of the research conducted. Business optimization problems become identification findings as to the basis for designing a data warehouse. Data collection is carried out through the observation process of a travel company and through the study of literature as support. Business optimization raises an analysis of how to improve it through various knowledge of business needs in e-travel. This business need is the foundation for the design of a data warehouse using fourth steps dimensional modeling through four stages, namely:

- 1) Select The Business Process
- 2) Select The Grain
- 3) Identify The Dimensions
- 4) Identify The Fact

Through these stages, the results of the data warehouse design are expected to be the company's answer need for improving e-travel business optimization. The design results will produce conclusions as an evaluation topic so that the design of the data warehouse can be effectively run and applied to e-travel companies.

IV. ANALYSIS AND RESULT

The results of this study are reviewed in several explanations according to the design methodology described in the previous section. The author designed the e-travel data warehouse in several stages.

A. Analysis of Business Needs

This stage is the main thing from what is needed when the data warehouse design is successfully realized. Based on the e-travel database generally, it has OLTP as shown in Figure 2 below.

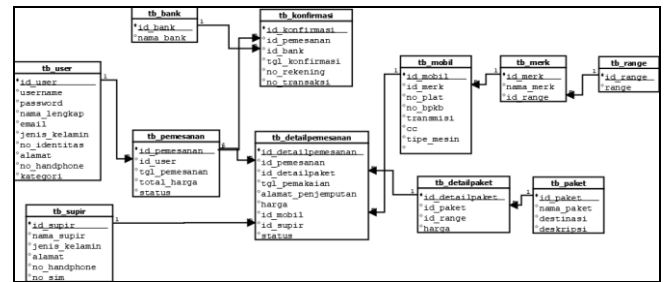


Figure 3 OLTP Travel Company

Figure 2 is an OLTP design as a database that serves various e-travel business daily needs. Based on these images, business processes that exist in travel have very diverse data, including:

1. Vehicle Data contains a company in carrying out travel services. The data concerns the name, brand, vehicle number, type of transmission, and range of passenger capacity.
2. Driver Data is a variety of stored data for information purposes of drivers working in company. Contains name, place and date of birth, mobile number, sex, and address.
3. Customer Data collects all important customer data, such as name, place and date of birth, address, email, mobile number and others.
4. Employee Data contains employee personal data who working at the company, which also contains the name, place and date of birth, address, and mobile number.
5. Tour Package Data is a collection of tour packages offered by the company. Contains data about package names, destinations, names of tourist attractions, and prices offered.
6. News Data stores various news about tourism. The purpose of this data is only for the interests of the company's website.
7. Bank Data contains a list of banks that work with companies to provide travel service payment transactions between the company and the customer.
8. Order Data contains detail transactions of the e-travel company to customers. All transactions using e-travel services will be recorded here.

From a variety of data that is a business requirement in e-travel, companies will use some data that can be needed to achieve their goals. The business needs of the e-travel data warehouse design are set on the business optimization side. The target needs are about the recapitulation of the orders number within a certain period of time, the use frequency of each vehicle, reports related the tour packages ordered within a certain period of time, and the service frequency of drivers working. The business optimization side is analyzed as a business need that requires some data according to the target, among others:

1. Vehicle Data. This data is needed to find out an entire history of vehicle usage. Intensity of each vehicle can

be determined by related data to Order Data.

2. Driver Data. Its importance is to find out an entire history of the services of each driver in serving travel services. Each driver will have a recorded trip that has been done in a certain period of time-related to Order Data.
3. Customer Data. Required with Order Data related, because it can be counted in a certain period of time for number customers use e-travel services.
4. Tour Package Data. Required to find out which tour package has highest interest by customer. So, there will be data records that are connected to the Data Order.
5. Order Data. It is data most important. Because it is a container for all other data. This data is a source of evaluation in the optimization of e-travel business.

Five data are estimates for authors to be able to focus on designing the e-travel data warehouse for increase business optimization.

B. Select The Business Process

This stage is the beginning of the process of fourth step dimensional modeling. The business process is chosen based on needs specified in business optimization according to the analysis described of the previous discussion. Business processes that want to be proposed design data warehouse, namely:

1. Amount of order
2. Frequency of vehicle use
3. Tour packages ordered
4. Intensity of driver services

The chosen business process is needed in the optimization of e-travel business to be able to be analyzed, studied, and identified data patterns in a certain period of time.

C. Declare The Grain

Grain on e-travel is determined to describe any data recorded in the fact table later. The Grain in the data warehouse design is used to record the process of transaction data from business processes offered by e-travel to customers. Determination of grain at this stage is illustrated in Table 1 below.

Table 1
Declare The Grain

Grain Dim	Amount of Order	Vehicle Use	Tour Package	Driver Service
Tour Package			✓	
Vehicle		✓		
Customer	✓			
Driver				✓
Time	✓	✓	✓	✓

Table 1 describes the process of determining grain with various dimensions needed to realize a fact table based on selected business processes.

D. Identify The Dimensions

Dimensions are identified to meet the fact table needs. Each selected dimension table is a data table that has important data records for the output of the data warehouse.

Table 2 until Table 11 is a collection of identified dimensions as supporting data tables for realization of the fact table needs. It was explained that designing a data warehouse for business optimization requires several dimension tables, such as:

Table 2
Member Dimension

Dim_Member
id_member
nama_member
no_identitas
jenis_kelamin
alamat
email
no_handphone

Table 3
City Dimension

Dim_Kota
id_kota
nama_kota
id_provinsi

Table 4
Province Dimension

Dim_Provinsi
id_provinsi
nama_provinsi
id_negara

Table 5
Country Dimension

Dim_Negara
id_negara
nama_negara

Table 6
Driver Dimension

Dim_Driver
id_driver
nama_driver
no_identitas
alamat

Table 7
Time Dimension

Dim_Time
id_time
tahun
bulan
tanggal
hari

Table 8
Vehicle Dimension

Dim_Kendaraan
id_kendaraan
nama_kendaraan
no_kendaraan
no_bpkb
transmisi
cc
mesin
kapasitas
id_merk

Table 9
Vehicle Brands Dimension

Dim_Merk
id_merk
nama_merk

Table 10
Package Tour Dimension

Dim_Paket
id_paket
nama_paket
id_destinasi

Table 11
Destination Dimension

Dim_Destinasi
id_destinasi
nama_destinasi
kategori

Table 12
Transaction Fact

Fact_Transaction
id_transaksi
id_member
id_paket
id_kendaraan
id_driver
jumlah_penumpang
alamat_penjemputan
id_time
total

Table 12 shows some data in the fact table which is formed from relationships between dimension tables. In this case, the author designed a fact table called fact_transaksi. This fact table is the center of all connectivity from various dimensions to meet the needs in terms of optimization according to the e-travel business process.

F. Design with Snowflake Schema

The design of the e-travel data warehouse for business optimization improvement uses the snowflake scheme as shown in Figure 4 as follows:

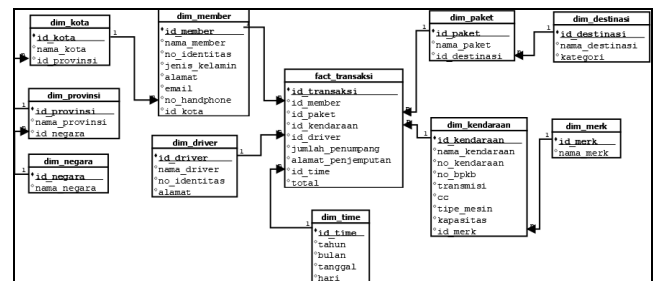


Figure 4 Design of Data Warehouse E-Travel

Figure 4 shows the snowflake scheme used to design the e-travel data warehouse. Author assumes that this scheme and design model, optimization of e-travel business can be improved by the ability of analysts to process various available data. Business processes that have been determined in the previous process help the author design an ideal and effective design for e-travel.

G. Analysis of Snowflake Schema

The implementation of the snowflake scheme has several advantages and disadvantages. According to [9], the snowflake scheme is better than other models, because more information can be used from data in the process of making decision. Some considerations that underlie to use this scheme are it relieves the memory load used, maintenance and update processes are easier, and query process is faster

E. Identify The Fact

At this stage, the fact table is selected based on the identified dimension table. A fact table in the data warehouse interprets the business processes that exist in e-travel and makes it easier for analysts to optimize business. The table of facts identified is shown in Table 12 below:

on ETL. The reason is dimension-to-dimension in the data structure, where each data category will have separate data units from each other. The weakness will have an impact on the level of query complexity, higher data structure and difficulties in finding data. However, the design of this scheme is still feasible to be used in a particular case study that involves dimension-to-dimension connectivity to support the need for fact tables.

V. CONCLUSION

The conclusion of this study is that the design of a data warehouse in an e-travel company will be very possible to improve business optimization. By analyzing the business needs expected in this design, the stages of the fourth-dimensional dimensional modeling can be implemented. This method is also very simple and can help in designing the design of the e-travel data warehouse.

REFERENCES

- [1] V. Georgiana, D. Kartawiguna, and Indrajani, "Radiology Data Warehouse Development As A Means Of Education, Research, and Quality Assuarance," *Journal of Theoretical and Applied Information Technology*, vol. 73, no. 1, pp. 65–71, 2015.
- [2] M. Y. Ricky, "Datawarehouse Sales and Supply of Goods Model Based On HTML5," *Journal of Theoretical and Applied Information Technology*, vol. 61, no. 1, pp. 175–182, 2014.
- [3] T. A. Abdulaziz, I. F. Moawad, and W. M. Abu-Alam, "Building Data Warehouse System For The Tourism Sector," in *International Conference on Intelligent Computing and Information Systems*, 2015, pp. 410–417.
- [4] I. P. A. P. WIjaya, Wahyudi, and M. Mataram, "Data Warehouse Implementation on Denpasar City Online Community Complaints System," *International Journal of Engineering and Emerging Technology*, vol. 2, no. 1, pp. 67–71, 2017.
- [5] BPS, "Berita Resmi Statistik Perkembangan Pariwisata Bali Maret 2019," 2019.
- [6] P. Widiadnyana, M. A. Maricar, I. N. Arnawan, and S. Ariyani, "Analysis of Data Warehouse for Transactional Analysis Methods Online Analytical Processing (OLAP) at Company XYZ Silver," *International Journal of Engineering and Emerging Technology*, vol. 2, no. 1, pp. 72–75, 2017.
- [7] K. A. B. Permana, G. B. Subiksa, and M. Sudarma, "Design Data Warehouse For Centralized Medical Record," *International Journal of Engineering and Emerging Technology*, vol. 2, no. 2, pp. 47–51, 2017.
- [8] M. P. A. Ariawan, I. M. D. Ardiada, and Y. P. Sudarmojo, "Design of Library Data Warehouse Using OLTP Result of Services Analysis," *International Journal of Engineering and Emerging Technology*, vol. 3, no. 1, pp. 62–65, 2018.
- [9] K. A. B. Permana, D. A. P. Wulandari, and P. A. Mertasana, "Online Analytical Processing (OLAP) for Disaster Report," *International Journal of Engineering and Emerging Technology*, vol. 3, no. 1, pp. 26–29, 2018.