Business Intelligence For Local Mining Company Acknowledgement Reporting System

DivyaYadav[a], Dr. Naveen Choudhary[b]

[a]*College of Technology and Engineering, MPUAT Udaipur, Rajasthan, India*
[b]*Professor and Head, Department of Computer Science and Engineering,College of Technology and Engineering, MPUAT Udaipur, Rajasthan, India*

Article History: *Do not touch during review process(xxxx)*

_____

Abstract: The association of mining service firms ensures that every mining service firm (their members) works in accordance with the legislation and constitution. The issue in this association is that the job is to create queries for every data from tables in MySQL to create reports quickly, efficiently, and automatically for the government,domesticinvestorsand internationalinvestors, and any institutions that need information such as market needs, resource developments, and improvement of mining technologies. For example, how many metals are produced and invested in a month or year; metal trading, metal cost, for each mining company; and the amount of project in mining companies based on how many metals are extracted.Data Warehouse is a recommended solution because it has structured storage and can analyse data with IN-MEMORY OLAP, which can calculate live data and process and visualise data at high speeds.The design used the Customised Kimball Method to build the Data Warehouse, so this association company will be able to generate reports based on the needs of each situation.

Keywords: Data Warehouse, Customised Kimball, Power BI, Report

_____

1. Introduction

Companies of mining service have a lot of potential, with private service clients and governmental service clientsinvestments increasing every year. This was also linked to the rate of growth of mining service companies on a small, medium, and large scale.You need to expand the capacity building of your organization by monitoring market needs, changing resources, and improving building technology. The company's expansion opens jobs, improves public infrastructure, and gives local mining an opportunity to compete with local and international markets. Therefore, mining service companies are one of the factors affecting development that are no less important than the development of education, economy and culture. This information is a valuable asset for any business, as knowledge plays an important role in decision-making for business analysts. For this reason, these assets must be secure and easily accessible when needed in the event of a security breach. However, information is difficult to extract due to a lot of data available. A data warehouse is a phenomenon in which a large amount of data from a certain period of time is processed into information in order to use the data for reporting purposes.In this case study by the Federation of Mining Services Companies, the association strives to nominate and register mining companies that have been certified by the Ministry of Public Works and Public Works since 2002. The association currently lacks a rapid reporting system. An efficient and automated way to generate reports from a database. Due to this limitation, the time required to generate the report is relatively long. Currently, all reports are generated by manually exporting the MySQL database to a Microsoft Excel file and then directly analyzing the Excel file. For this reason, you need a data warehouse to produce reports that are good, fast, and easy to understand.By implementing the data warehouse and IN-MEMORYOLAP analysis according to the mapping requirements, the results of the analysis can represent data from different reports more quickly when the assignment needs to make data-driven decisions. To make a decision, you need to analyze only a small amount of historical data to support or disagree with the hypothesis and make a decision based on the analysis. The methodology used in this solution uses a Customised Kimball method. Customised Kimball method is an approach that collect transaction data from all sources using queries to analyze data. Some of the terms used in data and transaction analysis technology are OLTP (Online Transaction Processing), OLAP (Online Transaction Processing), and IN-MEMORY OLAP. OLTP and OLAP play important roles in analyzing transactions and data. IN-MEMORY OLAP plays an important role in data analysis. This is the process of calculating live data directly from memory. These ideas for Ralph Kimball were still in use in the current implementation of a data warehouse.

2.Related Work

Business Intelligence (BI) is the skill or knowledge of how to process collected and stored data to provide organizations with information to help them make decisions. Business intelligence needs data and tools to store and process historical data for decision making. Help your business grow by making decisions based on business data.

A data warehouse is a relationship between databases that are built for query operations to facilitate reporting and data analysis. Data warehouses help companies analyze trends in corporate data warehouses over a period of time. The main function of a data warehouse is to facilitate the organization in long-term strategic planning based on the data stored and making quick decisions to address problems. Modeling dimensions are used to implement a data warehouse. Dimensional modeling is a data structure technology that is optimized and created to read, summarize, and analyze numerical information such as values, sums, and weights. There are few restrictions required to process this data in a simplified view.

A star chart is a template that has been traditionally used to represent multi-dimensional data containing two types of tables. B- Fact tables and multi-dimensional tables. These tables are de-formatted to form a star-shaped Entity Relationship Diagram (ERD). This diagram is drawn using the fact table centered in some dimension tables. A facts table is a table that contains facts. In this case, all historical data is used in the analysis process. In this case, the dimension table is a table containing the attributes that support the facts.

Data warehouses, or what is commonly referred to as Online Transaction Processing (OLAP), are market-driven compared to traditional OLTP. OLAP is a prerequisite for creating a Decision Support System (DSS) that uses this tool to make complex decisions. OLAP can be combined from many data sources to create multi-dimensional reports that are used for data analysis and decision making. There are three variations of OLAP extensions: relational OLAP (ROLAP), multidimensional OLAP (MOLAP), and hybrid OLAP (HOLAP). ROLAP acts as an interface between the Relational DBMS and OLAP users. MOLAP stores data from an OLTP database into cubes. HOLAP is a combination of the two.

OLAP Cube is an OLAP feature that groups all dimensions into a fact table. This helps you create and extract data in a multidimensional manner according to the hierarchy of each dimension to get the relevant information according to the level of detail you choose. A data cube stored in a cell is similar to the structure of a three-dimensional matrix. The process used in this method is the grouping process used to group the cells into a cube. The total operations used are SUM, AVG, COUNT, MIN, and MAX. An example of syntax used for an aggregate operation is AggFunction (CubeName, Measurement) [By Dimension *].

ETL (Extract, Transform, and Load) is a series of steps to process data from a database that includes the process of data extraction. Maintaining data quality, maintaining data standards, and re-converting data into a specific format (converting) and submitting it to a data warehouse (upload).

Pentaho is an open source application for performing business analytics, ETL, dashboards, and reporting tasks. Pentaho facilitates these business analysis tasks with an integrated solution for creating and developing business analytics projects. Pentaho's ETL process uses Pentaho Data Integration to simplify data extraction and transformation. In designing OLAP, Cube uses Mondrian software in its Pentaho Workbench blueprint. Pentaho Report Designer is used to design reports from the ETL process.

Qlik Sense is an interactive application for analyzing and visualizing large amounts of data to create presentations, dashboards, and reports. This app is easy to use due to its drag and drop functionality. Qlik Sense is available as a desktop application (Qlik Sense Desktop) and cloud hosting (Qlik Sense Cloud) and is free to use. For commercial use, we offer Qlik Sense Enterprise. This removes the limitations of the free version.

3.Proposed Method

This part provides steps to help you troubleshoot partnership issues, which we'll discuss later in this paper. Start with data collection and analysis, star schema design, ETL process to create OLAP IN-MEMORY cubes, and finally create reports and dashboards.

Start collecting data. Analyze all the data that the case study supports to get information about the calculations and metrics presented in the OLAP Analysis, report design, and metrics presented in the dashboard. This data serves as a reference for developing the data warehouse. The data source used to develop the data warehouse comes from an operational database derived from a Relational Database Management System (RDBMS).

Data extraction steps are performed using these data sources to extract data from the OLTP database into the IN-MEMORY OLAP database. This data is converted into filtered data that is used for analysis. Data is broken down into facts and dimensions. Dimensions are used to create categories that gather data from facts. On the other hand, facts serve as the main source of information to be analyzed. The data design is shown in a star schema in Figure 1. The star chart consists of one fact table and two dimension tables.
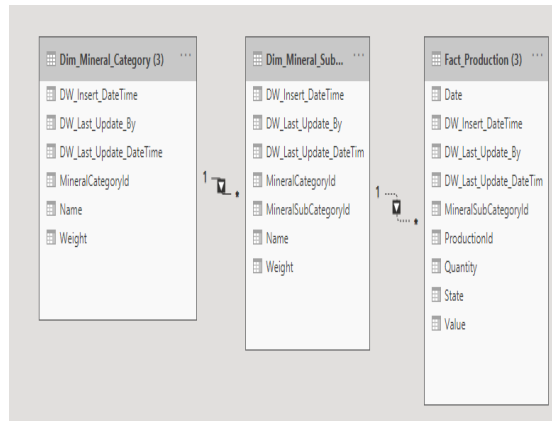
Figure 1: star schema

Power BI is used to implement ETL steps based on star schema data design facts and dimensions. The ETL process begins with selecting a table containing the required data and extracting the data from the OLTP database. Then select the data columns required for the IN-MEMORY OLAP analysis. The data is identified by an alternate key that represents the identifier of each piece of data. Finally, this data is stored in a table in the OLAP database. Figures 2, 3, and 4 illustrate ETL steps for creating a sequential dimension table. Get started with mining companies, categories, states, dates, times, and more.
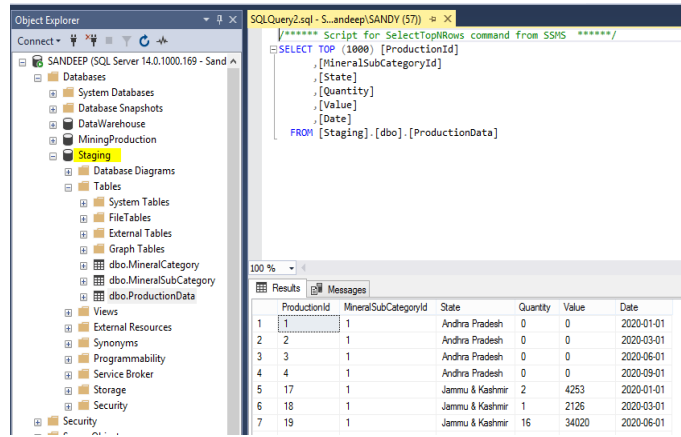

Figure 2: Source Data
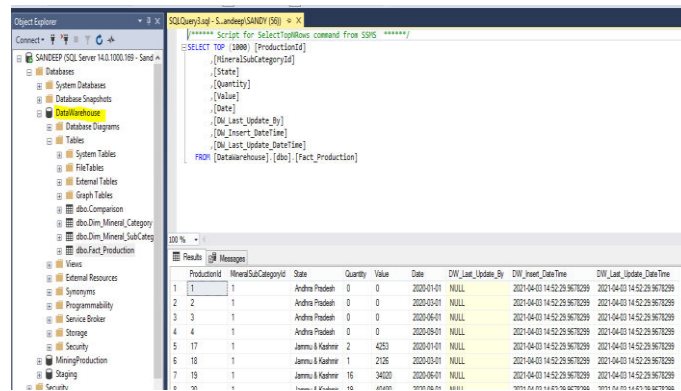
Figure 3: Data load into Staging area

Figure 4: Data load into Data Warehouse

For fact tables, use each alternate key in the dimensional table to associate related data. The data associated with the alternate key is chosen as a sample fact sheet data. As with dimension tables, you need to define columns to define the key values needed for your analysis. In this step, you can use the IN-MEMORY OLAP database not only for live calculations, but also for dashboards and reports.

The next step is to create an IN-MEMORY OLAP cube. The design of this case study is shown in Figure 5. The result of this IN-MEMORY OLAP cube design is saved as an XML file, also known as a Mondrian file.
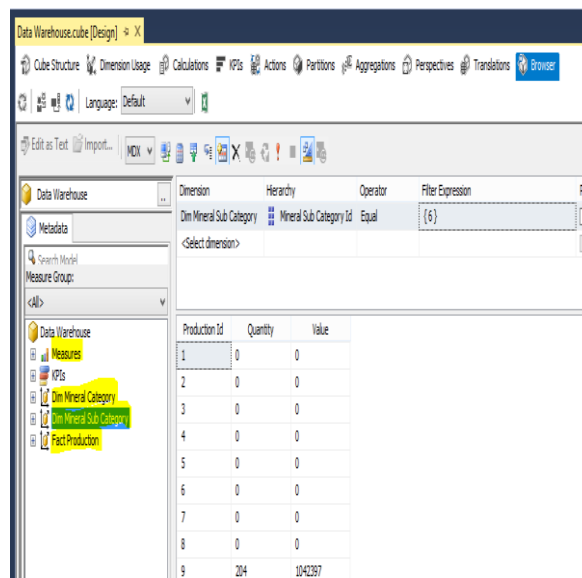

Figure 5: IN-MEMORY OLAP Cube

4.Discussion

Introducing a data warehouse application to generate reports and dashboards to support mining company association decision-making. The purpose was to better organize the challenging nature of the report manually. This method used the ETL process and IN-MEMORY OLAP to process the data in the production database into an easy-to-analyze star schema. The data is displayed with reports and dashboards that administrators can review and select.

5.Result

The Reports and Dashboards allows you to manipulate the data in your IN-MEMORYOLAP database into easy-to-understand charts or tables. Graphs that summarize data based on specific criteria make the report easier to understand. If you need to analyze the details of your data, dashboards provide a more comprehensive visualization.

In this case study, analysis will be done to find out summary of data with conditions:
● In what month mining projects have been done in a year.
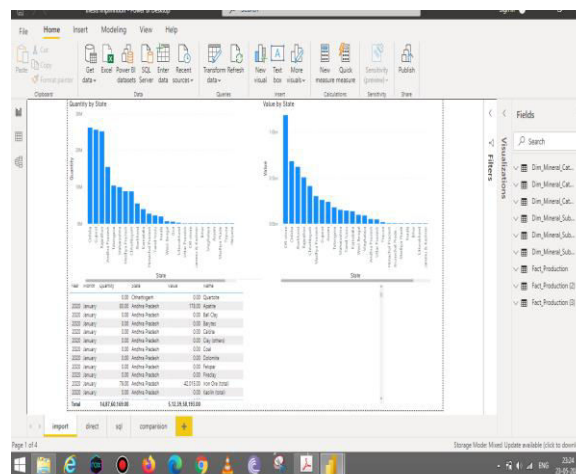●In which state minerals are produced more.

Figur 6: Reports of mining company

Each of these terms is represented in reports and dashboards. Figure 6 shows the month in which the mining company extracted most of the minerals. Figures 7 and 8 show the total processing time between the proposed and existing method (the traditional method and the Kimball method). Figures 7 and 8 clearly show that data processing takes less time as compared to the existing methods.
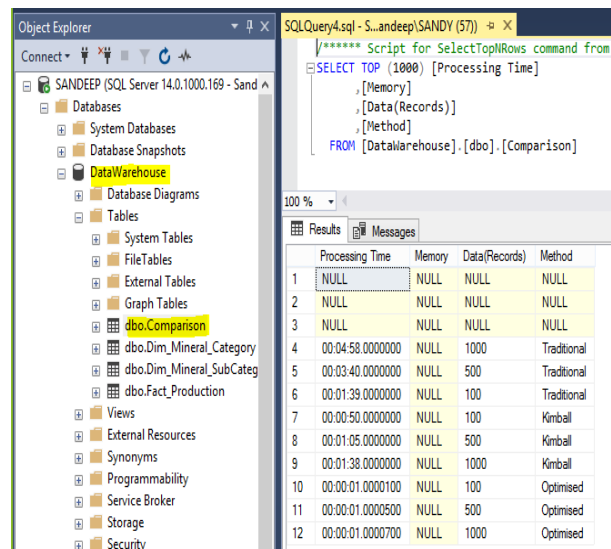

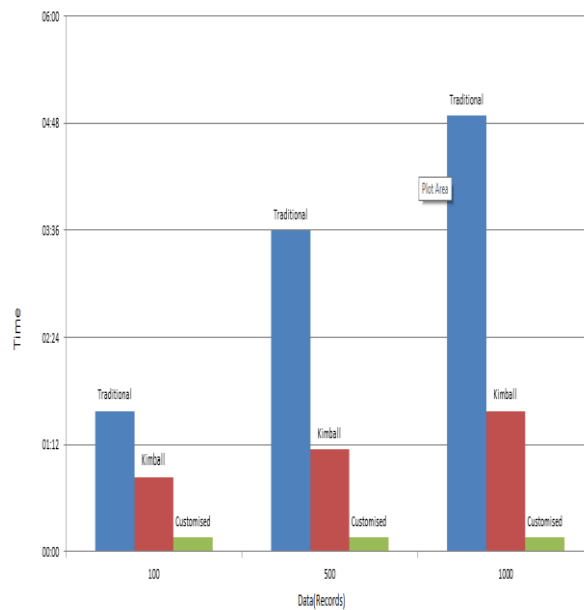Figure 7: Processing time between existing method and proposed method

Figure 8: Graph of processing time between traditional, Kimball and customised(proposed) method

6.Conclusion

Dashboards and reports make data analysis easier. Data warehouse provides the ability to process fact data with multiple dimensions in an IN-MEMORY OLAP database. The results of data processing can be displayed in various types of graphs, tables, and dashboards. The combination of reports and dashboards has made it easy for business analysts to make correct and fast decisions to solve problems.

For future work, we plan to create a decision support system to help managers and senior executives make decisions from these reports and dashboards. We will also explore several applications that support real-time operational data to create real-time business intelligence.

References

Girsang(a), A. S., Andita, E., Isa, S. M., Purnama, A. and Samuel, F., "Business intelligence for physical examination platform service reporting system," in Sustainable Information Engineering and Technology (SIET), 2017 International Conference, 2017.

Girsang(b), A. S., Satya, D., Isa, S. M., Fariz, S. Al, Susilo, B. and Ramdani, D., "Decision support system using data warehouse for hotel reservation system," in Sustainable Information Engineering and Technology (SIET), 2017 International Conference, 2017.

Goswami, R., Bhattacharyya, D. K., Dutta, V. and Kalita, M., "Approaches and issues in view selection for materialising in data warehouse," International Journal of Business Information Systems, vol. 21, no. 1, pp. 17-47, 2016.

Gounder, M. S., Iyer, V. V. and Mazyad, A. Al, "A survey on business intelligence tools for university dashboard development," in Big Data and Smart City (ICBDSC), 2016 3rd MEC International Conference, 2016.

Gudivada, V. N., Irfan, M. T., Fathi, E. and Rao, D. L., "Cognitive analytics: Going beyond big data analytics and machine learning," in Handbook of Statistics,vol. 35, Elsevier, 2016

Hart, R. and Kuo, A. M. H., "Meeting Health Care Research Needs in a Kimball Integrated Data Warehouse," in Data Science and Advanced Analytics(DSAA), 2016 IEEE International Conference, 2016.

Kimball, R. and Caserta, J., The data warehouse ETL toolkit: practical techniques for extracting, Cleaning, Conforming, and Delivering Data, 2004.

Larson, D. and Chang, V., "A review and future direction of agile, business intelligence, analytics and data science," International Journal of InformationManagement, vol. 36, no. 5, pp. 700-710, 2016.

Sidi, E., Merouani, M. El and Amin, A. A. El, "The impact of partitioned fact tables and bitmap index on data warehouse performance," International Journal of Computer Applications, 2016.

Vargas, V., Syed, A., Mohammad, A. and Halgamuge, M. N., "Pentaho and Jaspersoft: a comparative study of business intelligence open source tools processing big data to evaluate performances," International Journal of Advanced Computer Science and Applications, vol., 2016.

Wang, Z., Chu, Y., Tan, K. L., Agrawal, K. L. and Abbadi, A. E., "HaCube: extending MapReduce for efficient OLAP cube materialization and view maintenance," in International Conference on Database Systems for Advanced Applications, 2016.