THE UNIVERSITY OF ECONOMICS AND LAW

FALCUTY OF INFORMATION SYSTEMS

FINAL PROJECT REPORT

BUSINESS INTELLIGENCE AND DECISION SUPPORT SYSTEMS

TOPIC: BUSINESS INTELLIGENCE SOLUTION FOR PRODUCTION DEPARTMENT IN ADVENTURE WORKS CYCLES

Lecturer: MA. Nguyen Van Ho
MA. Le Ba Thien

Group Elite:

1. Nguyen Tran Thuy Quynh K204060307
2. Hoang Thi Thanh Phuong K204061445
3. Nguyen Hoang Tinh K204061451
4. Nguyen Thi Huyen Thuong K204061450

Ho Chi Minh City, May 8, 2023

ACKNOWLEDGMENT

We would like to express our sincere thanks to Mr. Nguyen Van Ho and Mr. Le Ba Thien for creating conditions for us to carry out this project as well as supporting us during the completion of the project.

Through the learning process of "Business Intelligence and Decision Support System", we not only receive new knowledge about Business Intelligence but also learn new skills and practical experiences. We really appreciate the lessons that the two teachers have taught, it will definitely be a source of motivation for us to continue to improve and develop more in the future.

During the implementation of the project, the team also encountered many difficulties and pressures in terms of time and knowledge. However, the team members tried to complete the project together, although there were still many shortcomings. We hope to receive comments and suggestions from you so that this project can be perfect.

Finally, to each member of the team, thank you for devoting time and effort on this project despite the tight demanding schedule to make this project a success.

Sincerely thank.

COMMITMENT

We ensure that this report “Business Intelligence Solution for Production Department in Adventure Works Cycles” is the original work that we researched and wrote. All the sources we have used or quoted have been indicated and acknowledged by complete references.

AUTHORS

TABLE OF CONTENTS

ACKNOWLEDGMENT i
COMMITMENT ii
TABLE OF CONTENTS iii
LIST OF FIGURES vi
LIST OF TABLES vii
LIST OF ACRONYMS viii
CHAPTER 1: INTRODUCTION 1
1.1 Business case for the project 1
1.2 Objectives of the project 3
1.2.1 General Objective 3
1.2.2 Specific Objectives 3
1.3 Research Objects 4
1.4 Scope of the project 4
1.5 Value and desired outcome of the project 5
1.6 Structure of the project 5
CHAPTER 2: THEORETICAL BASIS 8
2.1 Overview of Business Intelligence 8
2.1.1 What is Business Intelligence? 8
2.1.2 Business Intelligent Architecture 9
2.1.3 Advantage of Business Intelligence in enterprises 10
2.1.4 BI Strategy for Business 11
2.2 Extract Transform Load Process 12
2.2.1 What is ETL? 12
2.2.2 Why do we need ETL? 12
2.2.3 ETL Process 13
2.3 Data Warehouse and Data Mart 14
2.3.1 What are Data Warehouse and Data Mart? 14
2.3.2 Who needs Data Warehouse and Data Mart? 15
2.3.3 Advantages and Disadvantages of Data Warehouse 16
2.3.4 Galaxy Schema 17

2.4 Key Performance Indicator 18
2.4.1 KPI Definition.....18
2.4.2 The advantages and disadvantages of KPIs.....18
2.4.3 Categories of Production KPIs 19
CHAPTER 3: REQUIREMENTS ANALYTICS AND INTRODUCTION TO BI SOLUTION.....20
3.1 Production business processes 20
3.1.1 Production Department.....20
3.1.2 The purpose of Production.....20
3.1.3 Production process.....21
3.2 Business Requirements Analysis 22
3.2.1 Production quality control 22
3.2.2 Production performance control 23
3.2.3 Inventory control.....24
3.3 Data Source and Challenges.....25
3.3.1 Data source 25
3.3.2 Challenges.....26
3.4 IT requirements Analysis (IT & Infrastructure) 27
3.4.1 Microsoft Office.....27
3.4.2 Microsoft SQL Server.....27
3.4.3 Visual Studio Community 2019.....28
3.4.4 Power Level 28
3.5 Data visualization tools 28
3.5.1 Surveying and Evaluation.....28
3.5.2 Proposing BI solution for the project.....31
CHAPTER 4: BUILDING DATA WAREHOUSE AND INTEGRATING DATA...32
4.1 Designing Data Warehouse 32
4.1.1 Bus Matrix 32
4.1.2 Master Data.....33
4.1.3 Transaction Data 33
4.2 ETL processes 44
4.2.1 Dimension Table's ETL Process 44
4.2.2 Fact Table's ETL Process.....45

CHAPTER 5: RESULTS – DATA ANALYTICS AND VISUALIZATION 46
5.1 Data analytics with SSAS technology 46
5.2 Report and dashboard systems (structure) 48
5.3 Data analysis with Power BI 48
5.3.1 Dashboard about Product Quality Control 48
5.3.2 Dashboard about Production Performance 53
5.3.3 Dashboard about Inventory 58
5.4 Evaluation and discussion 63
CHAPTER 6: CONCLUSION AND FUTURE WORKS 66
6.1 Results 66
6.2 Limitations 67
6.3 Future works 68
REFERENCES 69
PROJECT EVALUATION 70
SIMILARITY RATE CHECKING 72
EDITTING APPENDIX 73

LIST OF FIGURES

Figure 2-1: The main components of BI Architecture 9
Figure 2-2: ETL Process using Data Warehouse.....13
Figure 2-3: Example of Data Warehouse 14
Figure 2-4: Galaxy Schema.....17
Figure 3-1: Production Process 21
Figure 3-2: Production database.....26
Figure 4-1: Bus matrix 32
Figure 4-2: Results of Dim Table ETL Process.....44
Figure 4-3: Results of Fact Table ETL Process 45
Figure 5-1: Complete data analytics with SSAS.....47
Figure 5-2: Cube model 47
Figure 5-3: General product quality performance dashboard over 2011-2014.....48
Figure 5-4: Top 5 manufacturing failed reason 49
Figure 5-5: Waste cost and Actual Cost by Year.....50
Figure 5-6: Scrapped Quantity by Location.....51
Figure 5-7: List of product subcategory of others.....52
Figure 5-8: General Production performance dashboard over 2011-2014 53
Figure 5-9: Overview of Production Performance.....54
Figure 5-10: Manufacturing Time and Delay Time by year.....55
Figure 5-11: Lead Time by Product Subcategories.....56
Figure 5-12: Late time and On-time Manufacturing by Product SubCategory 56
Figure 5-13: Table of Product Subcategory 57
Figure 5-14: General Inventory Performance Dashboard over 2011 – 2014.....58
Figure 5-15: Sum of Quantity by Year and Month.....59
Figure 5-16: Inventory Value by Year over 2011 - 2014 60
Figure 5-17: Inventory Value by Location 61
Figure 5-18: Inventory Value by Location Subassembly 62
Figure 5-19: Inventory Value by Location Paint Storage.....62
Figure 5-20: List of Product Subcategory Name's Inventory Quantity.....63

LIST OF TABLES

Table 1.1: Project Business Case 1
Table 3.1: Evaluation of Data visualization tools 28
Table 4.1: Master data 33
Table 4.2: Transaction Data 34
Table 4.3: Dimension Product Tables 34
Table 4.3: Dimension ProductCategory Tables 36
Table 4.3: Dimension ProductSubCategory Tables 36
Table 4.3: Dimension ScrapReason Tables.....37
Table 4.3: Dimension Location Tables 37
Table 4.3: Dimension Date Tables 38
Table 4.3: Fact Scrap Tables 40
Table 4.3: Fact Manufacturing Tables 40
Table 4.3: Fact Inventory Tables.....42
Table 4.4: Data warehouse description 43

LIST OF ACRONYMS

Acronym	Description
BI	Business Intelligence
ETL	Extract Transform Load
OLAP	Online analytical processing
KPIs	Key Performance Indicators
SQL	Structured Query Language
SSAS	SQL Server Analysis Services
SSIS	SQL Server Integration Services
IDE	Integrated Development Environment

CHAPTER 1: INTRODUCTION

Chapter 1 introduces the business case, objectives, research objects, scope, value, and desired outcome of the project. The business case outlines the rationale and benefits of the project, while the objectives are specific targets that the project aims to achieve. The research objects provide a broader picture of the project's investigation areas, while the scope clarifies the project's boundaries. The value is assessed in terms of benefits and the desired outcome is the end result and finally defines the project structure.

1.1 Business case for the project

Table 1.1: Project Business Case

Project Business Case			
Project name	Data Analytics Project		
Project sponsor	M.A. Nguyen Van Ho, MA. Le Ba Thien	**Project manager**	Nguyen Thi Huyen Thuong
Date of project approval	April 28th	**Last revision date**	April 28th
Contribution to Business Strategy	The project aimed to improve business effectiveness by reducing the negative effects of silos that hinder communication and collaboration between different departments. By exploiting data collected from users and suppliers, the company aimed to gain a better understanding of its customers and suppliers, identify new business opportunities and potential areas for growth, and make more informed decisions. The project was an essential part of the		

	company's long-term vision for success and represented a significant investment in its future. Improving data organization and analysis, the company was positioning itself for continued growth and success in a competitive marketplace.
Options Considered	A business intelligence and decision support system
Benefits	Optimize their production processes and increase efficiency Better manage their resources and identify quality control issues and take corrective action to improve product quality Improve customer satisfaction by ensuring that products are produced and delivered on time and at a high level of quality
Timescales	The estimated time for implementing the project is 2 months.
Costs	Server = $100.000 Software licensing = $90.000 Resource labor = $200.000 Training & other misc costs = $20.000 Total estimated cost = $410.000
Expected Return on Investment Risks	The project was initiated with the aim of improving the business efficacy of the company, which is a crucial component of its long-term business strategy. Its success would enable the company to achieve its phased Production plans, which would be advantageous to its business operations. However, if the project were to fail, it could considerably disrupt the company's phased Production plans, which would subsequently affect its business operations. Therefore, it is recommended that the process of deploying the data lake is thoroughly planned, taking into account all potential risks and

	uncertainties in order to ensure its successful implementation. This would require a careful assessment of the company's current resources, as well as the identification of any potential challenges that may arise during the deployment process, so that they can be addressed proactively. By doing so, the company can ensure that the deployment process is smooth and efficient, and that it achieves its intended goals of improving business efficacy and achieving phased Production plans.

1.2 Objectives of the project

1.2.1 General Objective

The main objective of this solution is to analyze business prospects. To achieve this, we have three core objectives:

- Production quality
- Production performance
- Inventory

1.2.2 Specific Objectives

Based on the general objective, we break this down into specific objectives as follows:

- Evaluate the quality of product production, thereby giving the rate of products with common defects and their causes.
- Support decision-making on the re-evaluation of production parts (machinery; labor).
- Consider replacing or purchasing more products instead of making your own.
- Monitor production schedule to ensure products are completed on schedule
- Sticking to the schedule of each department will notice the problem in time so that it can be fixed before it is due.

- Statistics of inventory quantity and comparison with Safety Stock and Reorder Point to make shopping/production suggestions.

1.3 Research Objects

The research object of the project is the AdventureWorks Database. This database is a product sample provided by Microsoft which is used as an online transaction processing (OLTP) database. The team uses data from the Production module to analyze the company's requirements for Production Quality, Production Performance, and Inventory. The AdventureWorks Database is a comprehensive database that contains a variety of tables and data sets, such as sales, products, and customers. By analyzing these data sets, the team can gain a deeper understanding of the company's requirements and develop more effective strategies to improve production quality and performance, as well as optimize inventory management. Furthermore, the team can also identify the key factors that contribute to the company's success and develop a plan to leverage these factors for further growth and expansion.

1.4 Scope of the project

The main objective of this project is to develop a data warehouse that will extract, transform and load data from the Adventurework database. The final system will be introduced to assist the production team, production managers or higher-level positions to get helpful information and view data insights needed to support decision-making. This project is fundamental to the company's long-term business strategy as it aims to improve the business effectiveness of the company.

The data warehouse will be built using SQL Server Management Studio, SSIS, SSAS, and Power Bi technology. The data will be analyzed and transformed to provide the required insights. The insights will be presented in a user-friendly way to the decision-makers, providing them with a big picture of the company's business. By bringing decision-quality information to the right people, they can be empowered and understand how to make better decisions because all the information and everything

they need is at their fingertips. To increase the effectiveness of the system, this project will use SSIS, SSAS approaches to analyze data.

The implementation of this project will improve the company's efficiency and effectiveness, leading to better decision-making and a competitive edge in the market. This project is therefore crucial for the long-term success of the company.

1.5 Value and desired outcome of the project

The BI solution project can bring significant value to the Production subsystem of AdventureWorks in various ways. One of the main benefits is optimizing resources in the production process. By exporting dashboards that predict product production trends, production costs can be monitored to make sure they are not exceeded, and manufacturing time can be evaluated to identify areas for improvement. Additionally, senior management levels can benefit from an overview of the company's production activities, which can help them make informed planning strategies and decisions. The project can also make the inventory control system work smoother and improve automation processes in production. Ensuring the consistency of information flowing between departments in the company is another important aspect that the project aims to address.

The desired outcome of the BI Solution project is to have a streamlined data warehouse that would allow tracking and retrieving production data more efficiently. The project aims to have charts and dashboards accurately reflect the requirements of the project, helping stakeholders to have a better understanding of production data. The project also aims to provide a centralized location for data storage, which can help reduce redundancies and discrepancies in data. Overall, the BI solution project can bring numerous benefits to the Production subsystem of AdventureWorks, and the desired outcome is to have a better and more efficient production process.

1.6 Structure of the project

The project is divided into seven chapters, each with its own focus and purpose. In addition to these main chapters, the project also includes a reference section that

provides guidance throughout the project, as well as an appendix that contains supplementary material. The content of the project is organized as follows:

- Chapter 1 Introduction: The definition of general goals, particular goals, project scope, and project objectives will all be covered. In addition, the project's values and anticipated outcomes are discussed.
- Chapter 2 Theoretical basis: In this project, we have utilized a range of theories to construct the models we aimed to build. We began with a thorough review of the relevant literature to identify the most suitable theories for our purposes. Once we had identified the most relevant theories, we conducted an in-depth analysis of each one to determine how it could be applied to our project.
- Chapter 3 Requirements analytics and introduction to Business Intelligence solution: A thorough requirements analysis that considers both economic and technical aspects can help ensure that the chosen solution is not only feasible and sustainable but also aligned with the organization's goals and objectives.
- Chapter 4 Building Data Warehouse and Integrating Data: Identifying data sources, determining data granularity, defining data transformation rules, and implementing ETL processes by extracting data from source systems, transforming it to meet business rules, and loading it into the Data Warehouse
- Chapter 5 Result - Data analytics and visualization: Analysts can extract relevant and meaningful insights from data by using technology and techniques that are best suited to the type of data being analyzed. To gain clearer insights into requirements, visualize data using charts or dashboards, and utilize different business intelligence (BI) tools. This will help support accurate decision-making.

- Chapter 6 Conclusion and future works: Summarize the results and limitations of the project's implementation. Additionally, it will provide guidance for future directions.

CHAPTER 2: THEORETICAL BASIS

Chapter 2, Theoretical Basis, is a critical component of this Business Intelligence project. It provides a deep dive into the theoretical underpinnings of our work. In this chapter, we explore the fundamental concepts and models that inform our research, including business intelligence, ETL processes, KPIs, and the MDX language for analyzing multi-dimensional data and OLAP. By establishing a strong theoretical foundation, we can develop a nuanced understanding of the complex business intelligence landscape and apply this knowledge to the practical challenges encountered in our work. This chapter offers a detailed exploration of the key theories and frameworks in this field.

2.1 Overview of Business Intelligence

2.1.1 What is Business Intelligence?

Business Intelligence (BI) is a technology-driven process for analyzing data and delivering actionable information that helps executives, managers, and workers make informed business decisions [1] . Specifically, BI involves collecting, integrating, and analyzing data from various sources, such as databases, spreadsheets, and other applications by using BI tools and techniques. They include data mining, predictive analytics, data visualization, and machine learning. After that, the data is transformed into meaningful information, which is presented in the form of reports, dashboards, and visualizations. These will enable businesses to identify patterns, trends, and relationships in their data.

Enterprise activities generate huge amounts of data related to customers, suppliers, orders or other information of the business. With a volume of data that is merely raw data, it cannot help businesses see the problem and provide useful information. Therefore, the use of BI is a solution to help businesses use data effectively and easily make optimal decisions.

2.1.2 Business Intelligent Architecture

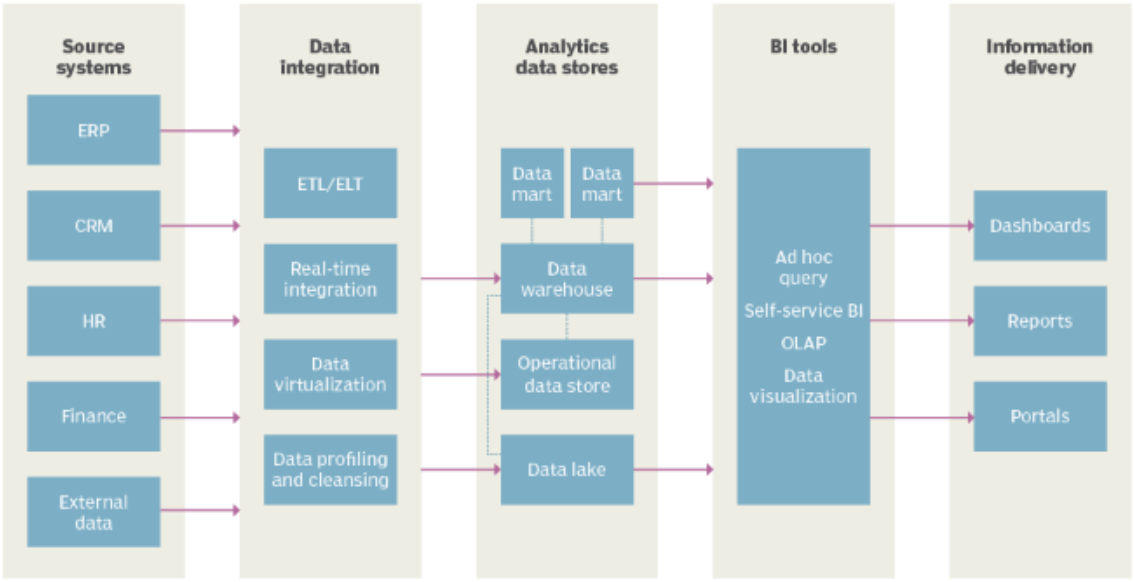


Figure 2-1: The main components of BI Architecture (Source: [2]*)*

With large amounts of data from many different sources, businesses need to handle, remove, and filter data and bring them to one source for easy control and analysis through the BI process.

In the figure 2-1, we can see overview of the main components of BI Architecture, it includes below:

- Data source: Businesses have data from various disparate sources that they need to use to satisfy their needs. This data source includes all systems and databases that store relevant data for the organization, such as customer data, sales data, production data, financial data, and more.
- Data integration: An enterprise must integrate and consolidate disparate data sets to develop unified perspectives of them in order to successfully evaluate the data acquired for a BI program. Therefore, this component is responsible for efficiently integrating data from disparate sources, transforming and cleaning data, and ensuring data quality.
- Data storage: This comprises all of the repositories where BI data is stored and managed. They are designed to store large volumes of data, usually for

long periods of time. Data Warehouses are optimized for querying and analysis, providing a single source of truth for the organization's data.

- BI tools: There are some BI tools such as Query and reporting tools, Data mining tools, Online Analytical Processing (OLAP) tools, Data visualization tools. These tools help users analyze, interpret and make sense of data, often using interactive visualizations to present data in a way that is easily understood by non-technical users.
- Information delivery: After using data visualization to provide users with predefined reports and visual dashboards, the enterprises will base on this to quickly identify insights, forecast trends, predict outcomes, and make data-driven decisions.

2.1.3 Advantage of Business Intelligence in enterprises

There are many advantages of implementing BI in enterprises that are important to the development of a business company. Some of the significant benefits are listed below:

- Data-driven decision-making: BI helps businesses make decisions based on accurate and timely data, rather than relying on intuition or guesswork. From raw data, BI makes it possible for businesses to analyze and interpret large volumes of data to identify trends, patterns, and insights that can inform strategic decisions.
- Improve operational efficiency: BI can help businesses identify issues that are causing inefficiencies and bottlenecks in their operations, allowing them to streamline processes and reduce costs. For example, a retailer can use BI to analyze sales data to optimize inventory levels and minimize out-of-stocks, providing effective sales and management solutions.
- Better customer insights: BI allows businesses to gain a deeper understanding of their customers' needs and preferences. Through analyzing customer data, businesses can identify trends, behavioral patterns, and buying habits that can inform marketing and sales strategies.

- Competitive advantage: BI can give businesses a competitive advantage by allowing them to make better decisions faster than their competitors. By leveraging BI, businesses can gain insight into emerging trends, market shifts, and customer preferences, allowing them to stay ahead of the competition.
- Improve financial performance: BI can help businesses improve financial performance by identifying areas where they can reduce costs, increase revenue, and optimize their operations. With data-driven decision-making, businesses can maximize profits and minimize risks.

2.1.4 BI Strategy for Business

The process of implementing and developing BI solution consists of 7 steps as follows:

- Step1: Collect and analyze user requirements.

  At this step, they will gather requirements from stakeholders, identify data needs and reporting needs, define a project charter and scope as well as metrics and KPIs.
- Step 2: Analyze system status and existing data.

  To understand current system and data status, they need to assess existing systems and data sources, identify limitations and gaps in data, identify data quality issues and develop a data governance plan.
- Step 3: Analyze and build data warehouse.

  In this step, they first identify the data sources needed, after that design the dimensional data model in order to build a data warehouse.
- Step 4: Analyze and build Extract Transform Load (ETL) process.

  To do ETL process, they identify and extract data from various sources, transform the data to fit the data warehouse schema and load the transformed data into the data warehouse.
- Step 5: Testing, checking metrics and processes.

At this testing step, they will test the ETL process and data quality, validate metrics and KPIs, verify the accuracy of the data.

- Step 6: Building analytical models, multidimensional database models (cubes)

  Referring to this step, they design and build multidimensional database models (cubes) and create OLAP cubes for ad-hoc queries and reporting.

- Step 7: Performance, analysis, results.

  In the final step, they focus on monitoring performance of the BI solution, analyzing the data using the BI tools, producing reports and visualizations, then evaluating the results and making improvements.

2.2 Extract Transform Load Process

2.2.1 What is ETL?

Extract Transform Load (ETL) is a process used in data warehousing to extract data from various sources [3], transform it to a desired format, and load it into a target database, data warehouse, or other data repository for analysis and reporting.

In this report, our group uses the SQL Server Integration Services (SSIS) tool to design the ETL process.

2.2.2 Why do we need ETL?

For businesses that want to manipulate data to produce useful information, ETL is important to filter out unused data as well as control and manage the data needed for them to analyze well. There are following reasons why we need ETL:

- ETL allows organizations to gather data from multiple sources, transform it into a common format, and load it into a target system where it can be analyzed by users.
- ETL helps to ensure that data is accurate, consistent, and up-to-date. This is important because data quality issues can lead to errors and inconsistencies in analysis, which can have serious consequences for business decisions.
- ETL can improve performance by reducing the amount of data that needs to be processed and analyzed. By transforming the data into a format that is

optimized for analysis, ETL can reduce the processing time and improve the speed of data retrieval.

- ETL can help organizations save time and effort by automating the process of collecting and transforming data from multiple sources efficiently and effectively.
- ETL is a critical component of business intelligence, providing the foundation for data analysis and reporting. By extracting, transforming, and loading data into a target system, ETL enables enterprises to analyze, derive insights from their data and make informed business decisions.

2.2.3 ETL Process

The figure 2-2 below shows an overview of the ETL process. It contains: Extract, Transform, Load.

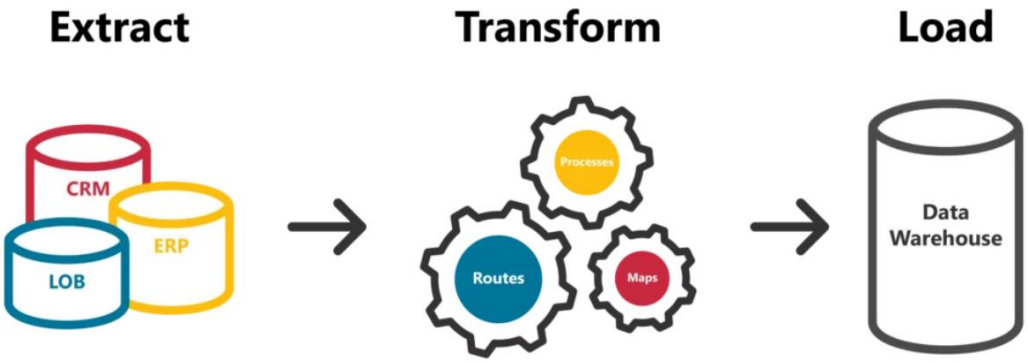


Figure 2-2: ETL Process using Data Warehouse (Source: [4]*)*

The ETL process is following these steps:

- **Extract:** Extraction is the initial stage in the ETL process. In this step, data is extracted from one or more source systems. This could include data from databases, flat files, web services, or other sources. The data may be extracted using SQL queries, APIs, or other methods. Because the extracted data is in various formats and can be corrupted, it is critical to extract it from various source systems and store it first in the staging area rather than directly in the data warehouse. As a result, importing information directly into the data warehouse may cause it to be damaged, and rollback will be

much more difficult. Therefore, this is one of the most crucial processes in the ETL process.

- **Transform:** Once the data has been extracted, it is transformed to fit the requirements of the target system. This could include cleaning the data, removing duplicates, standardizing data formats, or performing calculations or aggregations. Data may also be enriched by combining it with data from other sources.
- **Load:** This is the final step. The transformed data is then loaded into the target system, such as a data warehouse, database, or data lake. This could involve inserting the data into tables, updating existing data, or creating new data structures.

2.3 Data Warehouse and Data Mart

2.3.1 What are Data Warehouse and Data Mart?

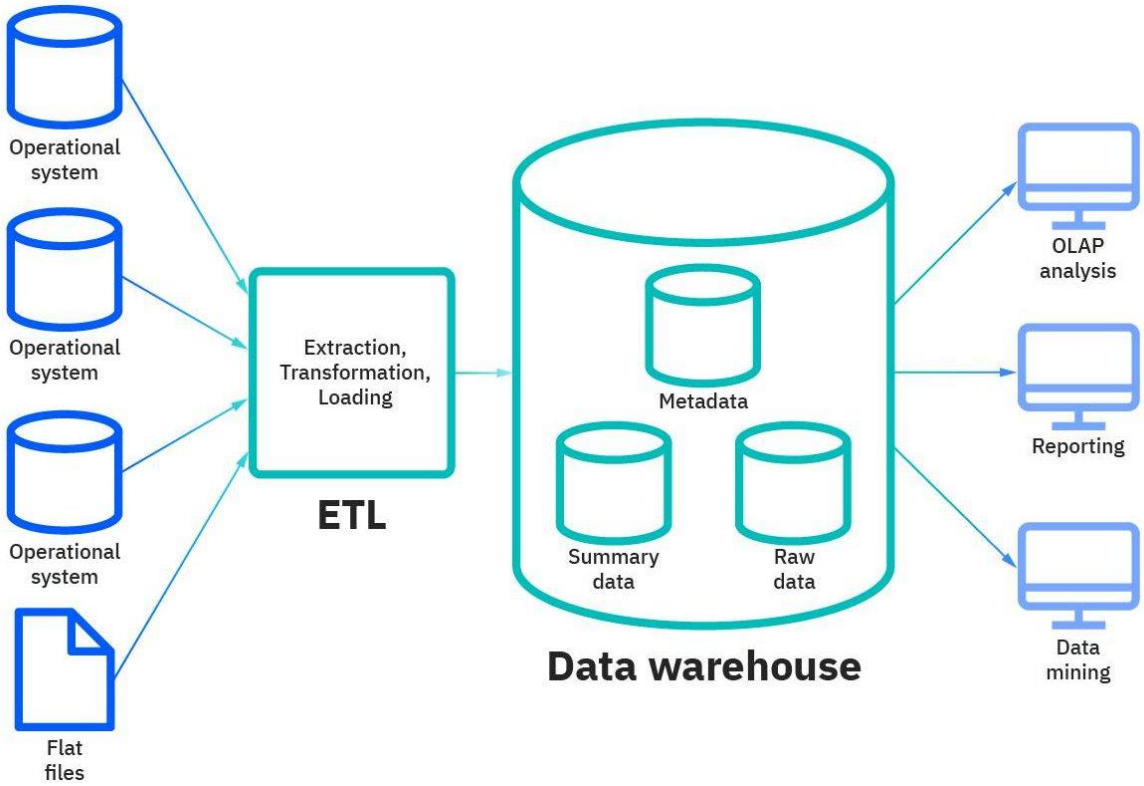


Figure 2-3: Example of Data Warehouse (Source [5]*)*

Data Warehouse is a collection of processes that collect, manage, and store historical and current data from a variety of sources to provide insights and support business decision-making. It is also an electronic repository of large amounts of business information designed for querying and analysis rather than transaction processing. Data warehouse is also the core of the Business Intelligence system built to analyze and report data for businesses. A data warehouse is a permanent, integrated, time-based, subject-oriented collection of data to support management decision-making.

The architecture of the Data Warehouse is structured in a separate layer model. The top tier presents the resulting information through reporting, analysis, and data mining tools. The middle tier of the architecture will include the analytics tools used to access and analyze the data. The bottom tier of this architecture is the database server, where data is downloaded and stored.

Data Mart is a subset of Data Warehouse that specializes in a certain department or business function in an organization or business. Companies or businesses use Data Mart to analyze information specific to each department so that Data Mart then provides the most concise and complete summary data to help make informed and quick decisions. more quickly.

There are two main types of Data Mart: Dependent Data Mart and Independent Data Mart:

- Dependent Data Mart: will be a subset of data created directly from the enterprise's existing Data Warehouse.
- Independent Data Mart: This data collection will be the opposite of Dependent Data Mart, which is not directly using the Data Warehouse or is known as an independent data system specializing in processing and analyzing a unit or part.

2.3.2 Who needs Data Warehouse and Data Mart?

Data Warehouse and Data Mart are needed for all types of users below:

- Users who need to make decisions based on large and extensive data.

- Users use sophisticated and complex collection and analysis processes.
- Users want access to data with simple and easy-to-use analytics technology.
- Individuals and users who want to access Data Warehouse and Data Mart to be able to make decisions.
- Those who want to quickly increase productivity and performance in the face of the huge amount of data required for grid or chart reports.

2.3.3 Advantages and Disadvantages of Data Warehouse

The advantages of Data Warehouse:

- Data Warehouse allows businesses or companies to quickly and accurately access important data from many sources and places.
- Data Warehouse provides consistent information about operations with many different functions such as assisting in data reporting or special data queries,...
- Data Warehouse also has the ability to integrate multiple data sources together to reduce the overload for the production system.
- Capable of integrating many features and refactoring to make it easier for users to report and analyze data.
- Save time retrieving users' data from multiple sources by allowing critical data to be accessed from multiple sources in a single place.
- Data Warehouse also stores a large amount of historical data, making it easy for users to review, compare and analyze data based on time points and make future predictions.

The disadvantages of Data Warehouse:

- Data Warehouse is not the choice for unstructured data.
- Data Warehouse can easily become obsolete in relatively quick time.
- Changing data types and ranges, data source schemas, and data queries is very difficult in the Data Warehouse.
- For users who are still quite vague, Data Warehouse is still very complicated for them to use proficiently.,

- The range of data will always tend to increase and become increasingly complex.
- Using Data Warehouse will cost an organization or business in terms of resources and costs for training and implementation purposes.

2.3.4 Galaxy Schema

The following iteration of the data warehouse schema is the Galaxy Data Warehouse Schema, sometimes referred to as a Fact Constellation Schema. The Galaxy Schema makes use of numerous fact tables joined by shared normalized dimension tables, unlike the Star and Snowflake Schemas. Galaxy Schema may be compared to a collection of connected, fully normalized star schema that eliminates any redundant or inconsistent data.

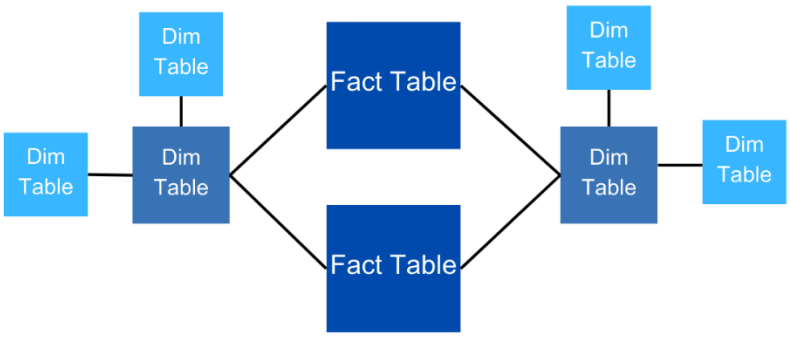


Figure 2-4: Galaxy Schema (Source: Authors)

Characteristics of the Galaxy Schema:

- For large database systems, Galaxy Schema's multidimensional nature serves as an important design factor.
- Because to normalization, Galaxy Schema decreases duplication to almost nil.

- Known for its high data quality and accuracy, Galaxy Schema facilitates efficient reporting and analytics.

2.4 Key Performance Indicator

2.4.1 KPI Definition

KPI stands for key performance indicator, a quantifiable measure of performance over time for a specific objective. KPIs provide targets for teams to shoot for, milestones to gauge progress, and insights that help people across the organization make better decisions. From finance and HR to marketing and sales, key performance indicators help every area of the business move forward at the strategic level [6].

2.4.2 The advantages and disadvantages of KPIs

There are many advantages of KPIs that help improve business performance. Firstly, Key Performance Indicators (KPIs) aim to monitor progress and provide measurable results in the form of amounts, metrics, or data. This enables individuals or organizations to easily track the development of their goals and identify areas that require more attention. KPIs can be set to provide periodic results, such as weekly or monthly, depending on the scope of the objective. Secondly, in large organizations with numerous employees, it can be difficult to monitor the progress of each individual. KPIs help to align everyone towards the same objective by making results accessible to all involved in the project. This fosters motivation and ensures that everyone is working in the same direction towards the common goal. Thirdly, tracking progress through KPIs enables supervisors to redesign or modify their plans based on the performance of previous targets. KPIs provide insight into the capacity, performance indicators, and productivity of individuals, helping organizations to establish future objectives and strategies. Finally, KPIs provide individuals with the opportunity to prove themselves and aid managers in assessing progress and rewarding them accordingly. This can motivate employees to work harder and strive for bonuses or recognition for their hard work. Additionally, KPIs allow employees to track their own performance and identify areas for improvement.

In contrast, KPIs have its disadvantage for the company. Firstly, a focus on short-term objectives can lead to a decrease in job quality. Financial objectives can take precedence over the credibility of the work, with metrics being valued more highly. Next, short-term oriented KPIs can help achieve immediate goals but may be detrimental to achieving long-term objectives. Success in one area may not be sufficient. Thirdly, standardization can reduce workers' creativity. Result-oriented goals may discourage innovative strategies. Finally, loyalty may be impacted by the focus on KPIs. By only demonstrating progress levels, it can be difficult to monitor the quality of work. This can weaken the bond between the company and the customer and may result in the company losing customers [7].

2.4.3 Categories of Production KPIs

Production KPIs can be categorized into three distinct areas.

- **Machinery or Asset Focused**: Those KPIs that are primarily responsible for monitoring the machinery include the machine uptime rate, which measures the time the machine is maintenance-free and working as expected. They also cover defects and change over time (time taken to set up the machine for another product line).
- **Unit or Product Focused**: KPIs that are focused on the products that are produced, such as the unit rate, Throughput, and cycle rate, fall into this category.
- **Cost Focused**: Costs that are attributed to the production process, such as avoided costs, staffing costs, or raw material costs would fall into this category. [8]

CHAPTER 3: REQUIREMENTS ANALYTICS AND INTRODUCTION TO BI SOLUTION

Chapter 3 will cover the details of business operations, including the production department and process, obstacles businesses face, and IT requirements. We'll delve into business requirements and data sources for precision and accuracy. We'll also discuss challenges and propose a Business Intelligence solution with data visualization tools for informed decision-making and improved operations.

3.1 Production business processes

3.1.1 Production Department

Adventure Works Cycles is a huge global manufacturing company that manufactures and sells metal, bicycles, bicycle components, and related accessories in commercial markets throughout North America, Europe, and Asia.

In the structure of this company, the Production Department is in charge of managing the manufacturing process, which includes overseeing the production schedule, inventory levels, product quality, and managing the workforce. This is the division whose activities are related to the production and assembly of products. This includes the production of bicycle frames, wheels, tires, and other components, as well as the assembly of complete bicycles. The Production department is also responsible for ensuring that the production orders are completed correctly, on time, and in a cost-effective manner.

3.1.2 The purpose of Production

The target of the Production Department in AdventureWorks is to ensure that the company's manufacturing processes are efficient, and cost-effective, and produce high-quality products that meet customer needs and expectations. It is critical to grasp the following specific purpose in order to accomplish the Production Department's tasks and responsibilities:

- Plan and schedule the production process to make sure that products are manufactured on time and in the most efficient way possible

- Manage the inventory of raw materials, work-in-progress, and finished goods, which need for the company to have the necessary resources to meet production schedules.
- Implement quality control processes in order that products meet the required quality standards and specifications.
- Manage production lines and ensure that production targets are fulfilled.
- Manage production costs which must remain profitable for the company. This involves monitoring production costs, identifying cost-saving opportunities, and optimizing the use of resources.

3.1.3 Production process

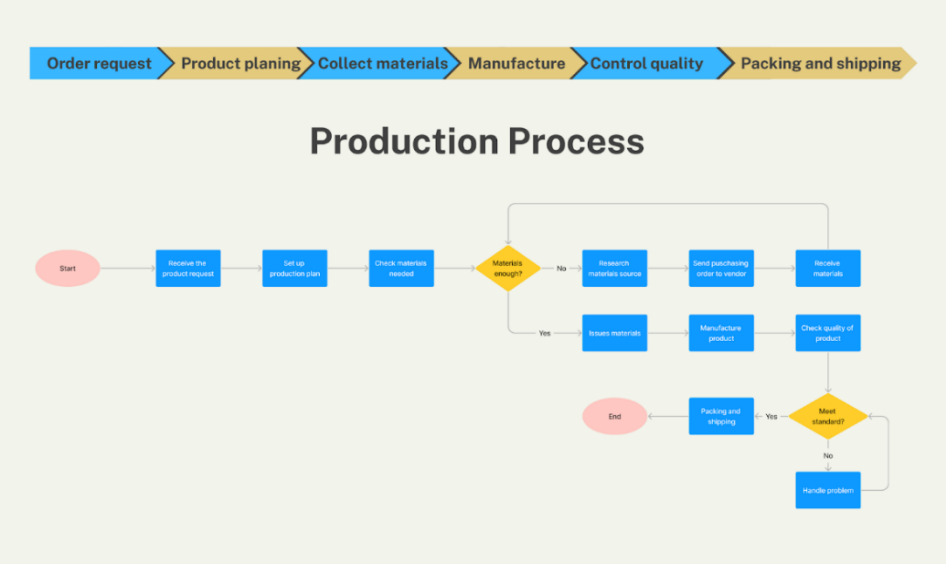


Figure 3-1: Production Process (Source: Authors)

In the Figures 3-1 about Production Process, it contains information on the steps to produce the product. In that process, it shows the direction of action in cases of insufficient raw materials or substandard products.

- Receive product order request: Orders for bicycles and related products are created by the Sales Department and sent to the Production Department. After receiving the request, The Production department reviews the orders, checks for available inventory, and determines the production schedule.
- Production planning: The Production department creates a production planning based on the available inventory, the number of orders, and the

production capacity. This plan includes production schedule, estimating the amount of raw materials needed, and creating a plan for the production process.

- Collect materials needed for production: In order to meet enough raw materials to carry out production, it is necessary to consider what materials we have and lack to produce products. After checking the quantity and condition of the materials, conduct sourcing and submit the purchase order to the supplier. This is the stage where it is necessary to ensure sufficient material requirements for the following stages.
- Issue materials and manufacture: At this stage, after satisfying the requirements, the quantity of materials required for production, the Production department issues materials to the production line. Materials include raw materials such as steel, aluminum, and rubber, as well as parts such as gears, chains, and wheels. After that, proceed to produce the products according to the required quantity.
- Control production quality: After being completed, the bicycle or bicycle components are subjected to a series of quality control examinations to ensure that it satisfies the needed criteria for safety, reliability, and performance.
- Packaging and shipping: Once the bicycles have been manufactured and quality checked, they are packaged and prepared for shipping to customers.

3.2 Business Requirements Analysis

3.2.1 Production quality control

The production department manufactures products to order. In addition to the requirement of on-time delivery, it is very important to ensure the correct delivery quantity.

Although the production department always tries to produce the correct quantity of products ordered, production problems can cause defective products. This leads to many serious problems in terms of production, sales and time and cost.

Therefore, managers need to control the company's production quality to make appropriate decisions:

- Evaluation of production quality
- Control issues that lead to defective products
- Inspect production parts and find out the main cause of defects
- Make a decision to repair the machine or replace the product if it is not possible to fix the cause of the error
- Make a decision about production in larger quantities than required or negotiate a longer delivery time if the product has a high defect rate

3.2.2 Production performance control

Evaluating the performance of the production department is essential for the company. Manufacturing performance is measured in terms of time and cost. In addition to completing the order quantity, the manufacturing department also needs to ensure that the orders are completed on time and at the expected cost. If the production is not on schedule, it will greatly affect the entire production, delivery, and sales process. This is really dangerous and affects not only the company's internal affairs but also the reputation of the company, the loss of a large number of customers as well as economic losses. In the case of production exceeding the expected cost, surely the biggest consequence is economic, affecting not only the profitability of a few orders but even the profitability of the whole company.

On the contrary, if production efficiency issues are well controlled, the company can save more time and costs, bringing economic and productivity benefits. From there, leaders can easily control production and make important decisions such as:

- Control the production schedule and the rate of completion of work on schedule to evaluate the production ability of the entire production department.

- Check production parts and specific orders that are not completed on time to find the cause. Consider whether the production department is overloaded, or because the expected delivery time is too short.
- Controlling production costs as expected or not, whether lower or higher than planned
- Make decisions in increasing/reducing costs and expected production time
- Provide solutions to reduce production time and production costs appropriately
- Make decisions to remedy in time if a production segment is behind schedule.

3.2.3 Inventory control

Warehouse management plays a very important role in every business. It governs the production and purchase and sale activities of the company. It also ensures normal business operations and can even boost corporate big business campaigns. For Adventurework, they have set safety thresholds and Re-order thresholds for each product. This is a very good point to be able to control their inventory at a safe level to ensure business activities take place at normal levels if there is not too big problem arising. However, managers can't check their inventory on a regular basis, so we need to build a real-time inventory control solution and provide solutions. detect warnings in a timely manner. From there, the leader can control and make the following decisions:

- Control the inventory quantity of products fully and continuously
- Make appropriate production or purchase and sale decisions for those inventory thresholds
- For products with high inventory, it is necessary to check the cause of inventory. If there is no influence by orders, the company needs to come up with appropriate sales strategies to consume products.

3.3 Data Source and Challenges

3.3.1 Data source

The AdventureWorks database is an extensive source of data that provides detailed information on the key transactions required by a bicycle manufacturing business. This comprehensive database includes data on various business aspects such as Sales, Purchasing, Human Resources, and Manufacturing. Our project has a specific focus on product manufacturing research and data analysis in order to meet the requirements and questions raised to achieve the project goal of improving decision making.

To achieve the project goal, we will delve into the Production Process which is one of the most crucial aspects of any manufacturing business. The original AdventureWorks database has 25 tables that provide essential information on the production process. However, we will need to supplement this data with additional research and analysis to gain a comprehensive understanding of the production process. By doing so, we aim to identify areas for improvement and optimize the manufacturing process to ensure efficient and effective production of high-quality bicycles.

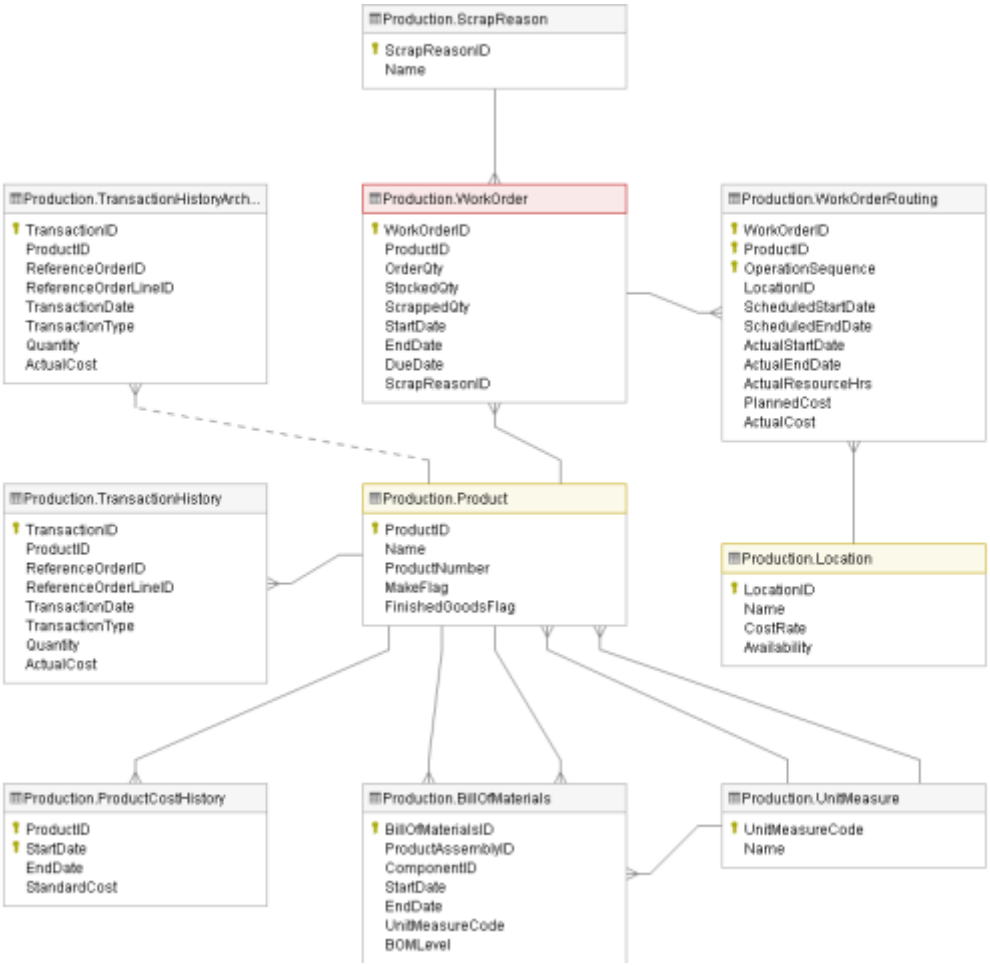


Figure 3-2: Production database (Source: [9]*)*

Our team built the database including 10 boards from the original data source including 3 fact tables and 7-dimension tables.

3.3.2 Challenges

AdventureWorks 2019 is a sample database provided by Microsoft that represents a fictitious bicycle manufacturer. Some of the challenges associated with data sources for the production module of AdventureWorks 2019 include:

- Limited data: Although AdventureWorks 2019 provides a comprehensive database, the data may not represent the complexities of a real-world manufacturing environment. The data may lack details such as machine parameters, employee performance metrics, and supply chain data.

- Data quality: The data in AdventureWorks 2019 may not be representative of real-world manufacturing data in terms of quality. The data may contain errors, omissions, and inconsistencies that can affect the accuracy and reliability of the production module.
- Data integration: The production module in AdventureWorks 2019 may require data from multiple sources, such as sales, inventory, and finance modules. Integrating data from these sources can be a challenge, as they may have different structures, formats, and levels of complexity.
- Scalability: AdventureWorks 2019 is designed to be a sample database for learning and testing purposes and may not be scalable for large manufacturing environments. In such cases, additional data sources and tools may be required to meet the scalability requirements of the production module.

3.4 IT requirements Analysis (IT & Infrastructure)

3.4.1 Microsoft Office

This is a set of applications that are widely used and grown in most countries around the world. This software is developed and founded by Microsoft. For the convenience and application of the utilities of this software for the project report, the whole team used the following three main software:

- Word: Used to write requirements, essays, reports as well as group meeting minutes.
- Powerpoint: Used to present as well as display the necessary and main content about the project for the group's presentation.
- Excel: Used to create tables or worksheets, measurement forms, resource calculations as well as project metrics for the whole team.

3.4.2 Microsoft SQL Server

To implement a subject BI project, performing data queries or building a data warehouse are indispensable core contents. The team used Microsoft SQL Server founded and developed by Microsoft for the above content.

3.4.3 Visual Studio Community 2019

Microsoft Visual Studio is an integrated development environment abbreviated as IDE developed by Microsoft. It is also known as the "World's Most Used Code Editor". With the main purpose of developing graphics, user interfaces, websites, information mobile applications and especially cloud utilities, ... With such useful features both the team used Microsoft Visual Studio Community 2019 in conjunction with Microsoft SQL Server in implementing project contents such as data analysis in the data warehouse, OLAP, SSAS and SSIS implementation.

3.4.4 Power Level

Performing the BI course project, after analyzing and querying the data, the whole team will use Power BI to visualize this data model. Power BI is a business analytics application developed and founded by Microsoft with the main purpose of providing visualization of data and business intelligence. The highlight is that Power BI allows importing data from many applications in the Microsoft system such as Excel, SQL Server,... data analysis services. After receiving information and analyzing them, Power BI will provide a variety of visualization charts from simple to complex to optimize data display for users to easily choose the right one for their needs. as well as its intended use.

3.5 Data visualization tools

3.5.1 Surveying and Evaluation

Table 3.1: Evaluation of Data visualization tools

Comparison Content	Power BI	Tableau	QlikView
Strengths	Because Power BI is developed by Microsoft, this advantage makes	Tableau's strength is its powerful and easy-to-use data visualization	The difference of Qlik compared to the other two applications is the

	Power BI can be easily combined with other applications in Microsoft such as Excel or SQL Server. Besides, Power BI can also integrate with other computing services like Azure and AWS. In addition, Power BI provides rich data visualization and features including real-time data analysis of the data.	capabilities. The highlight is that Tableau will focus on improving the user experience and increasing the scope of implementation and standardization.	ability to connect with other applications and utilities on a large scale. At the same time, QlikView will help direct the vision for the product.
Weaks	Focus on multiple Cloud Analytics (Azure). Users do not know many features when working with other products.	No complex model support yet. At the same time, the cost of application products is quite high.	The cost for operation and use is very expensive and requires users to have high skills in use. What is quite disappointing is that QlikView does not yet have self-service capabilities for analytics purposes.

The ability to analyze and exploit	- Power BI has the ability to analyze user behavior to help users better understand the relationship between user interactions and data. - Support for temporal data analysis, cloud data storage analysis and spatial data. - Support many kinds of multi-national languages to make it easier for users to use.	- Tableau also has powerful features in generating dynamic reports like Power BI. Besides, Tableau has the ability to annotate in the chart and display color details in the chart to help users have a deeper data orientation.	- In comparison to the two tools above, QlikView has less analysis and reporting features

Power BI, Tableau, and QlikView are all data analysis tools widely used in businesses and organizations. Each tool has its own advantages and applications, so the right choice will depend on the needs and requirements of each organization. For our project, we decided to use Power BI to analyze and visualize the data. Power BI is not the most powerful tool, but it is enough for us in terms of ease of use, aesthetics, and intuitiveness. For our less complicated requirements, Power BI has enough capacity to meet the needs of analyzing. In particular, Power BI is also the tool with the lowest cost of the above tools, so for a subject project, we were convinced by this condition. One key point that makes us use Power BI is its integration with other Microsoft products. During our entire process of working on this project, we used most of the Microsoft products. It was these persuasive factors that led our team to use Power BI to complete our report.

3.5.2 Proposing BI solution for the project

For data analysis in the project, we recommended using Microsoft's Power BI due to its various functionalities and benefits. Power BI is a comprehensive package that includes data extraction, exploration, interactive dashboards, and analytics tools. It is a desktop application that interacts with Power BI Service, a cloud-based SSAS report server, or Power BI Report Server, a local reporting system. Power BI Desktop is a free, stand-alone personal analysis application that allows advanced users to create complex data mixtures using on-premises data sources.

One of the main reasons to use Power BI is its cost-effectiveness. It ranks top in cost for a corporate-sized product compared to other data visualization and analysis technologies. Power BI includes most of the significant features on a regular basis, and the Pro version is affordable for businesses.

Power BI is also user-friendly and visually appealing, as its interface, procedures, and functionalities are similar to those in Excel. Users can easily convert to Power BI for repeated data processing jobs. Additionally, it offers a range of reporting capabilities that make it easier for consumers to use reports, such as Q&A features and bookmarking.

Moreover, Microsoft continuously updates and adds new data analytics features to Power BI, consolidating Azure cloud services that store data analytics in conjunction with Power BI, using Azure Machine Learning to make algorithmic analytics easier to apply, and using Power BI reporting on HoloLens' AR/VR platform. All these features make Power BI a product that is always on the cutting edge of technology and features.

CHAPTER 4: BUILDING DATA WAREHOUSE AND INTEGRATING DATA

In this chapter, the most important details when building a data warehouse and integrating data are designing the data warehouse, developing ETL processes to populate the data warehouse, cleaning and validating dimension tables, and loading fact tables. These steps will ensure that the data warehouse is properly designed and populated with high-quality data that can be used to drive business insights and decision-making.

4.1 Designing Data Warehouse

4.1.1 Bus Matrix

Bus Matrix is a very useful data warehouse design tool for organizations. Because it ensures consistency and integration across the enterprise, but also flexibility, and extensibility.

In this project, we use Bus Matrix to describe fact and dimension tables as well as the relationship between them from which to build Data Mart.

In Table 4.1 Bus matrix, each row represents a business requirement of the production department; Each column corresponds to one dimension of the business.

	DIMENSIONS								
BUSINESS REQUIREMENTS	Product	ProductCategory	ProductSubCategory	Inventory	Location	WorkOrder	WorkOrderRouting	ScrapReason	Date
Production quality control	X	X	X			X		X	X
Production performance control	X					X	X		X
Inventory control	X	X	X	X	X				X

Figure 4-1: Bus matrix (Source: Authors)

4.1.2 Master Data

Master data is consistent and continuously used data to describe the core entities of an enterprise. We have described the required master data in Table 4.1 Master data.

Table 4.1: Master data

Master data	Describe
Product	The data about products sold or used in the manufacturing of sold products. It includes the product's information like id, name, category, attributes, price, cost, and stock limit…
Product Category	The data about high-level product categorization. The company's product categories include Bikes; Components; Clothing; Accessories.
Product Subcategory	The data represents the product classification in more detail.
Location	The data about product inventory and manufacturing locations. It includes ID and location description.
Scrap Reason	The data about manufacturing failure reasons. It includes ID and failure description.

4.1.3 Transaction Data

Transaction data is data arising from purchase and sale transactions and business activities of enterprises. The characteristic of transaction data is that it

changes continuously and depends on the master data. Transaction Data in this project is shown in Table 4.2 Transaction Data.

Table 4.2: Transaction Data

Transaction Data	Describe
Product Inventory	The data about storage information of products in the warehouse as location, shelf, bin, and quantity.
Work Order	The data about manufacturing work orders. It includes information about product, quantity, scrap, and time.
Work Order Routing	The data about work order details. It indicates the manufacturing process sequence at each specific manufacturing location. It shows the production process sequence at each specific production location. It also includes actual and expected production time and costs

Fact & Dimension Tables

Table 4.3: Dimension Product Tables

Field Name	Descriptive	Type	isNull
DimProduct			
ProductKey (PK)	Primary key for Product records. Identity / Auto increment column	int	
ProductID	A code that identifies a specific product in the system. It can be a string or a numeric value.	nvarchar	

ProductSubCategoryID	A foreign key that associates a product with a specific subcategory of products.	nvarchar	yes
ProductName	The name of the product.	nvarchar	
ProductNumber	Unique product identification number	nvarchar	
Color	The color of the product.	nvarchar	
Size	The size of the product	nvarchar	yes
Weight	The weight of the product.	decimal	yes
ListPrice	The price of the product.	money	
MakeFlag	A binary flag indicating whether the product is manufactured or purchased from a supplier.	bit	
FinishedGoodsFlag	A binary flag indicating whether the product is a finished good or a component of a finished good.	bit	
SafetyStockLevel	The minimum amount of inventory that should be kept on hand to meet expected demand for the product.	smallint	
Reorderpoint	The level of inventory at which the product should be reordered.	smallint	
Standard Cost	The cost to produce or acquire the product.	money	
SizeUnitMeasureCode	The unit of measurement for the product size.	nvarchar	yes
WeightUnitMeasureCode	The unit of measurement for the product weight.	nvarchar	yes
DaystoManufacture	The number of days required to manufacture the product.	int	
ProductLine	The category of the product (e.g., bicycles, clothing, electronics).	nvarchar	yes

Class	The class of the product (e.g., economy, luxury, sport).	nvarchar	yes
Style	The style of the product (e.g., road bike, mountain bike, city bike).	nvarchar	yes
ActiveFrom	The date from which the product is available to purchase.	datetime	
ActiveTo	The date until which the product is available to purchase.	datetime	

Table 4.4: Dimension ProductCategory Tables

Field Name	Descriptive	Type	isNull
DimProductCategory			
ProductCategoryKey(PK)	The primary key of the ProductCategory table. It is a non-negative integer used to identify a specific category of products in the database.	int	
ProductCategoryID	A code that identifies a specific category of products in the system. It can be a string or a numeric value.	int	
ProductCategoryName	The name of the product category.	string	

Table 4.5: Dimension ProductSubCategory Tables

Field Name	Descriptive	Type	isNull
DimProductSubCategory			
ProductSubCategoryKey(PK)	The primary key of the ProductSubcategory table. It is a non-negative integer used to identify a specific subcategory of products in the database.	int	

ProductSubCategoryID	A code that identifies a specific subcategory of products in the system. It can be a string or a numeric value.	int	
ProductCategoryID	A code that identifies a specific subcategory of products in the system. It can be a string or a numeric value.	int	
ProductSubCategoryName	The name of the subcategory of products.	string	

Table 4.6: Dimension ScrapReason Tables

Field Name	Descriptive	Type	isNull
DimScrapReason			
ScrapReasonKey(PK)	The primary key of the ScrapReason table. It is a non-negative integer used to identify a specific reason for scrapping in the database.	int	
ScrapReasonID	A code that identifies a specific reason for scrapping in the system. It can be a string or a numeric value.	smallint	
ScrapReasonName	The name of the reason for scrapping.	string	
ActiveFrom	The date from which the reason for scrapping is active.	datetime	
ActiveTo	The date until which the reason for scrapping is active	datetime	

Table 4.7: Dimension Location Tables

Field Name	Descriptive	Type	isNull
DimLocation			

LocationKey(PK)	The primary key of the Location table. It is a non-negative integer used to identify a specific location in the database.	int	
LocationID	A code that identifies a specific location in the system. It can be a string or a numeric value.	string	
LocationName	The name of the location.	string	
CostRate	The cost rate for the location.	money	
ActiveFrom	The date from which the location is active.	datetime	
ActiveTo	The date until which the location is active.	datetime	

Table 4.8: Dimension Date Tables

Field Name	Descriptive	Type	isNull
DimDate			
DateKey(PK)	The primary key of the DimDate table. It is a non-negative integer used to identify a specific date in the calendar as a foreign key in other tables.	int	
DateID	A code that identifies a specific date in the calendar. It is in the format yyyymmdd.	string	
FullDate	The full date corresponding to the DateKey. Data type is datetime.	datetime	
DayOfMonth	The number of the day in the month, from 1 to 31	int	Yes
DayName	The name of the day of the week (Monday, Tuesday, etc.).	string	
DayOfWork	The number of the day in the week, from 1 to 7	int	
DayOfWeekInMonth	The number of this day of the week within this month, from 1 to 5.	int	yes

DayOfWeekInYear	The number of this day of the week within this year, from 1 to 366.	int	yes
DayOfQuarter	The number of this day within the quarter, from 1 to 92 or 91.	int	yes
DayOfYear	The number of this day within the year cycle, from 1 to 364 or 365.	int	yes
Month	The number of the month in the year, from 1 to 12.	int	yes
MonthName	The name of the month (January, February, etc.).	string	yes
MonthOfQuarter	The number of the month within the quarter, from 1 to 3.	int	yes
Quarter	The number of the quarter within the year, from 1 to 4.	int	yes
QuarterName	The name of the quarter (Q1, Q2, etc.).	string	yes
Year	The year, in four-digit format.	int	yes
YearName	The name of the year.	string	yes
MonthYear	The name of the month and year (January 2019, February 2019, etc.).	string	yes
MMYYYY	The month and year, in the format MMYYYY.	string	
FirstDayOfMonth	The first day of the month as a date.	date	
LastDayOfMonth	The last day of the month as a date.	date	
FirstDayOfQuarter	The first day of the quarter as a date.	date	
LastDayOfQuarter	The last day of the quarter as a date	date	
FirstDayOfYear	The first day of the year as a date.	date	
LastDayOfYear	The last day of the year as a date.	date	

Table 4.9: Fact Scrap Tables

Field Name	Descriptive	Type	isNull
FACT_SCRAP			
WorkOrderID	Primary key for WorkOrder records	int	
ProductID	A code that identifies a specific product in the system. It can be a string or a numeric value.	nvarchar	
ScrapReasonID	A code that identifies a specific reason for scrapping in the system. It can be a string or a numeric value.	smallint	
LocationID	A code that identifies a specific location in the system. It can be a string or a numeric value.	string	
OrderQty	The quantity of a product that has been ordered by a customer.	int	
StockedQty	The quantity of a product that is currently available in inventory and ready to be sold.	int	
ScrappedQty	The quantity of a product that has been removed from inventory due to damage, defects, or expiration.	int	
ActiveFrom	The date that a product is first available.	datetime	
ActiveTo	The date that a product is no longer available.	datetime	

Table 4.10: Fact Manufacturing Tables

Field Name	Descriptive	Type	isNull
FACT_MANUFACTURING			
WorkOrderID	Primary key for WorkOrder records	int	

ProductID	A code that identifies a specific product in the system. It can be a string or a numeric value.	nvarchar	
OperationSequenceID	This is a unique identifier that is used to identify a specific operation or task within a larger process or workflow. It is often used to help organize and track different stages of a project or production process	int	
StartDate	The date and time that a particular operation or task begins. It can be used to help track progress and ensure that all tasks are started on schedule.	datetime	
EndDate	The date and time that a particular operation or task is completed. It is often used to help determine whether a project or process is on schedule, and to identify any potential delays.	datetime	yes
DueDate	The date and time by which a particular operation or task is expected to be completed	datetime	
ScheduledStartDate	The date and time that a particular operation or task is scheduled to begin, based on the planned timeline or schedule.	datetime	
ScheduledEndDate	The date and time that a particular operation or task is scheduled to be completed, based on the planned timeline or schedule.	datetime	
ActualStartDate	The date and time that a particular operation or task is scheduled to be completed, based on the planned timeline or schedule.	datetime	
ActualEndDate	The date and time that a particular operation or task is scheduled to be completed, based on the planned timeline or schedule.	datetime	

PlannedCost	The date and time that a particular operation or task is scheduled to be completed, based on the planned timeline or schedule.	float	
ActualCost	The actual cost of completing a particular operation or task, as recorded during the production process	float	

Table 4.11: Fact Inventory Tables

Field Name	Descriptive	Type	isNull
FACT_INVENTORY			
ProductID	A code that identifies a specific product in the system. It can be a string or a numeric value.	nvarchar	
LocationID	A code that identifies a specific location in the system. It can be a string or a numeric value.	string	
Shelf	Storage compartment within an inventory location	string	
Bin	Storage container on a shelf in an inventory location	string	
Quantity	Quantity of products in the inventory location	int	
ActiveFrom	The date that a product is first available.	datetime	
ActiveTo	The date that a product is no longer available.	datetime	

4.1.5 Data Warehouse model (Snowflake or Star or Galaxy)

The data warehouse is created in accordance with the business's requirements and contains 3 fact tables, 6 dimension tables, and production department data organized using the galaxy model.

Table 4.12: Data warehouse description

No	Table	Describe
1	DimProduct	Products sold or used in the manfacturing of sold products.
2	DimProductCategory	High-level product categorization
3	DimProductSubCategory	Product subcategories, see ProductCategory table.
4	DimScrapReason	Manufacturing failure reasons lookup table
5	DimLocation	Product inventory and manufacturing locations
6	DimDate	Detailed time
7	FactScrap	Evaluate the quality of product production, thereby giving the rate of products with common defects and their causes, support decision making on the re-evaluation of production parts (machinery; labor), consider replacing or purchasing more products instead of making your own.
8	FactManufacturing	Monitor production schedule to ensure products are completed on schedule, stick to the schedule of each department will notice the problem in time so that it can be fixed before it is due.
9	FactInventory	To provide shopping and production recommendations, statistics on inventory volume and comparisons with Safetystock and Reordpoint are used.

4.2 ETL processes

4.2.1 Dimension Table's ETL Process

This is the process for ETL data by SSIS tool of DimProduct, DimProductCategory, DimProductSubCategory, DimScrapReason, DimLocation, DimDate:

Step 1: Create a new SSIS Project in Visual Studio.

Step 2: Add and configure an OLE DB connection manager to connect to the destination database.

Step 4: Create a Data Flow Task in the SSIS package.

Step 5: Add and configure a Flat File Source to extract data from the flat file.

Step 6: Add and configure transformation steps to transform data, we use Slowly Changing Dimension components here.

Step 7: Add and configure an OLE DB Destination to insert data into a table in the destination database.

Step 8: Check and handle errors if any.

Step 9: Run the SSIS package to perform the ETL process.

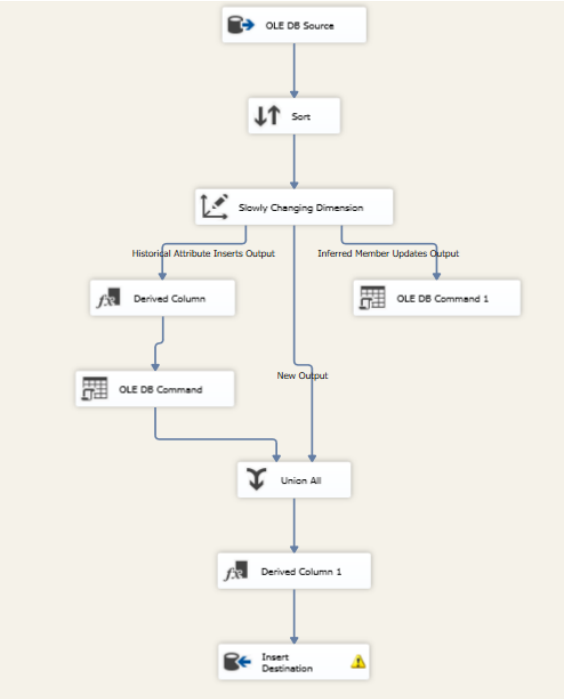


Figure 4-2: Results of Dim Table ETL Process

After following the steps, we have result as Figure 4-2.

4.2.2 Fact Table's ETL Process

To perform ETL for Fact table (Fact Scrap, Fact Manufacturing, Fact Inventory), the following steps:

Step 1: Create a new SSIS Project in Visual Studio.

Step 2: Create a Data Flow Task in the SSIS package.

Step 3: Add and configure connections to the source and destination databases.

Step 4: Add and configure an OLE DB Source to extract data from a table in the source database.

Step 5: Add and configure transformation steps to transform data, such as Derived Column, Aggregate, or Sort.

Step 6: Add and configure an OLE DB Destination to insert data into the Fact table in the destination database.

Step 7: Check and handle errors if any.

Step 8: Run the SSIS package to perform the ETL process.

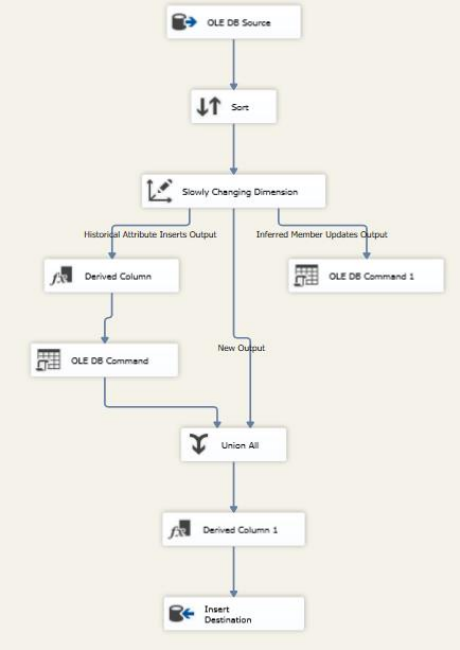


Figure 4-3: Results of Fact Table ETL Process

CHAPTER 5: RESULTS – DATA ANALYTICS AND VISUALIZATION

This chapter is a critical part of presenting research results. It focuses on data analytics and visualization, which identify patterns, trends, and relationships between variables. By using statistical techniques and software tools, we gain insights and communicates findings using charts, graphs, and tables. This chapter provides an overview of the findings, methods, and tools used to collect and analyze data. We aim to provide a clear and concise explanation of the research findings and demonstrate their significance to the research problem. This demonstrates to draw meaningful conclusions from data and support their argument with evidence-based results.

5.1 Data analytics with SSAS technology

The thesis makes use of the SSAS tool to analyze data from the DW data warehouse and create insightful reports to aid in decision-making. The Business Intelligence Development Studio (BIDS) tool that comes with SQL Server incorporates SSAS.

The cube is built in four steps:

Step 1: Launch SQL Server Management Studio and connect to the Analysis Services instance you want to use to create a Cube.

Step 2: Create a new Data Source

Step 3: Select the type of Data Source

Step 4: Create a new Cube

Step 5: Add Measures and Dimensions to the Cube.

Step 6: Create Calculations and KPIs (Key Performance Indicators) if necessary.

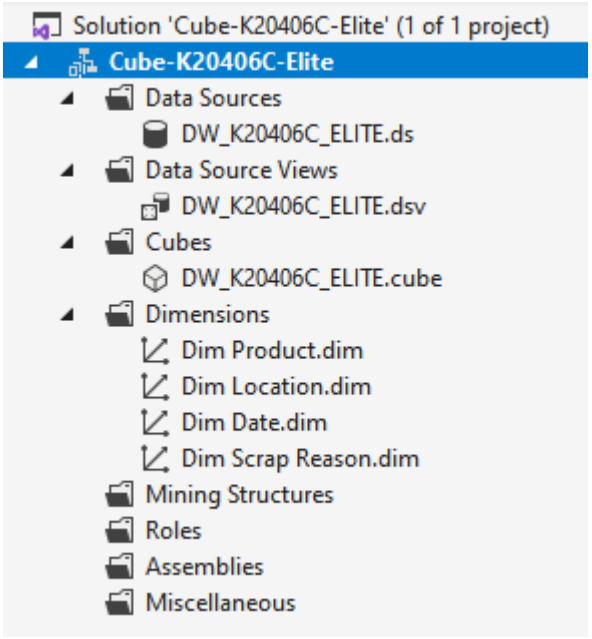


Figure 5-1: Complete data analytics with SSAS

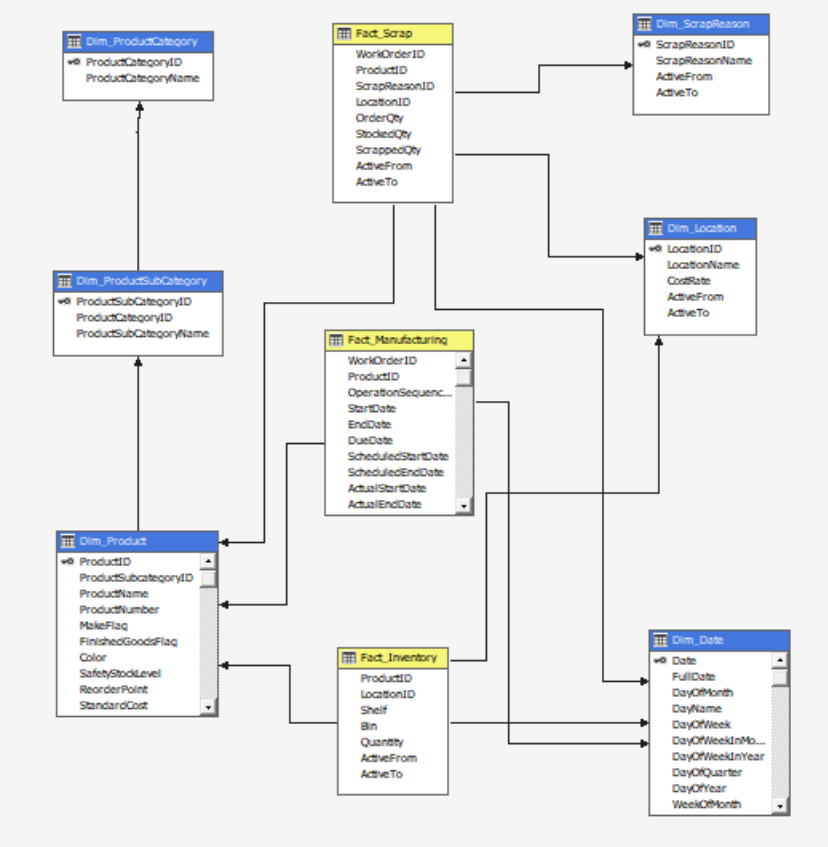


Figure 5-2: Cube model

5.2 Report and dashboard systems (structure)

The system of reports and dashboards is the result of data that has been transformed into charts that businesses can easily understand and provide useful information for them.

Our team used Power BI to analyze and present the following information:

- How is the production quality?
- What is the status and value of the products in stock?
- What are the production lead times and delays and on-time deliveries?

5.3 Data analysis with Power BI

5.3.1 Dashboard about Product Quality Control

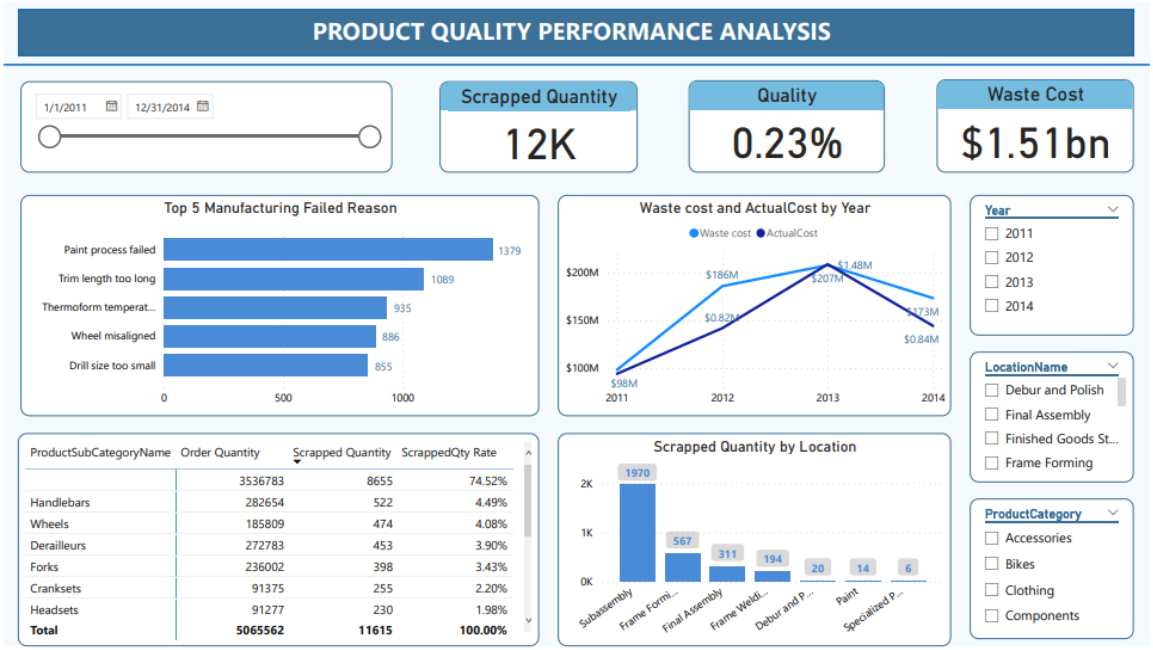


Figure 5-3: General product quality performance dashboard over 2011-2014

From figure 5-3 provides an overview of the product quality performance in the production of AdventureWorks, which includes visual charts and tables based on data from the team's data warehouse. The dashboard has been filtered by year, location name, and product category to provide specific insights into product quality.

The visual charts include a clustered bar chart that shows the top 5 manufacturing failed reasons, a line chart that shows the variation of waste cost and actual cost from 2011-2014, a column chart that shows scrapped quantity by location name, and a detailed table that shows product subcategory by order quantity, scrapped quantity, and scrapped quantity rate. These visual charts and tables help stakeholders understand the overall quality of the product, identify trends, and make informed decisions to improve the production process.

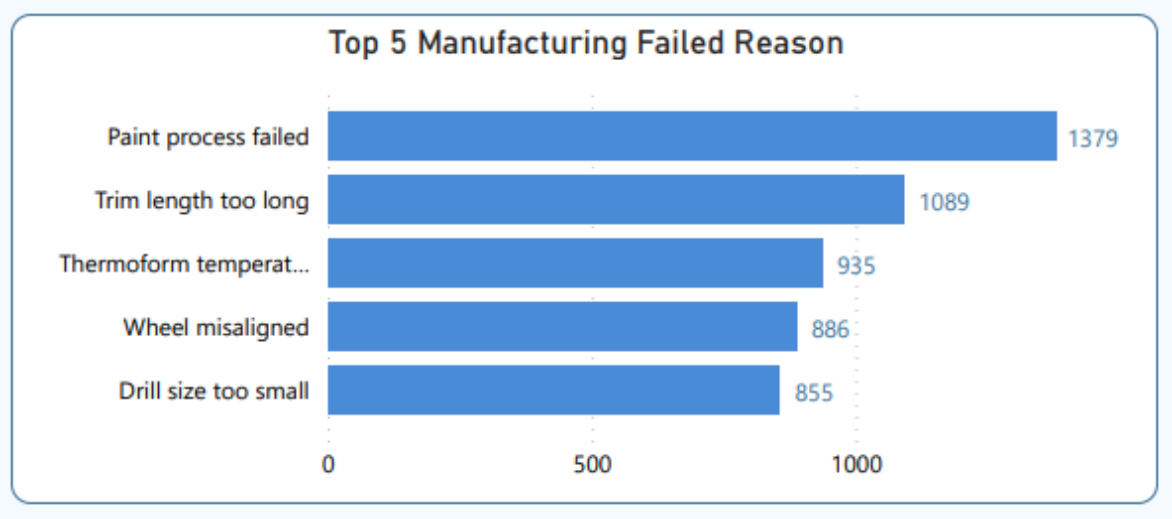


Figure 5-4: Top 5 manufacturing failed reason

From figure 5-4, we can see the clustered chart of the top 5 manufacturing reasons displays the number of defective products caused by the leading production issues. Specifically, the list includes 5 main causes: paint process failed with a total of 1379 defective products, trim length too long with 1089 products, thermoform temperature too long with 935 products, wheel misaligned with 886 products, and drill size too small with 855 products. These values are represented by vertical bars, with each bar representing a production issue arranged in descending order of the number of defective products. Displaying this information on the clustered chart helps users easily compare values and identify the main causes of product defects during the production process.

Therefore, it is necessary to address the issue of high numbers of defective products resulting from the top 5 manufacturing reasons, there are several solutions that could be implemented. Firstly, it is recommended to strengthen quality control measures and ensure that manufacturing processes are carried out correctly. This could help reduce the number of defective products as issues can be identified and corrected in a timely manner. Secondly, providing training to employees on manufacturing processes and monitoring product quality can help improve their skills. This can enhance their ability to recognize and address production issues before they become more serious. Finally, exploring technical or equipment improvements could also help reduce the number of defective products. For example, if small drill sizes are causing many products to be defective, using larger drill sizes could be considered to minimize the number of defects.

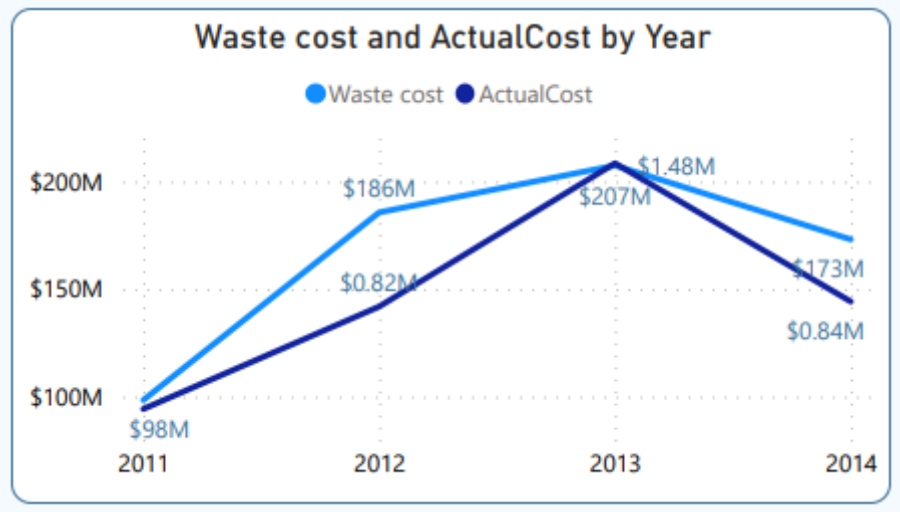


Figure 5-5: Waste cost and Actual Cost by Year

In figure 5-5, the line chart depicts the trends of Waste cost and Actual Cost from 2011 to 2014. In 2011, the Waste cost was $98M, which increased drastically in 2012 to $186M. However, the cost dropped significantly in 2013 to $1.48M, and then increased again to $173M in 2014. On the other hand, the Actual cost was $98M

in 2011, decreased significantly in 2012 to $0.82M, and then increased to $207M in 2013, and slightly decreased to $0.84M in 2014.

From the chart, we can see that there was a huge increase in Waste cost from 2011 to 2012, which then decreased sharply in 2013 before increasing again in 2014. In contrast, Actual cost showed a significant decrease in 2012 before increasing in 2013 and then decreasing slightly in 2014.

Based on these observations, we can identify several potential solutions to reduce the waste cost and actual cost. Firstly, it may be necessary to analyze the reasons for the sharp increase in waste cost from 2011 to 2012 and implement measures to reduce waste in the production process. These measures could include improving the efficiency of the production process and reducing the use of resources.

Secondly, it may be necessary to review the cost of production and identify areas where cost can be reduced without affecting the quality of the product. This could include optimizing the supply chain, negotiating better deals with suppliers, and reducing the overhead costs.

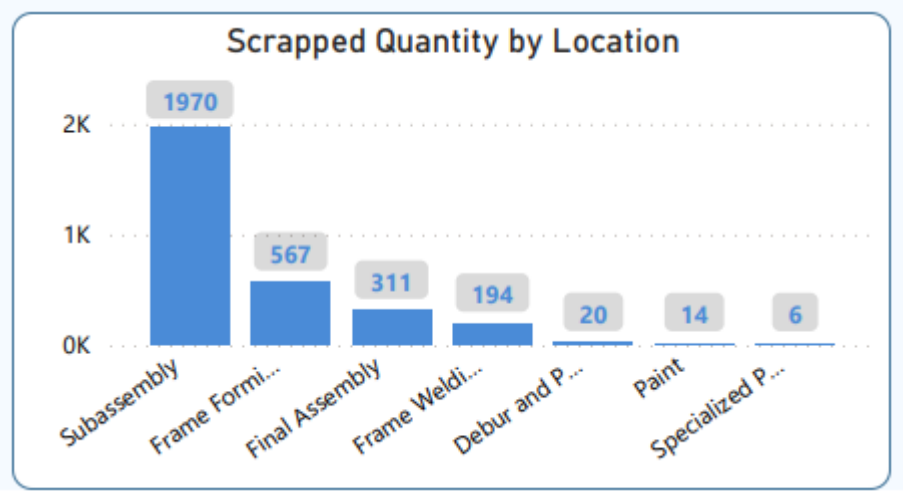


Figure 5-6: Scrapped Quantity by Location

In figure 5-6, the column chart displays the scrapped quantity by location name, including Subassembly with 1970, frame forming with 567, final assembly with 311, frame welding with 194, deburr and polish with 20, paint with 14, and

specialized paint with 6. Each column represents a location and the corresponding scrapped quantity, arranged in descending order. This chart helps users easily compare values and identify the top locations where scrap is generated in the manufacturing process.

Based on the observation above, we can propose solutions to reduce the number of defective products during the manufacturing process. Firstly, we can focus on the Subassembly area to address the issue and reduce the number of defective products. Specific measures could include enhancing quality control in this area, ensuring that manufacturing processes are carried out correctly, and training employees on the necessary skills and manufacturing processes to increase their ability to detect and repair technical errors before they become more severe.

Additionally, we can also consider technical improvements or equipment modifications to reduce the number of defective products. For example, if the subassembly part production is encountering issues due to inaccurate sizing, we could consider using new equipment or changing the manufacturing process to minimize technical errors.

ProductSubCategoryName	Order Quantity	Scrapped Quantity	ScrappedQty Rate
	3536783	8655	74.52%
Handlebars	282654	522	4.49%
Wheels	185809	474	4.08%
Derailleurs	272783	453	3.90%
Forks	236002	398	3.43%
Cranksets	91375	255	2.20%
Headsets	91277	230	1.98%
Total	**5065562**	**11615**	**100.00%**

Figure 5-7: List of product subcategory of others

Based on the figure 5-7, the given table which can be observed that the total order quantity is 5065562, while the total scrapped quantity is 11615, resulting in a

Scrapped Quantity Rate. It is noteworthy that the Bottom product subcategory has the highest Scrapped Rate of 74.52%.

To address this issue, it is recommended to focus on the Bottom subcategory and investigate the root causes of the high Scrapped Rate. This could involve conducting a thorough analysis of the production process and identifying any potential inefficiencies or defects that could contribute to the high rate of scrapped products.

Additionally, implementing quality control measures such as regular inspections and testing could help to identify and address any issues early on, before they become more significant and lead to higher rates of scrapped products. This can ultimately lead to a reduction in production costs and an improvement in product quality.

5.3.2 Dashboard about Production Performance

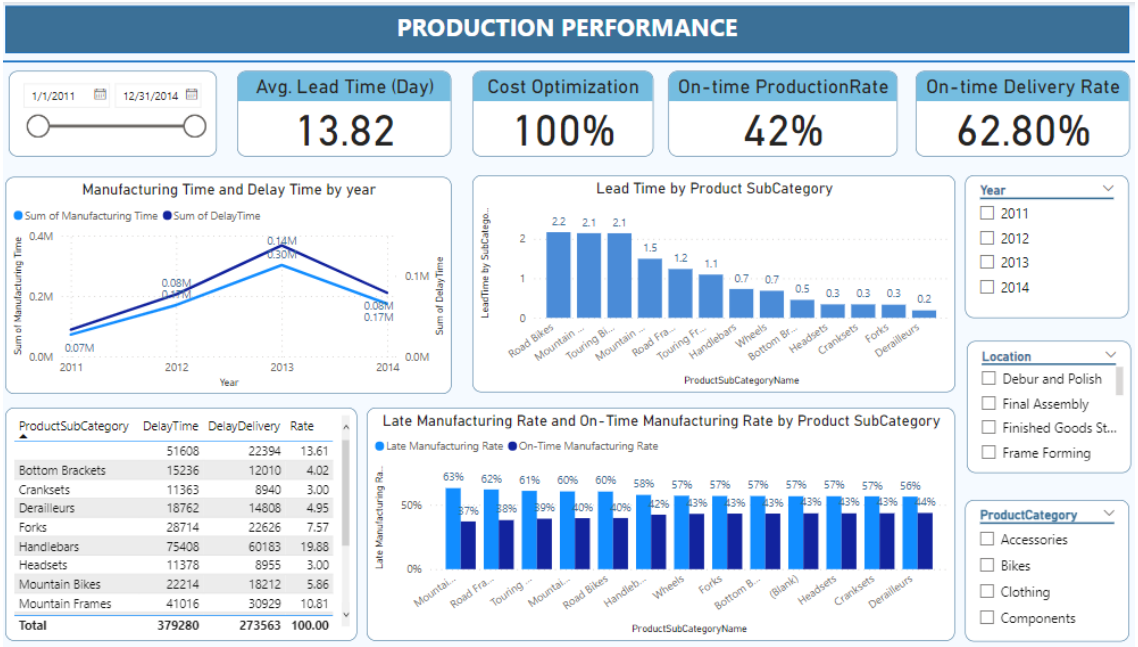


Figure 5-8: General Production performance dashboard over 2011-2014

Figure 5.8 shows the Production Performance Dashboard. Back to the requirements of the project, the manager wants to control production efficiency through two main criteria: time and cost. The dashboard has shown the production schedule values such as: Production time, production time exceeding the plan and related ratios.

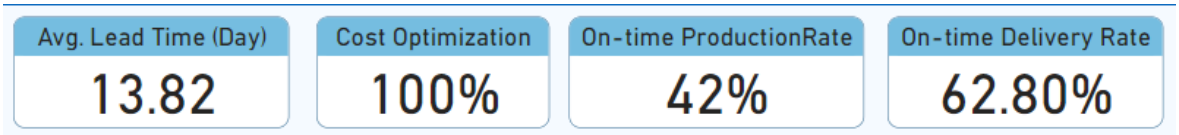


Figure 5-9: Overview of Production Performance

The general KPI values in figure 5-9 give us an overview and easy to understand about Production Performance.

- Avg Lead Time: The first value is the average lead time of an order (day). Lead time is calculated from the start of production to the end of production. Average lead time tells us about a business's ability to fulfill orders. Although the lead time of an order can vary greatly by quantity and product requirements. However, the average lead time gives an overview of current production capacity and suggests targets for shortening the lead time.
- Cost Optimization: This is the value that shows the ability of the business to control costs. Cost Optimization is calculated based on the ratio between the expected production cost and the actual production cost. Cost Optimization has a value of 100%, which means that the business is controlling its production costs extremely effectively, with no costs exceeding the expectations.
- On-Time: Production and business activities of enterprises often have problems with schedules. It can be said that this is one of the common problems and causes many impacts on businesses. The on-time production rate is calculated based on the difference between the expected production end date and the actual production date. The on-time delivery rate is calculated based on the difference between the production finish time and the delivery time.

- Currently, the enterprise has a relatively low on-time production rate of 42% and on-time delivery rate of 62.8%. Overall, this number needs further improvement. Leaders need to find the root cause of scheduling problems so that they can improve performance and gain other business benefits.

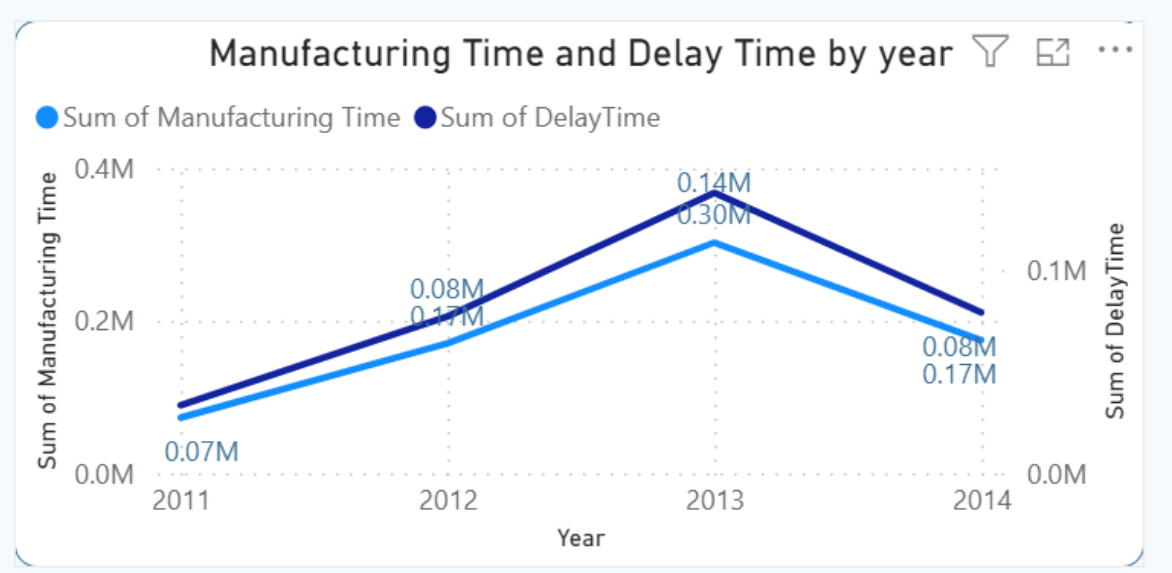


Figure 5-10: Manufacturing Time and Delay Time by year

The diagram in Figure 5-10 shows the total production time and the total delay time for 4 years. The diagram in Figure 5-10 shows the total production time and the total delay time for 4 years. We can see that production time and delay are proportional to each other year by year with the difference not too volatile. Production time gradually increased and decreased in 2014 due to statistical data until the second quarter of 2014, but according to current trends, it is likely that production time will continue to increase until the end of 2014.

The increase in production time can be a good sign because the size of the company's business and production is growing more and more.

In particular, compared to the beginning of 2011, delay time has decreased more than lead time. This shows that, at a time when the business is still young, there will be some difficulties in the production process leading to delays, but now, when the process is more and more stable and professional, we believe that the delay time will be reduced and businesses can optimize more on the production schedule.

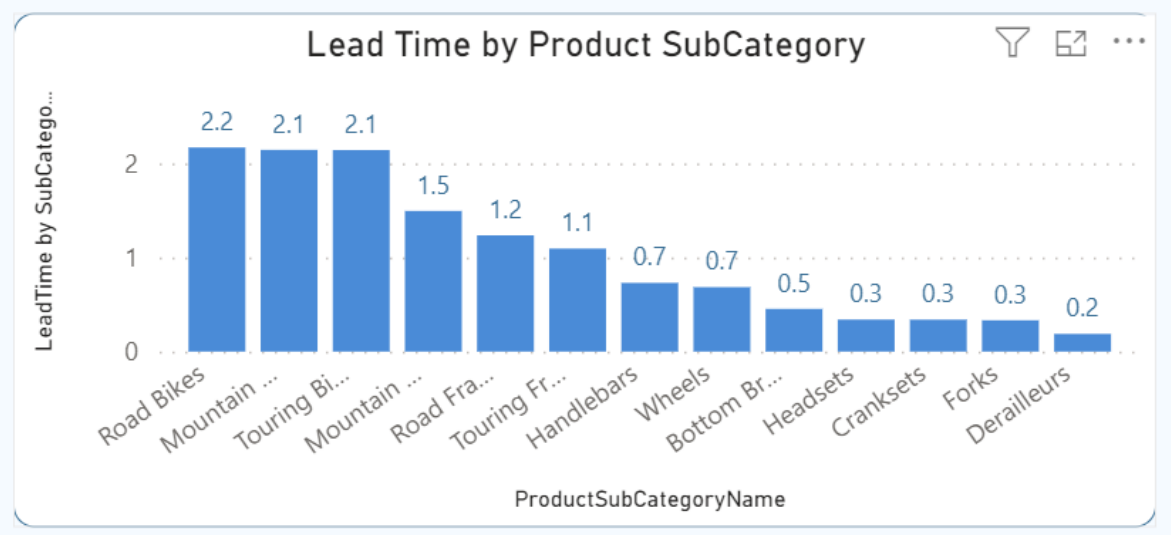


Figure 5-11: Lead Time by Product Subcategories

The column chart in Figure 5.11 shows the production time by SubCategory. This value shows the average total production time of each product type. This helps businesses control their production time more effectively by predicting the production time of products, thereby meeting customer needs, and taking measures to shorten production time.

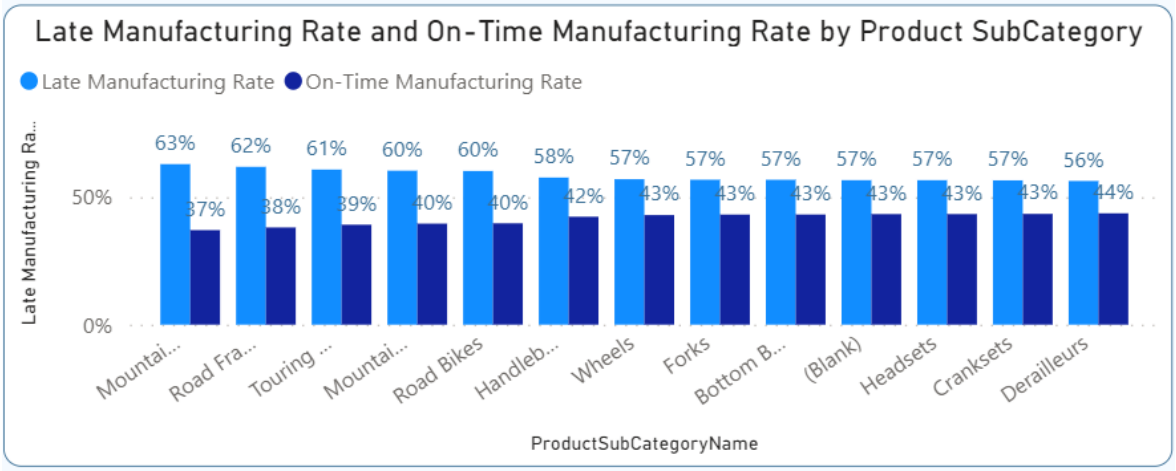


Figure 5-12: Late time and On-time Manufacturing by Product SubCategory

The chart in Figure 5-12 shows the comparison between the on-time production rate and the late production rate for each product. The graph shows that the difference is not too large between the two values. Although the on-time production rate is still not high, it is clear that all products have a higher on-time

delivery rate than the late rate. In general, the ratio between products is quite uniform. It can be concluded that the cause of production delay is not due to product characteristics.

ProductSubCategory	DelayTime	DelayDelivery	Rate
	51608	22394	13.61
Bottom Brackets	15236	12010	4.02
Cranksets	11363	8940	3.00
Derailleurs	18762	14808	4.95
Forks	28714	22626	7.57
Handlebars	75408	60183	19.88
Headsets	11378	8955	3.00
Mountain Bikes	22214	18212	5.86
Mountain Frames	41016	30929	10.81
Total	**379280**	**273563**	**100.00**

Figure 5-13: Table of Product Subcategory

The table in Figure 5.13 shows an overview of the production values that need to be considered, including: Lead time, late delivery time and lead time ratio. Managers can easily control production performance by product category to re-evaluate the expected lead time, take measures to reduce delay rates to improve production efficiency.

5.3.3 Dashboard about Inventory

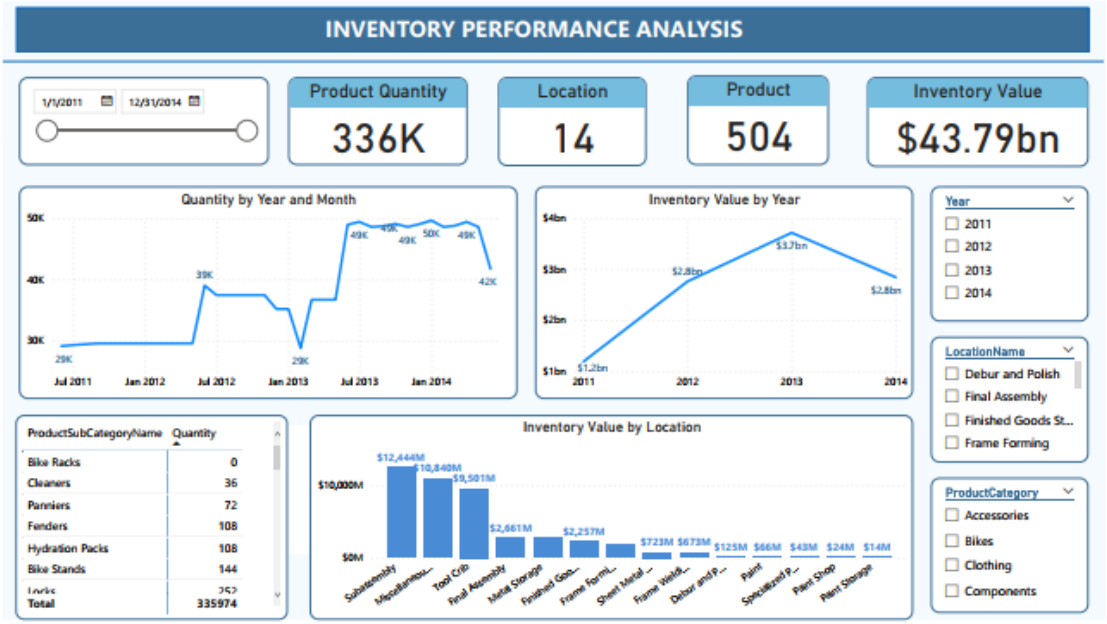


Figure 5-14: General Inventory Performance Dashboard over 2011 – 2014

From figure 5-14 above, between 2011 and 2014, there were 336K goods, 14 locations, and 504 inventory products totaled in the data inside the Inventory Performance Dashboard visualization team. An overview of AdventureWorks' production and retail operations from 2011 to 2014 is given in the figure below. This dashboard has undergone scrutiny. For thorough information on inventory kinds and their values, use the year, location Name, and product category filters. The graphs include a column chart that may be plotted, a line chart that shows the value of the inventory year over year, and a chart that shows the total number of products in stock over the course of each month and year. A table displaying the stock of Product Subcategory Name and the product inventory value for each Location. These graphs and the visualization up top assist stakeholders or decision makers in identifying trends and choosing which products to promote.

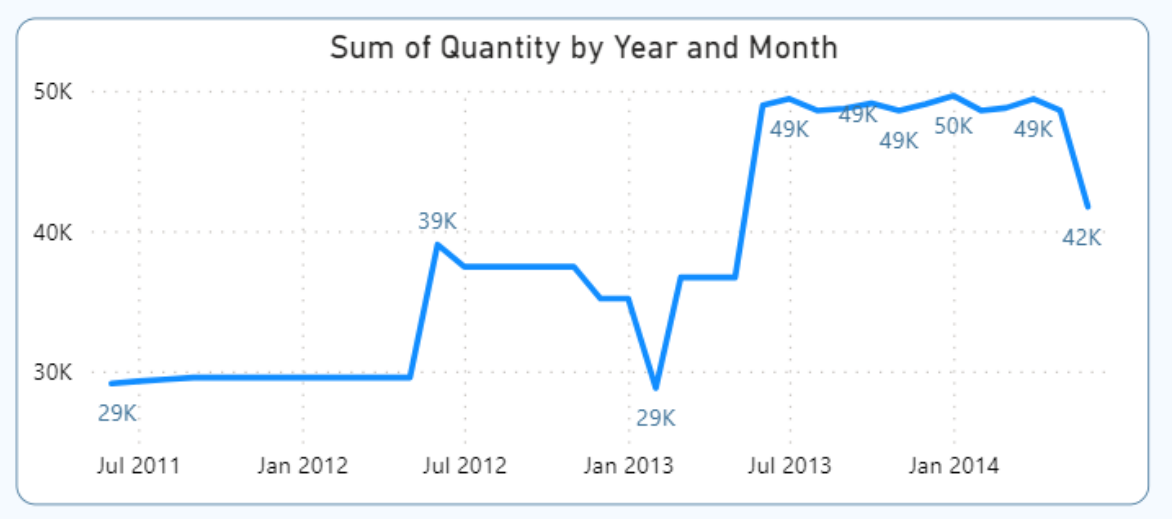


Figure 5-15: Sum of Quantity by Year and Month

From figure 5.15, this line graph shows the trend of changes in the total product inventory for the months of July 2011 and April, July, and February of 2012, 2013, and 2014. From July 2011 to February 2012, the quantity tended to stay at 29K. However, the number of inventories increased to 39K by July 2012. This figure typically declines and returns to 29K by February 2013. But by July 2013, there were 49K goods in stock, a substantial rise from the previous month. And from here until January 2014, there is a weak upward and downward trend, with the number of products in stock remaining at around the same level of 49K and 50K but down to 42K by February 2014.

These data allow us to conclude that inventories are growing annually. We can determine a possible remedy, such as. The first step is to do a market demand analysis to determine which products to make, sell, and develop marketing campaigns for. Second, to increase the amount of inventory sold to the market, this promotion could be applied to other products or the price of this product may be lowered.

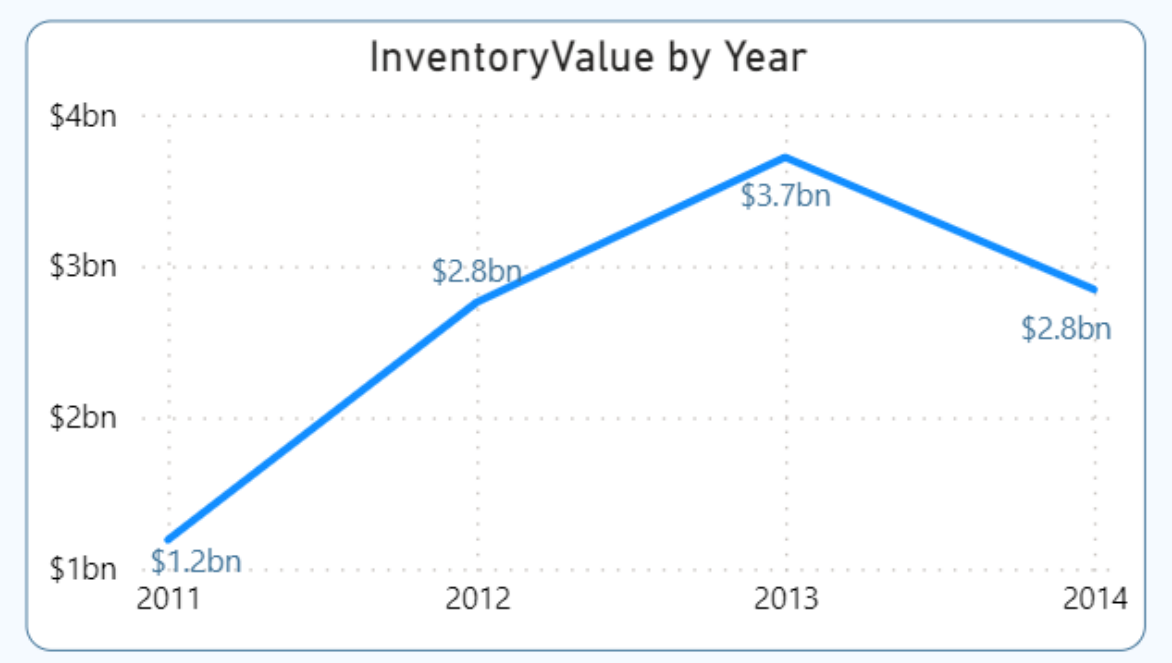


Figure 5-16: Inventory Value by Year over 2011 - 2014

From figure, this line chart shows that the value of inventory increases gradually from 2011 to 2012 and 2013 and has a decreasing trend from 2013 to 2014. In which, the inventory value of 2013 is considered to have the highest increase with a value of $3.7bn. The inventory value for 2012 and 2014 is equal at $2.8 billion and 2011 has a minimum value of $1.2 billion. Along with the increasing proportion of the number of products in stock, the value of inventory products also increases. This can cause the business value of the business to be lost or lost. In order to reduce the value of inventory, it is necessary to consider the costs as well as the transportation of the production so that it is less expensive and saves time in the most optimal way. This helps businesses produce faster but also save their own value.

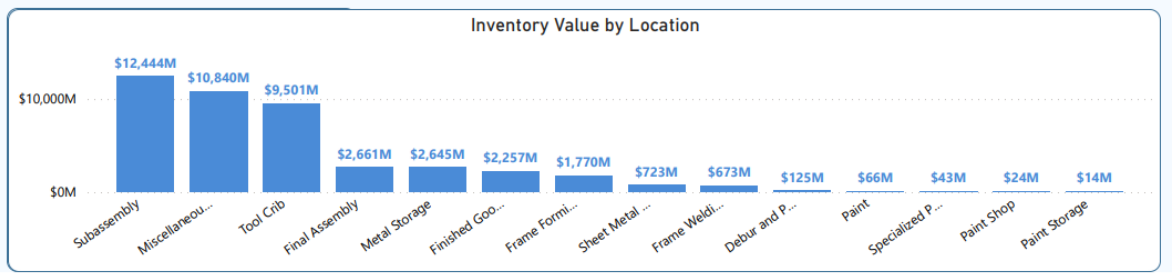


Figure 5-17: Inventory Value by Location

The column chart shows the inventory value by location name, with Subassembly having the highest value of $12.444M, followed by Miscellaneous storage at $10.840M and Tool Crib at $9.501M. Final Assembly has a value of $2.661M, while Metal storage and Finished goods storage have similar values at $2.645M and $2.257M respectively. Frame Forming has a value of $1.77M, Sheet metal racks at $723M, Frame welding at $673M, Debur and polish at $125M, and paint at $66M. The lowest values are for specialized paint at $43M, paint shop at $24M, and paint storage at $14M.

From the chart, it can be observed that Subassembly, Miscellaneous storage, and Tool Crib are the most valuable locations, with a significant difference in value compared to other locations. Therefore, it is important to pay attention to these locations and ensure that proper inventory management practices are in place to prevent loss or damage to inventory. On the other hand, the lower value locations such as paint, paint shop, and paint storage, may not require as much attention in terms of inventory management. However, it is still important to monitor inventory levels in these locations to ensure that there is no excess or shortage of inventory that could impact production and customer satisfaction.

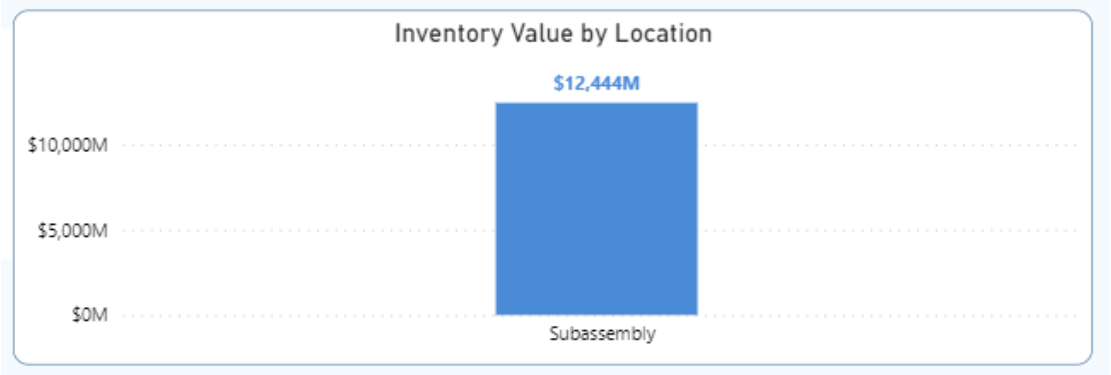


Figure 5-18: Inventory Value by Location Subassembly

In which, Location about Subassembly has the highest value of $12,444,080.008,2225 billion.

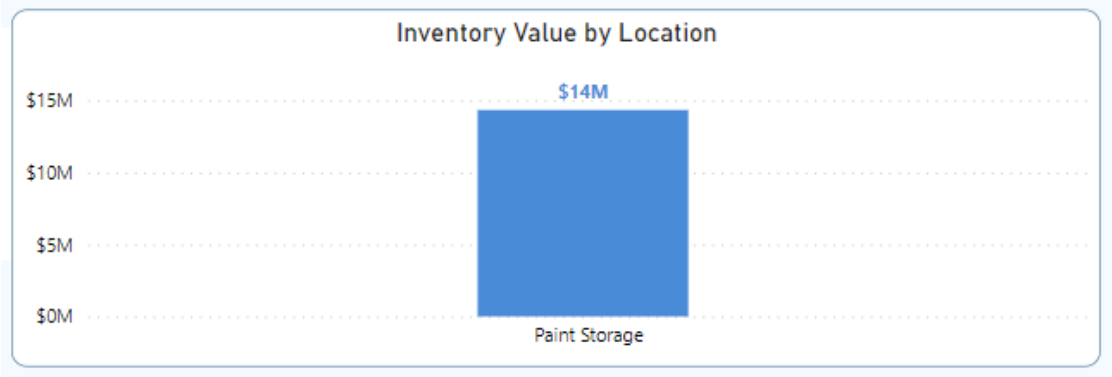
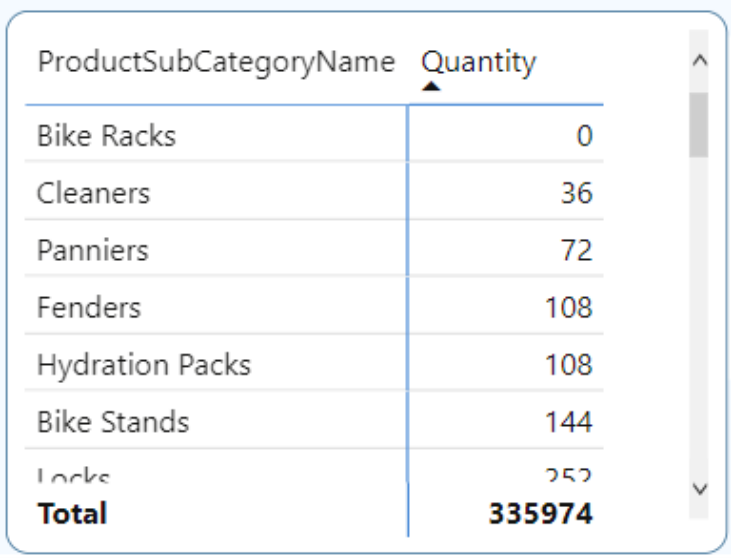


Figure 5-19: Inventory Value by Location Paint Storage

Location of Paint Storage has the lowest value of $14,336,948,175 billion.



ProductSubCategoryName	Quantity
Bike Racks	0
Cleaners	36
Panniers	72
Fenders	108
Hydration Packs	108
Bike Stands	144
Locks	252
Total	**335974**

Figure 5-20: List of Product Subcategory Name's Inventory Quantity

From figure, this table provides a full inventory of Product Subcategory Name based on the figure... The entire inventory then equals 335974. There is no inventory where Product Category Name is Bike Racks. Users can compare the production and distribution process to the market using this table to manage and see which Product Subcategory Names are in stock.

5.4 Evaluation and discussion

With a dashboard that includes many charts, it helps managers easily grasp the business situation of each department, thereby making predictions about the future human resources situation of each department. Moreover, with the combination of many types as an overall picture, it gives an overview for managers to make reasonable decisions for the company.

There are several solutions that can be employed to reduce the number of defective products during the manufacturing process. One approach is to strengthen quality control measures. This can be achieved by implementing stricter inspection protocols and using more advanced testing equipment. Another solution is to provide additional employee training that focuses on identifying defects and improving production processes. By enhancing the skills and knowledge of the workers, they

will be better equipped to identify and rectify issues before they result in defective products.

In addition to these solutions, exploring technical or equipment improvements can help to further reduce the number of defective products. For example, upgrading manufacturing equipment to more advanced models with enhanced precision and accuracy can lead to a significant reduction in defects. Another option is to consider modifications to the production line that can improve the flow of materials and minimize the potential for errors. To identify areas for improvement in the production process, it is important to analyze trends in waste cost and actual cost over the years. This can help to pinpoint potential issues and inefficiencies that may be contributing to defects. By addressing these areas, it is possible to achieve cost savings and increased efficiency, as well as a reduction in the number of defective products.

Performance management is a crucial aspect of any business. One way to achieve this is by understanding the production time of each product, which allows for the evaluation of expected production time. From there, an effective production schedule can be developed, and orders can be accepted. Negotiating flexible completion times also gives the business some leeway while ensuring that customers are satisfied with the quality of production. It is essential to implement measures that can shorten production time, such as hiring additional labor, expanding the factory, or repairing equipment. These measures can improve efficiency and ensure that production is completed on time. Another critical aspect of performance management is controlling delayed production time. It is essential to identify the causes of delays in progress, such as malfunctioning machines, inaccurate expected completion time, or unforeseen incidents. Once these causes are identified, improvements can be made to address the issue. Setting targets for the next quarter or year can help reduce the delay rate, ensuring that production runs smoothly. A target of 20% reduction in delay rate can be set, and measures put in place to achieve this target. Controlling production costs is another essential aspect of performance management. The business needs to control costs to reduce unexpected expenses as well as actual

production costs. Based on this evaluation, it is possible to determine the ability of the business to control costs and implement measures to improve this ability.

To effectively address the challenge of excess inventory, there are a range of marketing strategies that can be implemented. One such approach is to utilize advertising campaigns that highlight the unique features and benefits of the product with trendy content that appeals to the target audience. By showcasing the value of the excess inventory products and positioning them in a highly desirable light, these campaigns can help to generate greater consumer interest and increase sales.

Another effective strategy is to implement promotional programs that incentivize customers to purchase excess inventory products. For example, offering free excess inventory products when purchasing other items can be an effective way to drive sales and clear out the inventory. Similarly, providing vouchers with a certain percentage discount when buying excess inventory products can be a powerful motivator for customers to make a purchase.

Overall, applying a range of solutions to these challenges can help businesses effectively manage inventory, reduce defects, and boost sales while providing greater value to customers.

CHAPTER 6: CONCLUSION AND FUTURE WORKS

This study offers valuable insights on the topic, but further research is needed to understand it more thoroughly. Future research could focus on the impact of external factors, such as cultural background or socioeconomic status, on the results. Increasing the sample size would also help generalize the findings. Additionally, the study's findings could be applied in real-world settings to improve outcomes in related fields. This study is an important contribution to the literature and provides a foundation for future research.

6.1 Results

This project is a comprehensive solution that provides businesses and managers with the necessary insights to help them make informed decisions that are aligned with their goals. With our top-notch data warehouse and production information analytics, we can help companies gain a deeper understanding of their operations and identify areas for improvement. In addition to the data warehouse, we have also developed some model KPIs that can be used as benchmarks to measure performance. These KPIs cover a wide range of areas, from production efficiency to customer satisfaction, and provide a holistic view of the business's overall performance. To make it easier for managers to access and analyze these insights, we have also designed intuitive and customizable dashboards that display the data in a visually appealing way. With these dashboards, managers can quickly identify trends, spot anomalies, and drill down into specific areas for more detailed information.

- To reduce the number of defective products during manufacturing, there are several solutions. One solution is to improve quality control by using stricter inspection protocols and advanced testing equipment. Another solution is to provide employee training to identify defects and improve production processes. Upgrading manufacturing equipment and improving material flow can also help reduce defects. Analyzing waste cost and actual cost trends can pinpoint potential issues and inefficiencies to address.

- Performance management is crucial. Understanding production time allows for evaluating expected production time, developing an effective production schedule, and accepting orders with flexible completion times. Hiring additional labor, expanding the factory, or repairing equipment can shorten production time. Identifying causes of delays, such as malfunctioning equipment, inaccurate expected completion time, or unforeseen incidents, can help improve delayed production time. Setting targets for a 20% reduction in delay rate is an effective way to ensure production runs smoothly. Controlling production costs is also necessary to reduce unexpected expenses.
- To address excess inventory, businesses can use marketing strategies. Advertising campaigns that highlight the unique features of products with trendy content, promotional programs that incentivize customers to purchase excess inventory products, and sales on excess inventory products can all be effective. By applying a range of solutions, businesses can manage inventory, boost sales, and provide greater value to their customers.

6.2 Limitations

Besides the results achieved through the implementation of the project, our team still has the following limitations:

- Due to limited time, most of the knowledge and contents learned have not been studied deeply
- The team spent a lot of time in shaping, searching and starting to work on the topic as well as understanding the Adventurework dataset
- Difficulty in getting errors and fixing errors during ETL and SSAS process implementation
- The construction of the data model is not thorough, leading to a lack of measurement to perform visualization
- The project is at the level of learning and research, so the analysis is still not deep and meaningful, mainly data representation on dashboards

- In the process of making the project, there are still certain errors that are difficult to control

6.3 Future works

In the future, the whole team is expected to apply integrated analysis of many analytical application tools such as Tableau, QlikView, Python to apply to the research and development process of this subject along with open Scale up analysis of departments like Sales and Human Resources to improve product manufacturing processes, predict demand trends on which products can be produced to meet demand or can be allocated human resources how to finish the product production process quickly and save the most time. This combination will help to make standard decisions more clearly and help the company to grow and gain a foothold in the domestic and international markets.

REFERENCES

[1] C. Stedman, "What is Business Intelligence (BI)? | Definition from TechTarget," 2 2023. [Online]. Available: https://www.techtarget.com/searchbusinessanalytics/definition/business-intelligence-BI.

[2] M. K. Pratt, "business intelligence architecture," 7 2020. [Online]. Available: https://www.techtarget.com/searchbusinessanalytics/definition/business-intelligence-architecture.

[3] raman_257, "ETL Process in Data Warehouse," 2 2 2023. [Online]. Available: https://www.geeksforgeeks.org/etl-process-in-data-warehouse/.

[4] J. Parkinson, "Choosing the ETL tool for you: A crash course from Peak Ace," 6 6 2022. [Online]. Available: https://www.dall-int.com/choosing-the-etl-tool-for-you-a-crash-course-from-peak-aces-martech-team/.

[5] T. D. Thanh, "Data Warehouse là gì? Kiến trúc và lợi ích của kho dữ liệu," 16 9 2022. [Online]. Available: https://bkhost.vn/blog/data-warehouse-kho-du-lieu/.

[6] "What is a KPI?," [Online]. Available: https://www.qlik.com/us/kpi.

[7] S. Dharak, "Advantages and Disadvantages of Key Performance Indicators," 26 6 2022. [Online]. Available: https://www.solobis.net/limitations-of-key-performance-indicators/.

[8] S. Kinsey, "21 Production KPIs you should be tracking on dashboards and reports," 19 11 2021. [Online]. Available: https://www.simplekpi.com/Blog/21-Production-KPIs.

[9] "Production - Adventurework," 2023. [Online]. Available: https://dataedo.com/samples/html/AdventureWorks/doc/AdventureWorks_2/modules/Products_10/module.html.

[10] "Kho dữ liệu, hồ dữ liệu và tập hợp dữ liệu khác nhau như thế nào?," 2023. [Online]. Available: https://aws.amazon.com/vi/compare/the-difference-between-a-data-warehouse-data-lake-and-data-mart/.

[11] "So sánh công cụ Power BI - Tableau - Qlik," 2 2021. [Online]. Available: https://kpim.vn/so-sanh-power-bi-vs-tableau-vs-qlik/.

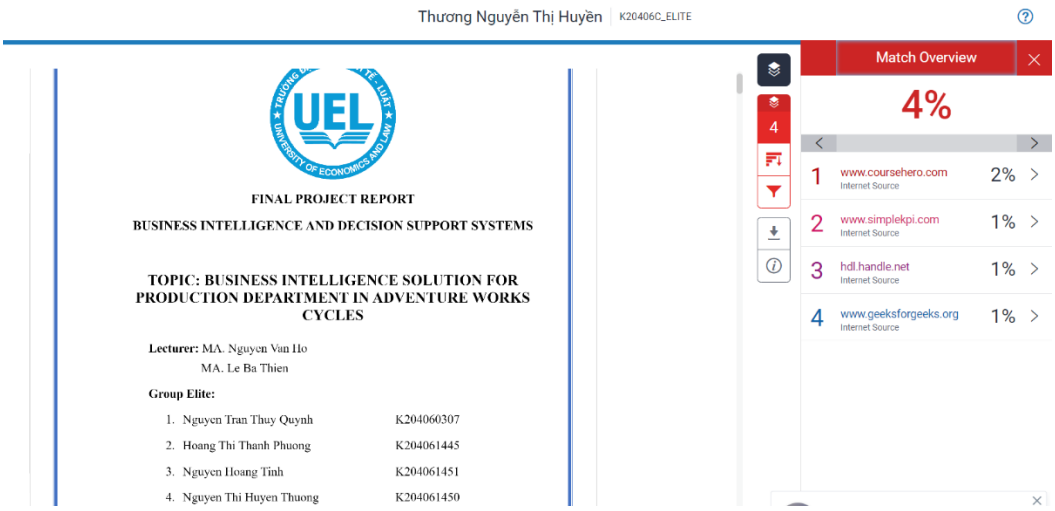
[12] D. Taylor, "Star Schema vs Snowflake Schema – Difference Between Them," 8 4 2023. [Online]. Available: https://www.guru99.com/star-snowflake-data-warehousing.html.

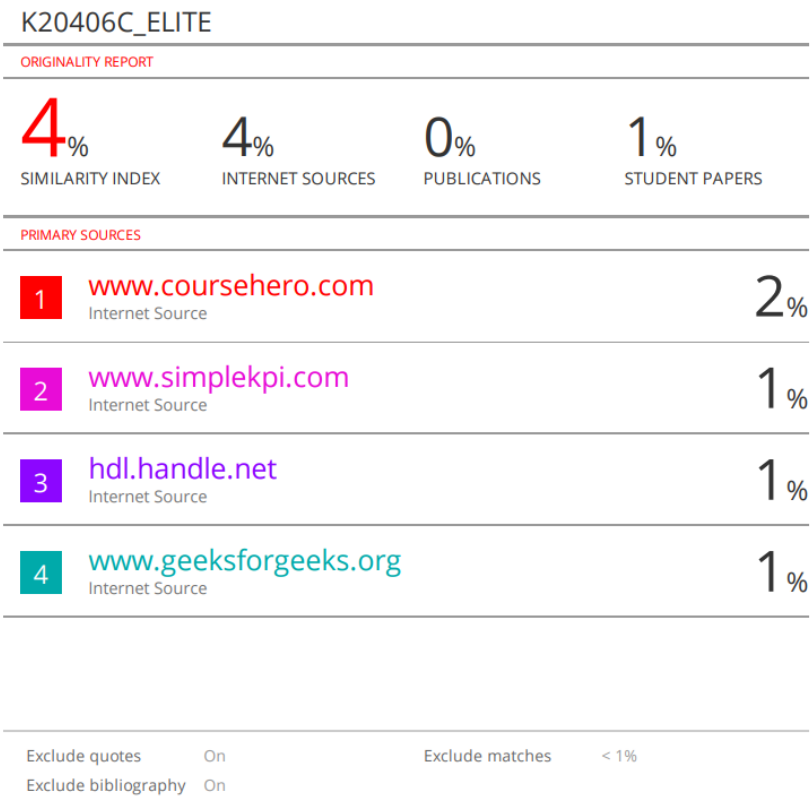
PROJECT EVALUATION

No	Full Name and Student Code	Levels of completions
1	Nguyễn Trần Thúy Quỳnh – K204060307 - Define the Business case, scope and structure of the project - Brainstorm ideas for business requirement - Build Cube, analyze data with SSAS - Analyze base on dashboard - Write report: Introduction, Theoretical basis, Data sources, Data analytics with SSAS, Data analytics with Power BI, Evaluation and Discussion, Result	100%
2	Hoàng Thị Thanh Phương – K204061445 - Define the Business case, scope and structure of the project - Brainstorm ideas for business requirement - Build Bus matrix - Visualize with Power BI - Analyze base on dashboard - Write report: Theoretical basis, Business Requirement analysis, Data analytics with Power BI	100%
3	Nguyễn Hoàng Tính – K204061451 - Define the Business case, scope and structure of the project - Brainstorm ideas for business requirement - Build Dim, Fact Table - Analyze base on dashboard - Design slide for Presentation - Write report: Theoretical basis, Business Requirement analysis, Data analytics with Power BI, Future works	100%
4	Nguyễn Thị Huyền Thương – K204061450 - Define the Business case, scope and structure of the project - Brainstorm ideas for business requirement - Build Dim, Fact Table	100%

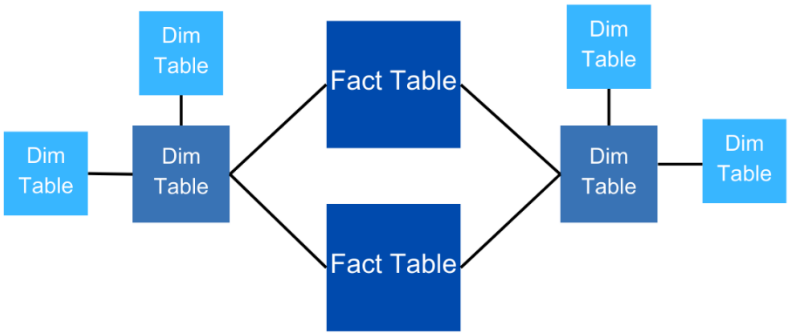
	- Build Datawarehouse, perform ETL process for Dim, Fact Table - Visualize with Power BI - Write report: Theoretical basis, Production Business Process, ETL processes, Dashboard structure, Limitations - Summary of content includes in the report	

SIMILARITY RATE CHECKING





EDITTING APPENDIX

Section	Amendment/Add
2.3.4	The following iteration of the data warehouse schema is the Galaxy Data Warehouse Schema, sometimes referred to as a Fact Constellation Schema. The Galaxy Schema makes use of numerous fact tables joined by shared normalized dimension tables, unlike the Star and Snowflake Schemas. Galaxy Schema may be compared to a collection of connected, fully normalized star schema that eliminates any redundant or inconsistent data.  Characteristics of the Galaxy Schema: For large database systems, Galaxy Schema's multidimensional nature serves as an important design factor. - Because to normalization, Galaxy Schema decreases duplication to almost nil. - Known for its high data quality and accuracy, Galaxy Schema facilitates efficient reporting and analytics.

3.5.1	Power BI, Tableau, and QlikView are all data analysis tools widely used in businesses and organizations. Each tool has its own advantages and applications, so the right choice will depend on the needs and requirements of each organization. For our project, we decided to use Power BI to analyze and visualize the data. Power BI is not the most powerful tool, but it is enough for us in terms of ease of use, aesthetics, and intuitiveness. For our less complicated requirements, Power BI has enough capacity to meet the needs of analyzing. In particular, Power BI is also the tool with the lowest cost of the above tools, so for a subject project, we were convinced by this condition. One key point that makes us use Power BI is its integration with other Microsoft products. During our entire process of working on this project, we used most of the Microsoft products. It was these persuasive factors that led our team to use Power BI to complete our report.