International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

Implementation of Extract, Transform, Load on Data Warehouse and Business Intelligence Using Pentaho and Tableau to Analyse Sales Performance of Olist Store

Tiara Difa Anggrainy[1], Agustin Rusiana Sari[2]
[1, 2]Business Information Systems, Gunadarma University, Jakarta, Indonesia, 10430
Email Address: [1]difatiara(at)gmail.com, [2]agustin(at)staff.gunadarma.ac.id

Abstract**—** *Business development in the world is increasing with the use of information technology which has an impact on local, national and international business activities, as well as generating big data. Big data that is generated needs to be processed, so that it can be used to analyse sales performance. Olist Store data processing is carried out by implementing the ETL process in the data warehouse using Pentaho and visualizing Business Intelligence on the smart dashboard using Tableau. The data warehouse design method used is a nine-step methodology. The result of this study is Olist Store data warehouse using PostgreSQL database and smart dashboard. The result of the Olist Store sales performance analysis shows that from 2016 to 2017 there was a significant increase, but in 2018 it decreased. Therefore, Olist Store management needs to improve the quality of sales by building new innovations, making promotion methods more attractive and improving the quality of the website to support customers in shopping.*

Keywords**—** *Business Intelligence, Data Warehouse, ETL, Visualization*.

I. Introduction

The business industry in the world is growing and is becoming the main target of society in surviving the COVID-19 pandemic era. Business processes cannot be separated from the important role of information technology as a medium used in every activity. Information technology is rapidly growing and has a lot of impact on business activities locally, nationally and internationally. A collection of data generated from business processes in large and complex quantities to be processed with conventional data processing technology is commonly referred to as big data [1].

Big data allows a business or even a company to collect data in real time that can be used to analyse the business processes that have been carried out, improve the efficiency of resource use and as a support for business evaluations to improve the quality of sales or marketing in the future. Big data processing does not guarantee that a business will run well and smoothly, but it can be optimized by utilizing a database architecture in the form of a data warehouse in order to produce rapid, accurate and secure information.

A data warehouse is a combination of technology and components that allow strategic data processing into information that can be used by business people or companies in assessing their business conditions and making decisions for business development [2]. The data warehouse is built for synchronizing transaction data that is manipulated using the ETL (Extract, Transform, Load) process to align raw data into a data warehouse. The ETL process can be carried out using various software, one of which is Pentaho Data Integration (PDI). Pentaho is able to collect data from multiple sources and manipulate data in many ways such as eliminating duplication, filtering, cleaning and formatting, as well as storing data with different formats and purposes.

Data processing in the data warehouse is closely related to the creation of Business Intelligence (BI). In it, BI is a data presentation technology that has been processed into useful information and has business value, then visualized into a smart dashboard in the form of diagrams, graphs and analytical reports. Making smart dashboards can be done with various BI tools, one of which is Tableau. The implementation of processed and visualized data into the smart dashboard describes the sales performance based on the period of the data studied, resulting in reports that can assist the authorized company in analysing data and making decisions.

Big data generated from the Olist Store business process continues to grow significantly, so it is necessary to process data that can be used to analyse sales performance and support business process quality improvement by implementing the ETL process in the PostgreSQL data warehouse using Pentaho, also visualizing BI on the smart dashboard using Tableau. The data used in this study is Olist Store transaction data for the 2016-2018 period. The results of the ETL implementation in the data warehouse and BI visualization are expected to be used as a reference in assessing the effectiveness and efficiency of sales at the Olist Store, as well as assisting in making decisions to determine future strategies.

II. Literature Review

A. Performance Assessment

The company's performance is the result of work achieved by a company during a certain period in accordance with the main tasks and functions that have been determined based on applicable regulations [3]. The company's performance needs to be known and confirmed to certain parties to be able to analyse the level of achievement of a company's results associated with the vision, as well as to know the positive and negative impacts of an operational policy [4].

Tiara Difa Anggrainy and Agustin Rusiana Sari, "Implementation of Extract, Transform, Load on Data Warehouse and Business Intelligence Using Pentaho and Tableau to Analyze Sales Performance of Olist Store," *International Research Journal of Advanced Engineering and Science*, Volume 7, Issue 2, pp. 368-374, 2022.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

B. Big Data

The origin of big data technology which started with the success of giant web service companies such as Google and Facebook in managing and utilizing unstructured data in the form of Consumer Generated Media (CGM) and Click Stream with very large volumes. Then, a concept known as big data emerged and became the center of attention in the world of information technology. In simple terms, big data is defined as a collection of data with a very large volume that is too complex to be processed with conventional data processing technology. The business world in this era of information technology cannot be denied that it is very dependent on the utilization of data and information, so the business world is one of the areas of application of big data [1]. Big data is also defined by three characteristics or attributes that were first proposed by Gartner and used by IBM which can be seen in Figure 1.

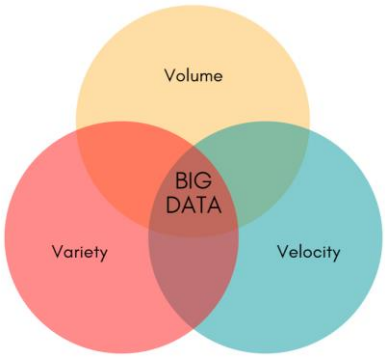

Fig. 1. Big Data Characteristics

C. Business Intelligence

The institution providing education and training in the data warehouse and BI industry, the Data Warehousing Institute defines business intelligence as a process, technology and tools needed to turn data into information, knowledge and plans that drive profitable business activities. Business Intelligence includes data warehouse processes, business analysis tools and content or knowledge management. This analytics combines data warehousing, data mining, multidimensional analysis, flow and graphical visualization to provide a vision that can enable immediate reactions to emerging opportunities while enabling environmental evaluation over time to find ways to improve and grow the business [5].

D. Database

The database consists of the word "base" as a headquarters or warehouse, while "data" is a record of a collection of facts that represent objects such as humans, animals, concepts, events and so on which are manifested in the form of letters, numbers, symbols, pictures, text, sound or other combination. Broadly speaking, the database is a collection of data groups that are interconnected and organized in such a way that they can be reused quickly and easily. The database can also be interpreted as a collection of data in the form of files, tables or archives that are interconnected and stored in electronic storage media, for ease of arrangement, sorting, grouping and organizing data according to the purpose [6].

E. Data Warehouse

A data warehouse is a blend of technologies and components which enable the strategic use of data with techniques for collecting and managing data from multiple sources to provide meaningful business insights. A data warehouse is also an electronic storage of large information in the form of business-generated data designed for query and transaction processing analysis. Data is processed, transformed and absorbed so that users can access the processed data in the data warehouse through business intelligence tools, namely SQL clients and worksheets. The data warehouse works as a central repository that combines information coming from different sources into one comprehensive database. Data enters a data warehouse from transactional systems and other relational databases [2]. The data warehouse architecture can be seen in Figure 2.

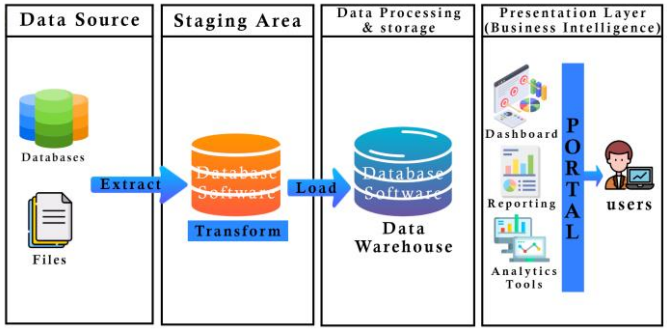

Fig. 2. Data Warehouse Architecture [7]

F. Extract, Transform and Load

Extract, Transform and Load (ETL) is a technology that processes data transfer between systems by accessing data from various sources and applying rules to clean the data, and it is stored in a data warehouse for analysis. The purpose of the ETL process is to collect, filter, process and combine relevant data from various sources to be stored in a data warehouse. The advantages of using ETL are ease of use, graphical flow, operational robustness, structural design, analysis, profiling and data cleansing, and complex data management [8].

G. Python

Python is a high-level programming language that is very popular today because it is considered powerful and close to human language. Python's popularity is also due to the increasing need in the fields of data science, machine learning, artificial intelligence, facial recognition and other fields. Big companies like Facebook, Google, Instagram, Netflix and other digital companies are relying on Python as part of their application programming language [9].

H. Related Research Journals

Several studies related to data warehouse design by implementing the ETL process have been carried out by previous researchers. Study of Implementation of ETL (Extract, Transform, Load) in Sales Data Warehouse using Pentaho Tools [10]. This study designs a data warehouse to process sales data into information that is used to analyse sales performance and become a reference in decision making. The data used is

Tiara Difa Anggrainy and Agustin Rusiana Sari, "Implementation of Extract, Transform, Load on Data Warehouse and Business Intelligence Using Pentaho and Tableau to Analyze Sales Performance of Olist Store," *International Research Journal of Advanced Engineering and Science*, Volume 7, Issue 2, pp. 368-374, 2022.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

Superstores's sales data in the United States for the 2014-2017 period and produced output in Microsoft Office documents, also data transformation to MySQL database.

Study of Data Warehouse Implementation and Its Application at PHI Minimart using Pentaho Tools and Power BI [11]. This study builds a sales data warehouse with the ETL process using Pentaho and visualization of sales information using Power BI. The result of the data warehouse design is analysed through the OLAP (On-Line Analytical Processing) process, the creation of a cube or schema workbench and the creation of a smart dashboard. This study produces sales data for PHI Minimart stores in 2008 which are displayed in the dashboard in the form of sales charts, items sold and total sales for each branch.

Study of Data Warehouse Design for Sales Transaction Data Using Snowflake Schema (Case Study: Online Market Dataset) [12]. This study uses snowflake schema data modeling and a nine-step methodology. ETL data processing using Talend Data Platform and Tableau to display informative data from the designed data warehouse. The result of this study is to produce a data warehouse design that can be visualized using Tableau which makes the processed data into useful information.

Another study conducted the Impact of Business Intelligence on the Quality of Decision Making, a Mediation Model [13]. This study uses bootstrap percentile and T-statistics. The quality of BI management has a positive direct and/or indirect effect on data quality, information quality, and BI solution coverage. This finding also supports the critical success factor (CSF) literature by providing the evidence of the importance of proper BI project management, but it is also found that high quality BI management translates into more comprehensive BI solutions and more robust deployment of BI applications across business functions.

III. Methodology

This research was conducted in several stages, namely data collection, data analysis, data cleansing, ETL design, data warehouse design and business intelligence visualization. The data collection stage is carried out by documenting public data and analysing sales performance at an e-commerce company. After the data collection stage, the data is cleaned from raw data to become clean data free from unwanted things, such as data duplication, unnecessary punctuation, adjustment of records with columns and others. Next is the ETL design stage, where at this stage the data will be extracted, transformed and loaded into the PostgreSQL database. The next stage is to design a data warehouse which consists of making fact tables and dimension tables, as well as data modeling using a star schema. The last stage is exploration of report results and analysis to design processed data reporting or visualize business intelligence, and enter data into the data warehouse, so that sales performance analysis results can be seen based on available data.

This research was conducted using a qualitative approach with a case study method. The data collection technique used in this research is literature study and documentation. The stages of the research can be seen in Figure 3.

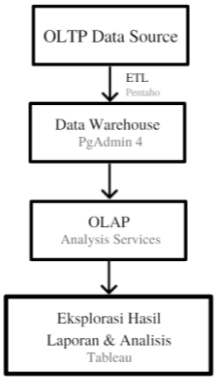

Fig. 3. The stages of the research

The data warehouse design uses a nine-step methodology with snowflake schema modeling which con-sists of one fact table and many dimension tables as a result of the normalization process. The nine-step methodology consists of:

1. Choose a process
2. Choose the grain
3. Identification and adjustment of dimensions
4. Choose facts
5. Save the pre-calculation in the fact table
6. Complete the dimension table
7. Selecting database duration
8. Trace the dimension change slowly, then identify the type of dimension change as follows.
9. Determine the priority and query mode

IV. Results

A. Data Warehouse Design

The design of the data warehouse is carried out using the nine-step methodology with nine stages as follows.

1. Choose a process: Based on the scope, this research selects the process of recording sales transaction data and the process of reporting sales transactions. The list of order data attributes can be seen in Table 1.

TABLE 1. Order Data Attribute List

No	Attribute	Description
1	order_id	Order code as the primary key as well as relate to the dimension table of the order of goods
2	customer_id	The customer code corresponding to the customer dimension table
3	order_status	Order status consisting of approved, canceled, created, delivered, invoiced, processing, shipped and unavailable
4	order_purchase_timestamp	Date and time of purchase order
5	order_approved_at	Order approval date by seller
6	order_delivered_carrier_date	Order delivery process date by logistics partner
7	order_delivered_customer_date	Order delivery date to customer
8	order_estimated_delivery_date	Shows the estimated delivery date before the order is automatically cancelled

Tiara Difa Anggrainy and Agustin Rusiana Sari, "Implementation of Extract, Transform, Load on Data Warehouse and Business Intelligence Using Pentaho and Tableau to Analyze Sales Performance of Olist Store," *International Research Journal of Advanced Engineering and Science*, Volume 7, Issue 2, pp. 368-374, 2022.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

2. Choose the grain: Grains are selected based on report needs, namely reports containing order quantity information that can be analysed based on item orders, customers, products, product name translations, sellers, order payments, order reviews, and geolocation.
3. Identification and adjustment of dimensions: The dimensions used consist of 8 dimensions, namely the dimensions of goods orders, customer dimensions, product dimensions, product name translation dimensions, seller dimensions, order payment dimensions, review dimensions and geolocation dimensions.
4. Choose facts: The fact chosen in this study is the number of orders.

TABLE 2. Dimension Table Detail

Dimension	Attribute	Description
Order Items	Order ID Item Order ID Product ID Seller ID Delivery Date Price Shipping Load	Reports can be generated based on the delivery date, price and delivery charge
Customers	Customer ID Customer Unique Id Customer Postal Code Customer City Customer Province	Reports can be generated by zip code, city and province
Products	Product ID Product Category Name Product Name Length Product Description Length Product Photo Quantity Product Weight Product Length Product Height Product Width	Reports can be generated based on product name, product name length, product description length, product photo quantity, product weight, length, height and width
Products Name Translation	Product Category Name (in Portuguese) Product Category Name (in English)	Reports can be generated by product category name in English
Sellers	Seller ID Seller's Zip Code Seller City Seller Province	Reports can be generated by zip code, city and province
Order Payments	Order ID Payment Order Type of payment Installment Payment Payment Value	Reports can be made based on the order of payment, payment type, installment and payment value
Reviews	Review ID Order ID Review Score Review Title Message Review Date Review Created Date of Review Replied	Reports can be generated based on the score, title and message of the review, the date the review was created and the date the review was replied to
Geolocation	Geolocation Postal Code Latitude Longitude	Reports can be generated based on latitude, longitude, city and province geolocation
	Geolocation City Province Geolocation	

5. Save the pre-calculation in the fact table: The calculation in the fact table of orders is the number of orders which shows the number of products ordered.
6. Complete the dimension table: The description of each predetermined dimension table can be seen in Table 2.
7. Selecting the duration of the database: This study uses order data for the data warehouse from 2016 to 2018 with the number of transactions of almost 100,000 data.
8. Track dimension changes slowly: Attributes that may change are product category, seller, customer, payment type, delivery date and review score. The Olist Store data warehouse uses the third type, namely the dimension attributes that have changed giving rise to alternatives so that the old and new attribute values can be accessed together on the same dimension.
9. Determine the priority and query mode: Every month updates and passwords will be given to the DBMS as an effort to maintain security, as well as to protect Microsoft Excel files that are used as OLAP.

B. Data Warehouse Modeling

The modeling used is a snowflake schema with one fact table, namely the core order table and 8 dimension tables, namely the goods order table, order payment table, customer table, order review table, seller table, geolocation table, product table and product name translation table. The relationship be-tween one fact table and 8 dimension tables is depicted in the snowflake schema which can be seen in Figure 4.

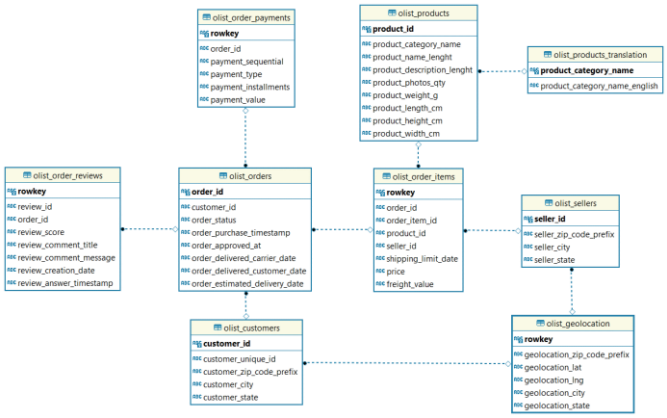

Fig. 4. Snowflake Schema

C. Data Cleansing

The data cleaning process is done by creating a Python script, where the script can clean data such as changing the name of the province in the buyers table and geolocation table. The name of the province contained in the data is an abbreviation or initial of the name of the province, such as the province of Sao Paulo which is written with "SP".

Data cleaning is done by creating a Python script using Jupyter Notebook. Jupyter Notebook is a popular tool for data processing with Python programming and can integrate scripts with output in one interactive page. The installation of Python with Jupyter Notebook can be done by installing on Anaconda, where Anaconda provides installation packages for the Python programming language and Jupyter Notebook.

Tiara Difa Anggrainy and Agustin Rusiana Sari, "Implementation of Extract, Transform, Load on Data Warehouse and Business Intelligence Using Pentaho and Tableau to Analyze Sales Performance of Olist Store," *International Research Journal of Advanced Engineering and Science*, Volume 7, Issue 2, pp. 368-374, 2022.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

D. ETL Process

The process of making ETL is done using Pentaho version 7 which has elements to create ETL, namely Job, Transformation and Database Connection. Jobs have the function of scheduling, maintaining and managing the transformation process, while transformation is the transformation process that will be carried out by the Job. Database Connection is used to make a connection to the database that is used as the Olist Store data warehouse, namely PostgreSQL. The stage of making the ETL process begins with making a connection to the PostgreSQL database with the database name that has been created, namely "oliststore". After the connection is made, proceed to the stage of making the transformation, where the entire series of processes in data management to loading data into the data warehouse is made at this stage. The entire series of ETL processes in managing Olist Store sales transaction data can be seen in Figure 5.

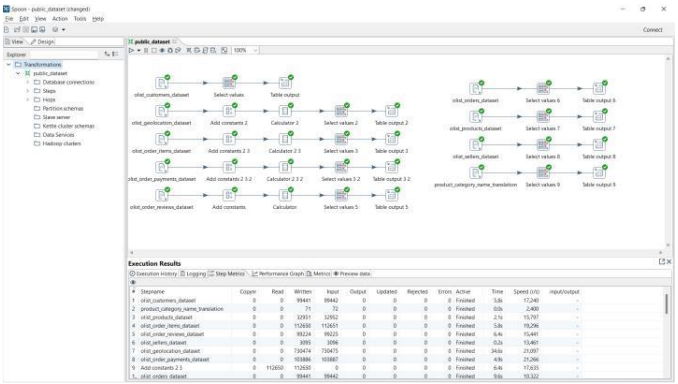

Fig. 5. ETL Process

The results of the entire series of ETL processes that have been run can be seen in the PostgreSQL database using PgAdmin4 as can be seen in Figure 6.

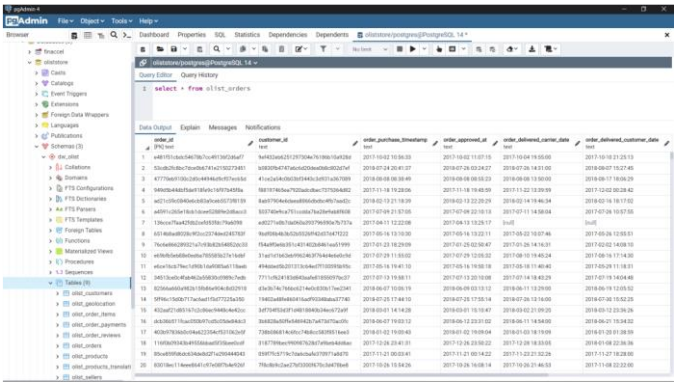

Fig. 6. Olist Store Data Warehouse

E. OLAP

The next process is the application of OLAP, which is a process to request data in a complex form and analyse the large volume of data. In this study, the OLAP process connects the server connection between the Tableau Desktop application to the oliststore database in PostgreSQL as the data source. The process of connecting the server connection from Tableau Desktop to the oliststore database in PostgreSQL can be seen in Figure 7.

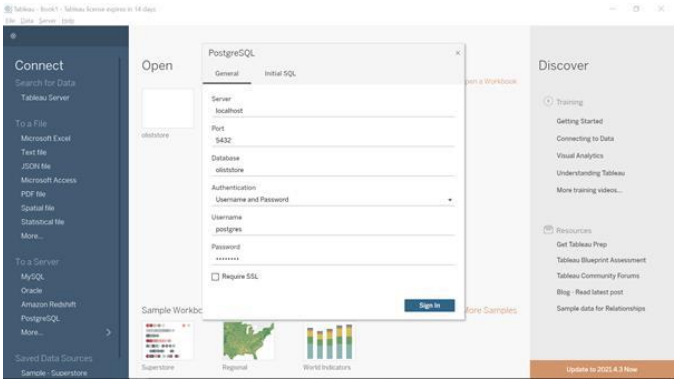

Fig. 7. Tableau Desktop Connection

F. Exploration of Report Results and Analysis

There are several data segmentations that can be used as reports in assessing sales performance at the Olist Store, which consist of sales segmentation, product category segmentation, payment type segmentation, profit segmentation, order assessment segmentation, customer city segmentation and buyer city segmentation. Exploration of report and analysis results or what can be called business intelligence is visualized in the form of a smart dashboard created using Tableau. Visualization of business intelligence can be seen in Figure 8.

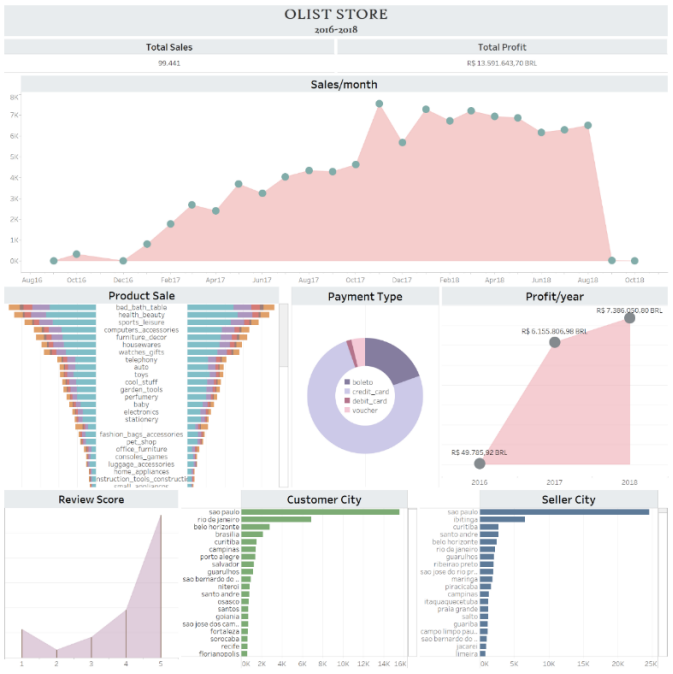

Fig. 8. Business Intelligence Visualization

Based on the results of the visualization of business intelligence that has been made, the monthly sales segmentation experienced a significant increase in December 2017, then slowly experienced a decrease in sales. It can be concluded that customer interest in ordering products at the Olist Store is slowly decreasing, so the management, especially in the field of

Tiara Difa Anggrainy and Agustin Rusiana Sari, "Implementation of Extract, Transform, Load on Data Warehouse and Business Intelligence Using Pentaho and Tableau to Analyze Sales Performance of Olist Store," *International Research Journal of Advanced Engineering and Science*, Volume 7, Issue 2, pp. 368-374, 2022.

marketing, needs to make new innovations to attract customers' interest in ordering products at the Olist Store. Total sales within two years (2016-2018) reached 99,441 sales.

In product category segmentation, product sales performance analysis is carried out based on customer ratings. After analysing product sales, it can be concluded that the products that have the highest number of orders and quite a number of customers give good value are the bed bath table product category or products related to tables, bathrooms and bedrooms. A visualization that displays the results of the analysis on product sales segmentation has been sorted by product category that has the highest number of orders to the least demand. In this case, the management needs to update the products in terms of quality and price, or it can also be done promotion of product categories that are less desirable.

Next is the payment type segmentation which has been visualized using a pie chart with details on the type of payment. Based on the visualization that has been made for the segmentation of the type of payment that is most in demand is a credit card, where by using a credit card payment can be made in installments. Therefore, the management can improve the business by collaborating with various banks to support the payment process by customers. This makes it easier for customers to make transactions with credit cards from any bank without restrictions, so that buyer interest will increase.

Profit segmentation is visualized annually based on the number of orders and in Brazilian Real (R$). Based on the data that has been visualized from 2016 to 2018, the profits that Olist Store gets from sales continue to increase. For the management, this can be used as a reference in increasing profits in the following year by applying a better sales method.

Rating segmentation with a score of 5 is the highest score given by customers for their satisfaction with product orders. The thing that needs to be a concern for management is that the assessment with a score of 1 or the worst score given by the customer looks more than a score of 2, 3 or 4. So, the management needs to minimize the possibility of a score of 1 increasing by increasing the quality of sales or product quality.

The last is the customer city segmentation shows that the most interest is in the cities of Sao Paulo and Rio De Janeiro, while in the city segmentation the sellers are in the same city, namely Sao Paulo, but the second most is in Ibitinga. The management needs to do product marketing to several other cities to attract customer interest, so that more customers from other cities are also interested in Olist Store.

V. Conclusion & Recommendation

A. Conclusion

Based on the results of the implementation of ETL on data warehouse design and visualization of business intelligence on the Olist Store, the following conclusions can be drawn.

1. Olist Store data warehouse has been generated using a PostgreSQL database with PgAdmin 4 which is useful as a well-integrated data center. The creation of this data warehouse was successfully carried out by implementing the ETL process, where there is a process of extracting, transforming and loading data into the target database.
2. A smart dashboard has been generated that provides information related to Olist Store sales performance with a data period from 2016 to 2018. The smart dashboard was created using the Tableau Desktop application with the report results in the form of sales performance graphs.
3. The results of the analysis of business intelligence visualization on Olist Store sales data with the data period from 2016 to 2018 that there is an increase in the number of orders every year, especially from 2016 to 2017 which has increased significantly. However, from 2017 to 2018 it can be said that it has slowly decreased slightly, so the management needs to improve the quality of Olist Store sales. There are also several segmentations that are analysed and visualized, so it can be concluded that the sales quality of the Olist Store is quite good, but it is very necessary to improve it to maintain sales stability and prevent a decrease in customer interest in the Olist Store. Things that may be done by the Olist Store management in improving the quality of sales and maintaining business stability are reviewing the quality and price of products at the Olist Store, making product marketing programs to various cities and being able to collaborate with various banks to support the customer payment process.

B. Recommendation

Based on the research that has been carried out, it is hoped that the development of a data warehouse at the Olist Store in the future is to apply the use of a cron job to run the Python script in the initial data cleaning automatically, because the sales transaction data per day is generated in large quantities. This cron job functions as a scheduler to run Python scripts at a predetermined time and allows automation of system maintenance, disk space monitoring and backup monitoring.

References

[1] Wijaya, Wayan M. *Teknologi Big Data: Sistem Canggih Di Balik Google, Ya-Hoo!, Facebook, IBM (Teori Hingga Tutorial)*. Nilacakra, 2019, https://www.google.co.id/books/edition/Teknologi_Big_Data/VeNDwAAQBAJ?hl=id&gbpv=1&dq=big+data+adalah&printsec=frontcover.

[2] Nordeen, Alex. *Learn Data Warehousing in 24 Hours*. Guru99, 2020, https://www.google.co.id/books/edition/Learn_Data_Warehousing_in_24_Hours/wgf9DwAAQBAJ?hl=id&gbpv=0.

[3] Warella, Samuel Y., Erika Revida, Leon A Abdillah, Delyana R Pulungan, Sukarman Purba, Erwin Firdaus, Diena Dwidienawati Tjiptadi, Muhamad Faisal, Lie Darwin, Marisi Butarbutar, Iskandar Kato. *Penilaian Kinerja Sumber Daya Manusia*. 1st Editio, Yayasan Kita Menulis, 2021, https://www.google.co.id/books/edition/Penilaian_ Kinerja_Sumber_Daya_Manusia/Vmo1EAAAQBAJ?hl=id&gbpv=1&dq=penilaian+kinerja&printsec=frontcover.

[4] Rismawati, and Matalata. *Evaluasi Kinerja Penilaian Kinerja Atas Dasar Prestasi Kerja Berorientasi Kedepan*. 1st Editio, Celebes Media Perkasa, 2020, https://www.google.co.id/books/edition/Evaluasi_Kinerja_Penilaian_Kinerja_Atas/ni9tDwAAQBAJ?hl=id&gbpv=1&dq=penilaian+kinerja&printsec=frontcover.

[5] Loshin, David. *Business Intelligence: The Savvy Manager's Guide*. 2nd Editio, Elsevier Science, 2012, https://www.google.co.id/books/edition/Business_Intelligence/L7SLNlS1ao8C?hl=id&gbpv=1&dq=business+intelligence+is&printsec=frontcover.

[6] Sutanta, Edhy. *Sistem Basis Data*. TIGA Ebook, 2004, http://grahailmu.co.id/.

[7] Qalam, Yance Ibnu. "Hubungan Data Warehouse Dengan Business Intelligence Dan ETL." *Kepo.Co*, 2020, https://kepo.co/hubungan-data-warehouse-dengan-business-intelligence-dan-etl/.

[8] Andoyo, Andreas, Elisabet Yunaeti Anggraeni, Ahmad Khumaidi, Adi Prasetia Nanda, Agus Suryana, Sucipto, Andino Maseleno, Panji Andhika Pratomo, Suyono, Satria Abadi. *SISTEM PENDUKUNG KEPUTUSAN Konsep, Implementasi & Pengembangan*. Penerbit Adab, 2021, https://www.google.co.id/books/edition/SISTEM_PENDUKUNG _KEPUTUSAN_Konsep_Implem/YTgmEAAAQBAJ?hl=id&gbpv=1&dq=ETL+adalah&pg=PA167&printsec=frontcover.

[9] Rangkuti, Yulita Molliq, Said Idrus Iskandar Al, and Dewan Dinata Tarigan. *Pengantar Pemrograman Python*. Media Sains Indonesia, 2021, https://www.google.co.id/books/edition/Pengantar_Pemrograman _Python/2ftLEAAAQBAJ?hl=id&gbpv=1&dq=python+adalah&printsec=frontcover.

[10] Prasetia, I. Putu Widia, and I. Nyoman Hary Kurniawan. "Implementasi ETL (Extract, Transform, Load) Pada Data Warehouse Penjualan Menggunakan Tools Pentaho." *TIERS Information Technology Journal*, vol. 2, no. 1, 2021, pp. 39–47, doi:10.38043/tiers.v2i1.2844.

[11] Udayana, Gede Acintia, I Made Yoga Mahendra, I Kadek Anom Sukawirasa, Gde Deva Dimastawan Saputra, and Ida Bagus Made Mahendra. "Implementasi Data Warehouse Dan Penerapannya Pada PHI-Minimart Dengan Menggunakan Tools Pentaho Dan Power BI." *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana)*, vol. 10, no. 1, 2021, p. 163, doi:10.24843/jlk.2021.v10.i01.p19.

[12] Marbun, Ivan Rivaldo, and Ramos Somya. "Perancangan Data Warehouse Untuk Data Transaksi Penjualan Menggunakan Schema Snowflake Studi Kasus : Online Market Dataset." *Universitas Kristen Satya Wacana*, vol. 5, no. 1, 2021, pp. 87–91.

[13] Wieder, Bernhard, and Maria Luise Ossimitz. "The Impact of Business Intelligence on the Quality of Decision Making - A Mediation Model." Procedia Computer Science 64 (2015): 1163–1171.