UNIVERSITY OF ECONOMICS AND LAW

FACULTY OF INFORMATION SYSTEMS

FINAL PROJECT REPORT

BUSINESS INTELLIENCE & DECISION SUPPORT SYSTEMS

TOPIC: PROPOSED BUSINESS INTELLIGENCE SOLUTION FOR PROCUREMENT EXCELLENCE AT ADVENTURE WORKS

Group: 3

Instructors:

1. **Assoc. Prof. Ho Trung Thanh, Ph.D.**
2. **Nguyen Van Ho, M.A.**
3. **Le Ba Thien, M.Sc.**

Ho Chi Minh City, April, 2024

Members of Group 3

No.	Full name	Student ID	Point / 10 (Individual Contribution	Signature
1	Trần Sĩ Đan	K214061258	10	
2	Nguyễn Thiên Huy	K214060396	10	Huy
3	Giả Hoàng Nam Phương	K214061744	10	
4	Trịnh Quốc Thịnh	K214060412	10	Thịnh
5	Dương Văn Nhựt Duy	K214060391	10	

Acknowledgements

First of all, we would like to thank the University of Economics and Law - Vietnam National University, Ho Chi Minh City, for including Business Intelligence and Decision Support Systems in the curriculum.

In particular, to complete this report, we would like to express our deep gratitude to Assoc. Prof. Ho Trung Thanh, Ph.D., Nguyen Van Ho, M.A. and Le Ba Thien, M.Sc. who is currently teaching Business Intelligence and Decision Support Systems for our class. The lectures that provide useful and interesting knowledge, along with the enthusiasm of the teachers and sisters in the teaching process, are a great inspiration to help us complete this report.

Knowledge is limitless, but each person's own knowledge always has certain limitations. Therefore, in the process of making the report, it will be difficult for us to avoid errors. Our team is looking forward to receiving suggestions from teachers so that the group can improve further. This means a lot to our group.

Wishing you good health and more success in your teaching career. We sincerely thank you!

Group 3

Commitment

We hereby commit that the following results are purely the application of our knowledge based on the knowledge taught from the Business Intelligence and Decision Support Systems subjects of Assoc. Prof. Ho Trung Thanh, Ph.D., Nguyen Van Ho, M.A. and Le Ba Thien, M.Sc., combined with references from books, newspapers, and other media.

Ho Chi Minh city, April 2024

Group 3

Table of Content

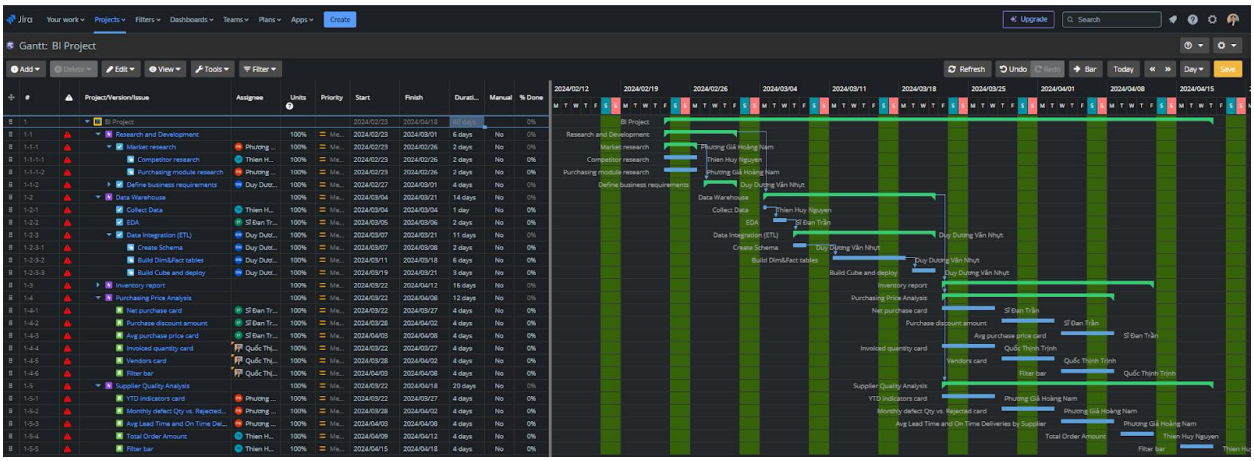
List of Figures

List of Tables

List of Abbreviations

KPI	Key Performance Indicator
ERP	Enterprise Resource Planning
ETL	Extract, Transform, and Load
CAGR	Compound Annual Growth Rate

Gantt Chart


Jira
Gantt: BI Project
Project/Version/Issue
Assignee
Units
Priority
Start
Finish
Manual
% Done
BI Project
Research and Development
Market research
Competitor research
Purchasing module research
Define business requirements
Data Warehouse
Collect Data
EDA
Data Integration (ETL)
Create Schema
Build Dim&Fact tables
Build Cube and deploy
Inventory report
Purchasing Price Analysis
Net purchase card
Purchase discount amount
Avg purchase price card
Invoiced quantity card
Vendors card
Filter bar
Supplier Quality Analysis
YTD Indicators card
Monthly defect Qty vs. Rejected...
Avg Lead Time and On Time Del...
Total Order Amount
Filter bar

Chapter 1: Project Overview

The introductory chapter is intended to introduce and provide an overview of the topic, as well as to establish foundational content such as setting goals, defining the object and scope, outlining research methods, detailing the implementation process, and describing the overall structure of the study project.

1.1. Business Case for the Project

1.1.1. Business case overview

1.1.1.1. Market research

a. Global market

The bicycle industry on a global scale has experienced significant growth, with projections indicating a robust trajectory. According to Statista (2024), in 2024, the projected revenue in the global Bicycle market is estimated to reach an impressive $62.80 billion. Looking ahead, the market is expected to experience a steady annual growth rate (CAGR 2024-2029) of 2.19%. This growth will result in a projected market volume of approximately $70.00 billion by 2029. Additionally, unit sales of the Bicycle market are expected to reach 138.30 million bicycles in 2029. When examining the pricing aspect of the Bicycle market, the volume weighted average price in 2024 is anticipated to be $0.45 thousand USD. From an international perspective, it is evident that China will lead in generating the most revenue in the Bicycle market, with an estimated $12.37 billion in 2024. The global bicycle market is experiencing a surge in revenue due to the increasing popularity of electric bicycles and cycling as a recreational activity and environmentally friendly mode of transportation.

One notable trend contributing to market expansion is the rising popularity of electric bicycles, which seamlessly integrate traditional biking with electric motor assistance, is reshaping consumer preferences. These electric bikes are particularly favored for their eco-friendliness and ability to alleviate physical exertion, making them ideal for longer commutes and challenging terrains. Concurrently, there’s a notable demand surge for specialized bicycles tailored to specific needs, such as mountain or road biking, leading

to the emergence of niche manufacturers. Local circumstances, such as urban congestion and strong cycling cultures, further contribute to market dynamics. Moreover, macroeconomic factors like government incentives for sustainable transportation and a growing focus on health and fitness are propelling bicycle adoption globally.

b. Vietnam market

The Vietnamese bicycle market reflects a promising growth trajectory, albeit on a smaller scale compared to the global landscape. According to Statista (2024), a steady increase in revenue, with expectations of reaching approximately $88.85 million in 2024 and $99.94 million by 2028, representing a compound annual growth rate (CAGR) of 2.98% during the period 2024-2028. Despite its relatively small size, the domestic market holds potential for expansion, driven by factors such as increasing urbanization, economic development, and a growing focus on personal mobility solutions. In terms of unit sales, the Vietnamese market is anticipated to achieve 0.54 million units by 2028, with an average weighted price of $159.30 per unit. However, it is worth noting that bicycle consumption in Vietnam remains modest compared to leading nations like China, primarily influenced by factors such as prevailing living standards and awareness of alternative transportation options.

To capitalize on the growth potential, efforts to enhance infrastructure, promote cycling culture, and raise awareness of the benefits of bicycles as a sustainable mode of transportation are crucial. By addressing these factors, the Vietnamese bicycle market can further solidify its position and contribute to the overall development of the country's transportation ecosystem.

1.1.1.2. Company Overview

According to Microsoft Learn, AdventureWorks is a fictional company, Adventure Works Cycles, known as a large-scale multinational bicycle manufacturing enterprise. With its headquarters in Bothell, Washington, the company takes pride in its dedicated and professional team of nearly 300 employees.

With the aim of expanding market share and enhancing business operations, Adventure Works Cycles has continuously invested in growth and development. In 2000,

they acquired the Importadores Neptuno manufacturing plant in Mexico, marking a significant step in expanding production and diversifying product lines. This facility not only plays a crucial role in producing key components for bicycles but also serves as the sole location for manufacturing the touring bicycle product line.

Particularly, to strengthen market presence and optimize business operations, Adventure Works Cycles is focusing on developing its website system to provide comprehensive and convenient product information to customers. At the same time, they are striving to reduce production costs to maximize profits and deliver the best value to customers.

1.1.1.3. Purchasing relative department

According Indeed Editorial Team, a purchasing department, also known as a procurement or buying department in some companies, is a crucial division responsible for acquiring the necessary goods and services essential for the business's seamless operation. These departments play a vital role in assisting companies in meeting their day-to-day requirements as well as accomplishing their long-term strategic objectives. The extent of responsibilities held by a purchasing department may vary based on the company's size, but typically, they play a key role in monitoring the supply chain and effectively managing vendor contracts to ensure the efficiency of the company's operations.

The Purpose of Purchasing Department

The purchasing department is instrumental in an organization, serving as a key player in the effective acquisition of goods and services vital to operational efficiency. Its core mission involves identifying reliable suppliers, securing cost-effective deals through adept negotiation, and aligning with budgetary constraints. The department manages supplier relationships, negotiates contracts for favorable pricing and terms, and ensures timely delivery of goods to specified standards.

Additionally, the purchasing department minimizes inventory costs while maintaining sufficient stock levels, requiring careful management. It plays a pivotal role in risk management, identifying and mitigating disruptions, price fluctuations, and quality

issues within the supply chain for seamless operations. Strict compliance with laws, regulations, and organizational policies upholds ethical and legal standards.

Strategically, the department aligns purchasing strategies with long-term goals, staying attuned to market trends, and contributing insights to the strategic planning process. By achieving these objectives, the purchasing department optimizes resource allocation, minimizes risks, and enhances the organization's overall efficiency and competitiveness in the dynamic marketplace.

Purchasing Process

The purchasing process is detailed as in the BPMN diagram below.

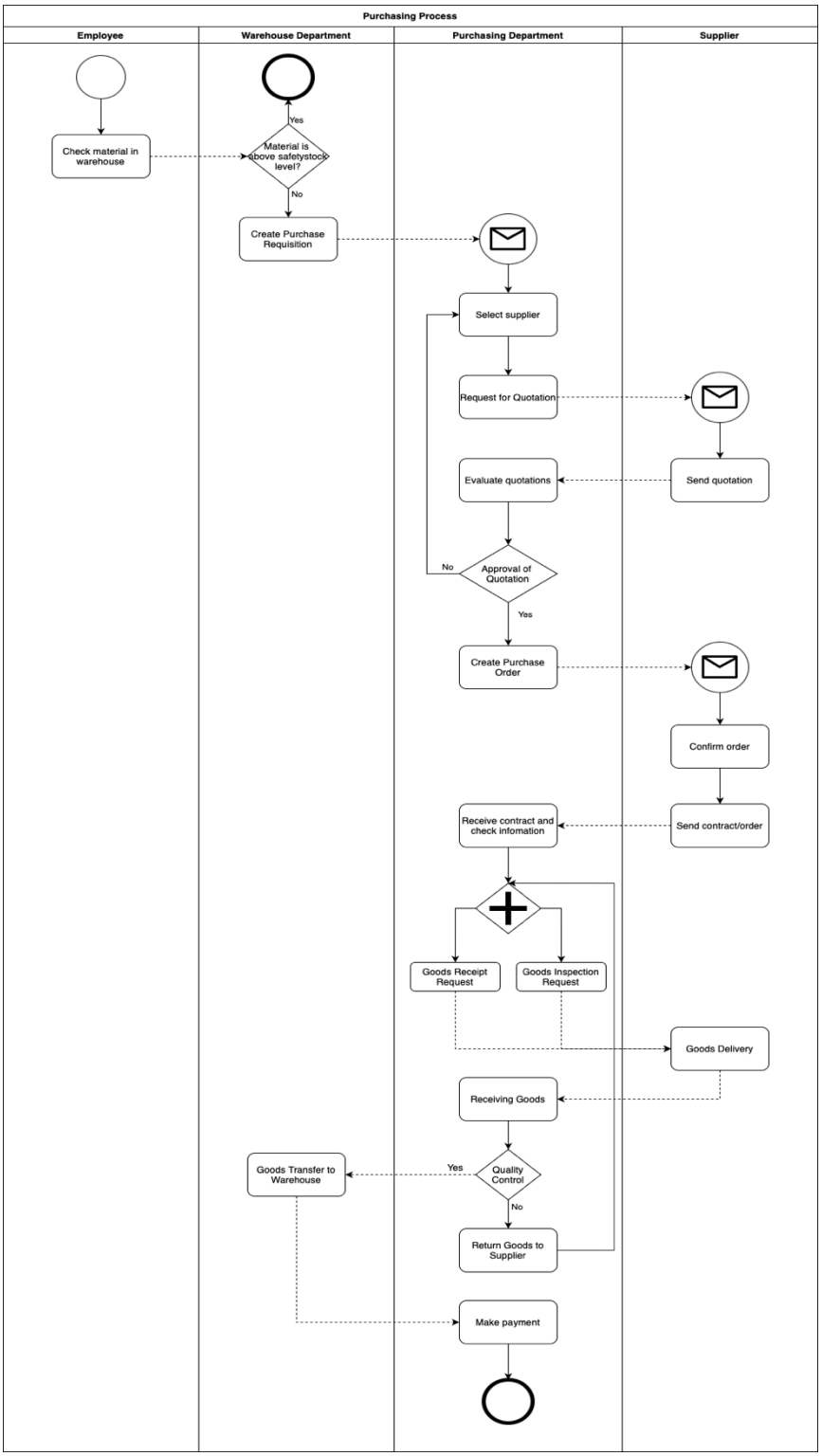


Figure 1.1. Purchasing Process (Source: Authors propose)

1.1.2. Business demands/problems

In order to define the main business problems of Adventure Works Cycles, being well understanding of the context and how the enterprise operates is essential. To get a more comprehensive perspective on the company, the project team drew up a Business Model Canvas and SWOT analysis as below.

1.1.2.1. Business Model Canvas

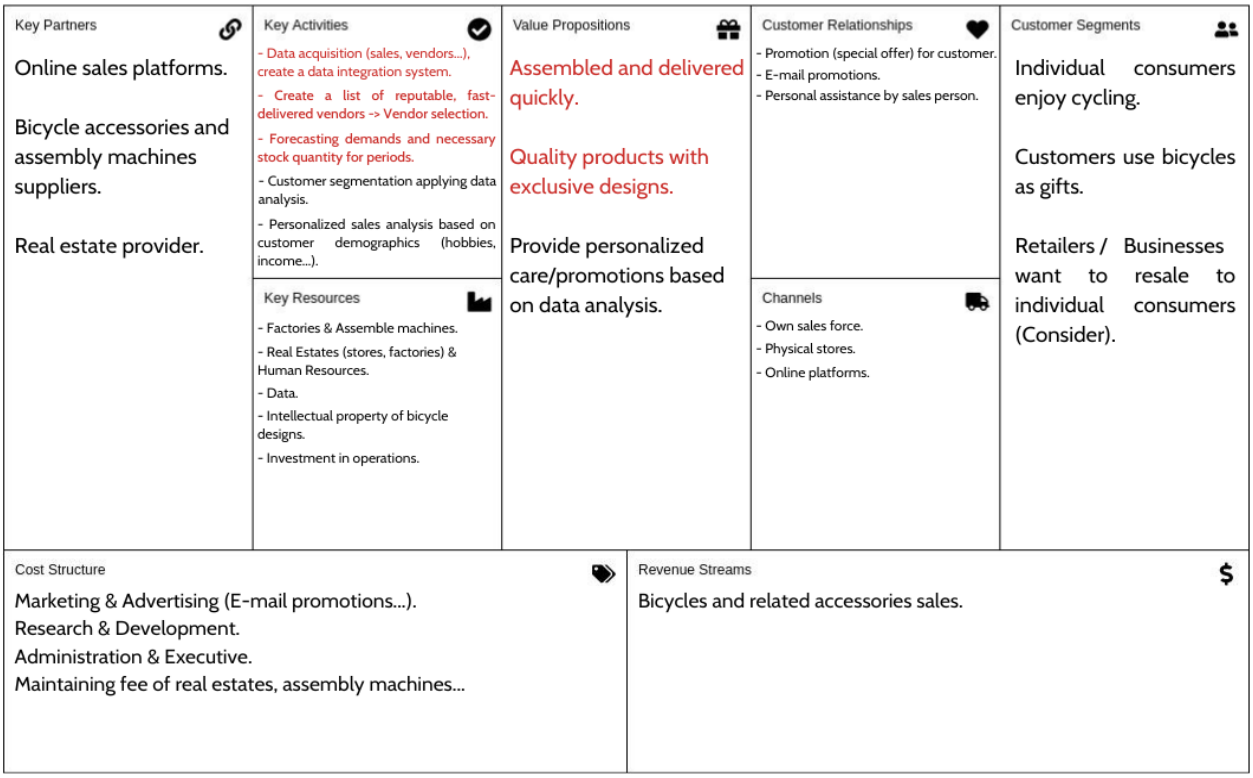


Figure 1.2. Business Model Canvas of Adventure Works Cycles (Source: Authors propose)

Business Model Canvas is a strategic planning support tool developed by Alexander Osterwalder (Founder and CEO of Strategyzer). The template only has one page which describes nine key elements that strongly affect the company.

- Customer Segments: Adventure Works Cycles (AWC) serve individuals who are cycling enthusiasts (B2C). Additionally, the project team assumes that it also works with other businesses (B2B) because there are many SalesOrderDetail records showing a great amount of order quantity.

- Value Proposition: the company’s values focus on the quality of each product; in the Qualitative aspect, AWC wants to serve their customers in a personalized way to bring the most suitable promotions or campaigns that meet customers’ needs - and this requirement can be supported by data analysis.

- Channels: based on the context and description of the AdventureWorks2019 dataset, physical stores and online platforms are two channels that AWC uses to distribute its products.

- Customer Relationships: the table Sales.SpecialOffer and the attribute UnitPriceDiscount in table Sales.SalesOrderDetail demonstrates that AWC has promotions and special offers for customers to stimulate purchases. In table Person.Person, there is also an EmailPromotion attribute that speaks to AWC's use of customer segmentation strategies for email marketing. At the stores, there are sales persons responsible for communicating with and supporting customers.

- Revenue Streams: AWC’s revenue mainly comes from bicycles and related accessories sales.

- Key Resources: as a manufacturer company, in order to operate effectively, AWC cannot lack quality high-tech factories and assembly equipment. Furthermore, the company has to invest in some real estate to expand its physical store chain, thereby increasing its human resources. Another important point is intellectual property rights, AWC needs to have copyright for its product designs to avoid having ideas stolen from competitors.

- Key Activities: The main and future activities of AWC need to be proposed based on the core values that the company wants to bring to its customers.

+ To be able to provide personalized services and marketing programs to customers, AWC needs to build an effective data integration system, thereby analyzing customer subgroups to make decisions for each segmentation.

+ To meet the demand for good quality products, low prices and fast shipping, choosing the right supplier, controlling inventory and improving production processes is extremely important. Additionally, forecasting

demand for each period of time helps AWC planning for stock quantity accurately, avoid excess or shortage of products.

- Key Partnerships: AWC works with the online sales platforms provider, real estate provider and vendors who provide bicycle accessories and assembly equipment.

1.1.2.2. SWOT analysis of the bicycle industry in the global market

Table 1-1. SWOT analysis of the bicycle industry in the global market (Source: Authors propose)

Strengths	**Weaknesses**
- Sustainable Transportation: Bicycles are seen as environmentally friendly and promote sustainable transportation, which is increasingly valued in today's environmentally conscious world. - Health and Fitness Trends: Growing emphasis on health and fitness benefits of cycling is driving demand for bicycles and related products. - Urbanization: Increasing urbanization in many parts of the world promotes cycling as a convenient mode of transportation for short distances, reducing traffic congestion and pollution. - Innovation: Constant innovation in bicycle technology, materials, and design leads to the development of	- Seasonal Demand: The bicycle industry often experiences seasonal demand patterns, with sales declining during colder or wetter seasons in many regions. - Safety Concerns: Safety concerns related to traffic accidents and infrastructure limitations can deter some potential cyclists, particularly in urban areas with inadequate cycling infrastructure. - Competition from Alternative Transportation: Competition from alternative transportation modes such as electric scooters and ride-sharing services can affect bicycle sales, especially for short urban commutes. - High Initial Cost: High initial cost can be a barrier for some consumers,

new and improved products, attracting consumers seeking cutting-edge features. - Government Support: Many governments promote cycling through infrastructure investments, bike-sharing programs, and incentives for bicycle commuting, contributing to market growth. For example, Copenhagen, Denmark has built many bridges for bicycles, 4 of them have been finished.	particularly for high-end bicycles and specialized equipment. - Fragmented Market: The global bicycle market is highly fragmented with numerous small and medium-sized manufacturers, leading to intense competition and pricing pressures.
Opportunities	**Threats**
- Emerging Markets: Emerging markets in Asia, Latin America, and Africa offer untapped potential for bicycle sales due to increasing urbanization, rising disposable incomes, and infrastructure development. - Bike-sharing Programs: Expansion of bike-sharing programs in urban areas presents opportunities for manufacturers and rental service providers to cater to the growing demand for shared mobility solutions. - Sustainable Tourism: Growing interest in eco-friendly and	- Regulatory Changes: Changes in government regulations related to safety standards, tariffs, or environmental regulations can impact manufacturing costs and market dynamics. - Substitution: Substitution threats from alternative modes of transportation such as electric scooters, public transit, or ride-sharing services can affect bicycle sales. - Supply Chain Disruptions: Disruptions in the global supply chain, such as raw material shortages or logistics

sustainable tourism presents opportunities for bicycle rental and tour operators in popular tourist destinations. - Customization and Personalization: Increasing demand for customized and personalized bicycles presents opportunities for manufacturers to offer bespoke products tailored to individual preferences.	challenges, can impact production and distribution of bicycles. - Intense Competition: Intense competition within the industry, including both traditional manufacturers and new market entrants, can lead to price wars, margin pressures, and market saturation.

1.1.2.3. Business problems

Through analysis of the bicycle industry in the market and the Business Model Canvas, some current business problems for AWC can be mentioned as:

- Intense Competition and Brand Differentiation: The highly competitive nature of the bicycle industry may pose challenges for AWC in distinguishing its products and brand from competitors.

- Understanding Customer Preferences and Trends: Adapting to shifting customer preferences and lifestyle trends related to health, fitness, and urban living is crucial for Adventure Works Cycles to effectively target and retain its customer base. Therefore, customer segmentation and personalized analysis are two of the key activities in the Business Model Canvas.

- Global Expansion and Market Entry: Venturing into emerging markets in Asia, Latin America, and Africa offers growth opportunities for Adventure Works Cycles, but requires thorough market analysis, strategic planning, and adaptation to local regulations and consumer preferences.

- Seasonal Demand Fluctuations: Adventure Works Cycles may experience fluctuations in sales due to seasonal factors, such as reduced demand during colder or wetter seasons, necessitating effective inventory and production management. This

problem is also mentioned in the Business Model Canvas about the forecasting demand activity in order to avoid excess and shortage of products.

- Market Saturation and Pricing Pressure: The industry's competitiveness could lead to market saturation and pressure on pricing, potentially impacting Adventure Works Cycles' profitability and market share.

- Innovation and Technological Adaptation: Staying abreast of technological advancements and innovations in bicycle design and materials is essential for Adventure Works Cycles to meet evolving customer preferences and maintain competitiveness.

- Supply Chain Management: Challenges in the global supply chain, such as material shortages or logistical issues, could disrupt Adventure Works Cycles' production and distribution processes, impacting its ability to fulfill orders, meet customer demand and delivery time.

Deeper researching on the Purchasing domain in this field, our team found that there are 4 main issues that directly affect this module: Seasonal demand fluctuations, Market saturation and pricing pressure, Innovation and technology adaptation, Supply chain management. Referring to Purchasing's perspective, the research team divided the problems into two more specific groups: “How to reduce overstock or shortage of goods?” and “How to choose reputable vendors that provide high-quality products, with fast shipping time?”

1.2. Objectives of the Project

1.2.1. General Objective

The general objective of implementing Business Intelligence (BI) solutions at AdventureWorks companies is to facilitate data-driven decision-making and promote a data-centric organizational culture by providing an advanced strategic tool. BI is considered a strategic tool that offers a comprehensive view of both current and historical data, providing valuable insights into various aspects of the business. The main goal is to enhance overall operational efficiency and quickly respond to dynamic market conditions. By using BI and data visualization tools, such as dashboards and reports, the solution aims to optimize procurement processes, minimize costs and waste, as well as enhance overall

efficiency within the organization. Furthermore, integrating data across different departments fosters collaboration between functions, from purchasing to sales, warehouse management, and production, creating a more cohesive and collaborative working environment.

1.2.2. Specific Objectives

- Analyze Business Requirements: Understand the needs and challenges of various departments, identify KPIs and metrics related to inventory management, purchasing, and analyze current processes to uncover inefficiencies and areas for improvement.

- Survey Data Sources: Conduct a comprehensive assessment of data sources within AdventureWorks, evaluating the quality, completeness, and relevance of data from each source, identifying any gaps or discrepancies that may need to be addressed.

- Build Data Warehouse and Data Marts: Develop a robust data warehouse architecture and data marts to efficiently store and organize data for analysis and reporting purposes.

- Integrate Data into Data Model and Data Warehouse: Implement data integration processes to seamlessly combine data from multiple sources into the data model and warehouse infrastructure.

- Visualize Results: Utilize data visualization techniques, such as dashboards and reports, to visually represent key metrics and insights derived from BI analysis, facilitating understanding and decision-making.

- Propose BI Solutions: Recommend tailored BI solutions aligned with the specific needs and objectives of AdventureWorks, considering factors such as scalability, flexibility, and compatibility with existing systems.

1.3. Business Questions

In the business, there are too many questions that need to be answered to make the business work better and bring more and more profit for the company.

In inventory, there are some questions that are usually asked:

"What is the status of the specific product in terms of quality, quantity, and location?"

The first question is: "What is the status of the specific product in terms of quality, quantity, and location? Why does the company ask like that? It is because they want to know better about their products. They want to know more information about their products in terms of ID, name, its quality (good or bad), its quantity (is it enough for the next week, or is it under stock or overstock, its inventory cost,...). Having a knowledge of that information will help them to capture in detail about their product, make the right decision when purchasing a product and above all is their can minimize the inventory costs and expenses.

"Do we need to make a purchase for the specific product?"

The second question is a question that the companies are always asking. They want to know when we need to make the purchasing decision to reduce the costs and expenses of the inventory management as much as possible.

"What is the status of the product in terms of current quantity, safety stock?"

The third question is asked when the manager wants to know if their products are understock or overstock so they can fulfill the order of customers.

In vendor management, there are also some questions are usually asked:

"Who are our vendors?"

"Which are our best partners, which are our new vendors?"

"What products do they supply us?"

Those questions above are always asked when the manager wants to know which products come from the vendor. They also want to know who are the best vendors that they have had and use it to strengthen their business relationship.

"Can we compare the vendors based on the different scores (time, quality, return rate)?"

Sometimes, the manager wants to compare the suppliers together based on some indicators such as: price, delivery time, return rate,... to help them choose the best vendor for the order.

In terms of Pricing negotiation, there are several questions as: "What is the appropriate price for the product?", "What is the trend in price of the product?", "How can we choose the right vendor in terms of pricing?" that are asked when the manager and

purchasing team want to know if it is the best price that they have or not, or how they can negotiate with the supplier and which data they can rely on to have the best and effective negotiation with their vendors.

1.4. Research Objects

The research objective of this project is to gain insights from purchasing departments of the fictional company Adventure Works, thereby developing a BI solution for analytics to support better decision-making. Research components include:

- **Adventure Works** is the fictional company that the BI solution is developed for.
- **Datasource**: AdventureWorks 2019 database.
- SQL Server.
- Procurement business process.
- Microsoft Visual Studio Application.
- Multidimensional data analysis technique OLAP.
- KPI performance evaluation solution.
- Visualization tools such as Power BI.

1.5. Scope of the Project

- Time scope: Timeframe from February 22, 2024 to May 05, 2024.
- Dataset time scope: Including data from 2011 to 2014.
- Space scope: AdventureWorks Company, especially focusing on Purchasing department.

1.6. Project budgets (Write after learning PMIS)

In order to gather sufficient information for conducting a cost-benefit analysis of this project, the research team will proceed under the assumption that Adventure Work Cycles company contracted us for constructing a data warehouse on a cloud computing platform (Microsoft Azure). Moreover, as a global corporation, the company also bought the Microsoft Fabric Capacity Reservation SKUs to improve the experiences of data integration, data engineering, real-time analytics and business intelligence with PowerBI on a lake-centric SaaS solution.

Additionally, due to the lack of actual cost and benefit information, some information below is just estimated.

1.6.1. Net economic benefits

The research team has conducted a thorough analysis and identified several benefits. These benefits are based on the team's research of the industry's insights.

Table 1-2. Tangible benefits of project (Source: Authors Propose)

Building a data warehouse on Cloud Platform	
Tangible benefits	
A. Cost savings from optimizing hardware expenses	$12300 per year
B. Cost savings from streamlining data storage expenses	$5720 per year
C. Increased revenue from enhanced decision-making and insights	$4150 per year
D. Reduced labor costs through automated data processing	$3000 per year
Total:	**$25170 per year**

Through the team's analysis, it is evident that implementing a data warehouse on a cloud platform like Azure can harness significant tangible benefits for organizations. These tangible benefits include cost savings achieved through hardware optimization and streamlined data storage expenses. Additionally, enhanced decision-making and insights can lead to increased revenue, while automated data processing can result in reduced labor costs. The estimated total monthly benefits of $25170 further support the financial case for the project.

The projected tangible and intangible benefits that the company expects to achieve can be assessed through the implementation of the project. By quantifying the economic impact, including projected cost savings, revenue increase, and other financial advantages, the net economic benefits can be estimated.

1.6.2. One-time costs and recurring costs

In this project, the one-time costs and recurring costs associated with building a data warehouse must be considered.

Table 1-3. One-time costs of project (Source: Authors propose)

One-time costs	**Description**	**Estimated Cost**
System Development Costs	Costs for designing and building the data warehouse on Azure	$400
Data Migration and System Conversion Costs	Costs for data migration and system conversion	$300
User Training Costs	To ensure successful adoption of the new data warehouse and associated BI solutions, we need costs for user training within the organization	$500

Table 1-3. provides an overview of estimated one-time costs associated with the project. These costs encompass expenses related to system development, data migration and system conversion, as well as user training. It is essential to consider these one-time costs during project planning and budgeting processes, as they contribute to the overall project cost.

Table 1-4. Recurring costs of project (Source: Authors Propose)

Microsoft Azure Estimate					
Estimated Monthly Cost					
Service category	**Service**	**Region**	**Type**	**Description**	**Estimated Monthly Cost**

Databases	Azure SQL Database	East US	Pay-as-you-go	4 vCore, Business Critical, Azure Hybrid Benefit	$899
Analytics	Microsoft Fabric	East US			$263
Total:					$1162

Table 1-4. shows an estimate of the monthly cost based on power and usage of Azure services. Since Azure's billing policy is pay-as-you-go, the above cost may vary depending on the level of tasks. However, the above cost level will be the cost that the team decides to use to perform the cost-benefit analysis of the project, the services and resources used on each service are calculated to match the process project implementation. The cost per month is $1162, per year recurring costs are $13944.

1.6.3. Economic Feasibility Assessment

Based on the net economic benefits, one-time costs, and recurring costs, an economic feasibility assessment could be established. This assessment will involve calculating the Net Present Value (NPV) of the project, which takes into account the expected cash flows, discount rate, and time horizon. By comparing the NPV with the initial investment, the stakeholders can determine the profitability and financial viability of the project. Moreover, the project's viability and estimated return on investment (ROI) can also be calculated.

A note is that the cost data provided for the tangible benefits and one-time costs are estimated based on industry research and benchmarks. For accurate information regarding Azure pricing and recurring costs, it is recommended to refer to the official Microsoft Azure website and consult with Azure representatives. This will ensure accurate and up-to-date cost information for the company’s specific project.

The economic feasibility assessment helps managers a lot in decision-making and determining the financial viability of building a data warehouse on Azure. It is essential to

consider both the tangible benefits and the costs, including their sources and basis, to make informed decisions and develop a robust business case.

1.7. Project deliverables

1.7.1. Value of the Project

The value of this project is multifaceted, encompassing a comprehensive understanding of purchasing processes, BI solutions, KPI evaluation methodologies, and business reporting systems. By delving into the intricacies of purchasing processes, the project aims to enhance operational efficiency and identify areas for optimization. The exploration of BI solutions facilitates informed decision-making through data-driven insights, while a nuanced grasp of KPI evaluation methodologies ensures the accurate measurement of key performance indicators.

Furthermore, the project underscores the importance of comprehending the business reporting system, facilitating efficient and accurate communication of insights. One of its key objectives involves the meticulous analysis of Adventure Work's data to construct a data warehouse. This initiative not only aids in understanding the company's data landscape but also sets the foundation for informed decision-making and strategic planning. Ultimately, the project's value lies in its holistic approach to data management, analysis, and reporting, empowering the organization with invaluable insights for future endeavors.

1.7.2. Desired Outcome of the Project

The project unfolds across four stages, each directed towards specific facets of Business Intelligence (BI) system implementation:

- Stage 1 (Chapter 1): The focus here is on understanding the BI project, encompassing documentation techniques, in-depth business requirements analysis, and KPI examination. This yields a comprehensive project document, KPIs reflecting business needs, and a framework for BI and Data Warehouse. Additionally, a detailed design for the Purchasing module's Data Mart is outlined.

- Stage 2 (Chapter 2 and Chapter 3): This stage centers on data preparation and the deployment of the Data Warehouse on SQL Server. Outcomes involve processed and

integrated data in alignment with the design strategy, a detailed data model, a bus matrix, and the implementation of ETL/ELT tools for the Purchasing module.

- Stage 3 (Chapter 4): Emphasizing the importance of a multidimensional database, this stage explores multidimensional data analysis within the Purchasing module and applies OLAP techniques on SSAS. The result is a functional multidimensional database capable of robust data analysis.

- Stage 4 (Chapter 5): In the concluding stage, Power BI is leveraged to fashion dashboards and charts that mirror the requirements and KPIs. The ultimate deliverables include intuitive dashboards offering profound insights into business performance within the Purchasing module.

Upon project completion, a comprehensive Business Intelligence system is delivered, featuring a Data Warehouse, user manuals, and potent Power BI dashboards. This delivery adds substantial value to the organization, facilitating strategic decision-making grounded in accurate and detailed data analysis.

1.8. Research Process

The research process of the project can be divided into 9 following steps below:

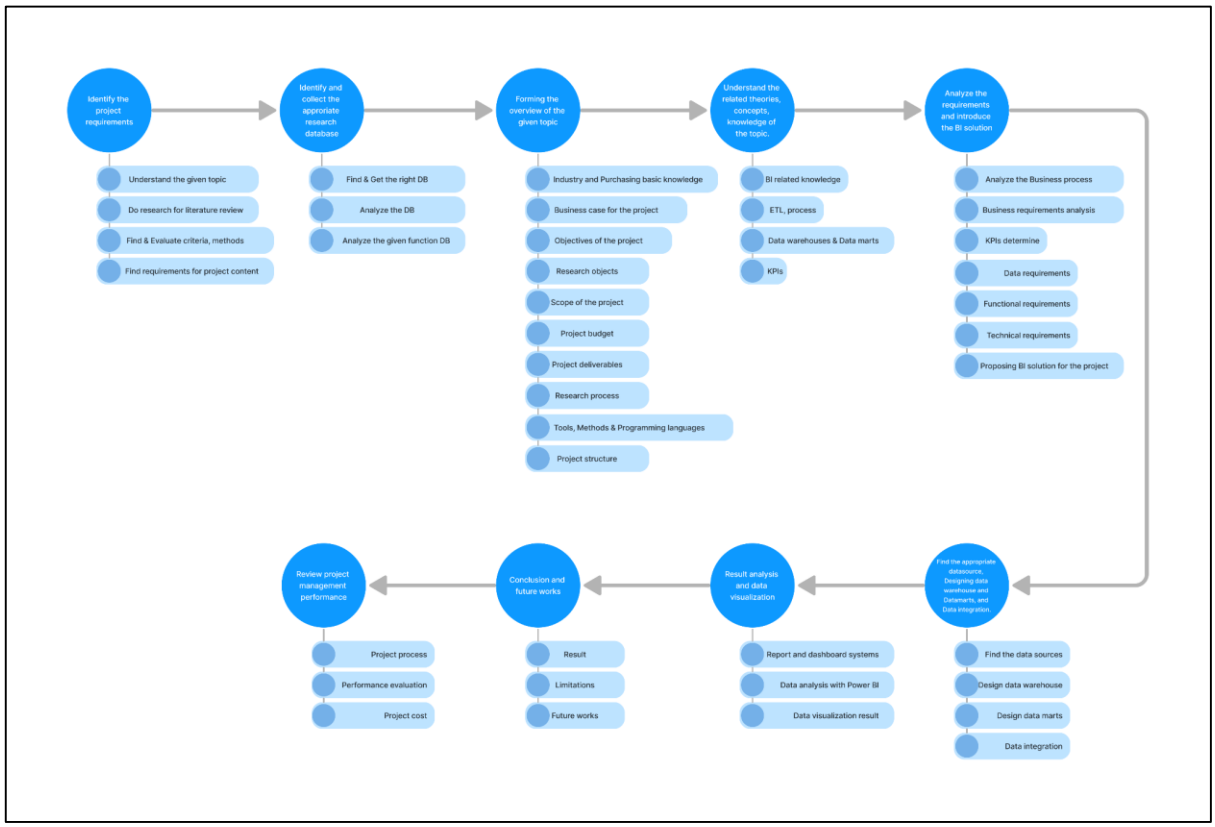


Figure 1.3. The research process (Source: Authors proposed)

- Identify the project requirements: In this step, our team delves into understanding the topic of improving the purchasing function within Adventureworks database. We conduct thorough research to review existing literature, evaluate criteria and methods, and define the project's content requirements, ensuring a clear direction for our BI solution.

- Identify and collect the appropriate research database: Our team focuses on sourcing and analyzing the right database for our project needs. This involves identifying and acquiring relevant databases, as well as meticulously analyzing both the overall database structure and the specific purchasing function data within Adventureworks.

- Forming the overview of the given topic: Here, we establish a comprehensive overview of the project, including fundamental industry and purchasing knowledge, the business case, project objectives, research objects, scope, budget, deliverables, process, tools, methods, programming languages, and project structure, ensuring clarity and alignment among team members.

- Understand the related theories, concepts, knowledge of the topic: We dive into relevant theories and concepts surrounding business intelligence (BI), including ETL processes, data warehouses, data marts, and key performance indicators (KPIs), equipping our team with the necessary theoretical foundation to address the project requirements effectively.

- Analyze the requirements and introduce the BI solution: This step involves a detailed analysis of the business process, requirements, and KPIs, determining both functional and technical requirements. Leveraging this analysis, we propose a tailored BI solution to address the purchasing function challenges within Adventureworks database.

- Find the appropriate data source, Designing data warehouse and Datamarts, and Data integration: Our team focuses on sourcing suitable data sources, designing an efficient data warehouse structure, creating data marts to serve specific analytical needs, and seamlessly integrating data from disparate sources to ensure a cohesive BI solution.

- Result analysis and data visualization: Here, we develop reporting and dashboard systems and utilize Power BI for data analysis and visualization, providing stakeholders with actionable insights derived from the BI solution implemented for the purchasing function enhancement.

- Conclusion and future works: We summarize the project's results, including its limitations, and outline potential future avenues for improvement or expansion, ensuring a thorough understanding of the project's outcomes and its potential implications for future endeavors.

- Review project management performance: Finally, our team conducts a comprehensive review of the project management performance, evaluating the overall project process, assessing performance against predetermined metrics, and analyzing project costs to identify areas for optimization and improvement in future projects.

In conclusion, through a systematic approach encompassing thorough research, strategic analysis, and implementation of a tailored BI solution, our team successfully addressed the purchasing function challenges within Adventureworks database, laying the

groundwork for enhanced efficiency and informed decision-making, while also identifying opportunities for future refinement and expansion.

1.9. Tools, Methods, and Programming Languages

1.9.1. SQL Server

- Objective: SQL Server is a relational database management system (RDBMS) developed by Microsoft. The objective is to store and manage data efficiently using a relational database.

- Task: AdventureWorks data is stored in SQL Server databases. The data can be queried, manipulated, and organized using SQL queries and stored procedures.

1.9.2. Microsoft Visual Studio

- Objective: Visual Studio is an integrated development environment (IDE) used for developing, testing, and deploying applications.

- Task: Developers can use Visual Studio to create custom applications, reports, or data integration solutions tailored to the specific needs of analyzing AdventureWorks data.

1.9.3. Multidimensional Data Analysis Technique (OLAP)

- Objective: Multidimensional analysis involves examining data from multiple perspectives to gain a comprehensive understanding.

- Task: OLAP techniques applied to AdventureWorks data allow users to analyze data based on different dimensions (e.g., time, product, geography) to uncover trends, patterns, and anomalies.

1.9.4. KPI (Key Performance Indicator) Performance Evaluation Solution

- Objective: KPIs are quantifiable measures used to evaluate the success of an organization or a particular activity.

- Task: Implementing a KPI solution allows tracking and evaluating key metrics within AdventureWorks data to assess performance and make informed decisions.

1.9.5. Visualization Tools (Power BI)

- Objective: Visualization tools are used to represent data visually, making it easier to understand and interpret.

- Task: Power BI can connect to AdventureWorks data, create interactive dashboards, and generate visualizations (charts, graphs, etc.) for better comprehension of data trends and patterns.

1.10. Structure of Project

The report is divided into 7 sections, including the following:

Chapter 1: Project Overview

Chapter 2: Defining requirements business/KPIs, data and quality

Chapter 3: Data preparation and Data modeling

Chapter 4: Data Integration

Chapter 5: Multi-dimensional data analysis

Chapter 6: Data Analytics and Visualization

Chapter 7: Conclusion and Future Works

Chapter 2: Defining requirements business/KPIs, data and quality

The initial phase of a BI project, where the chapter's content involves identifying and analyzing business requirements, along with concurrently conducting source data collection and analysis of the current purchasing process, aiming to develop a BI solution for the purchasing system and ensure the satisfactory resolution of the request.

2.1. Business requirements analysis

2.1.1. Business Requirements

2.1.1.1. General Business requirements of Adventureworks company

To achieve overarching goals of improving profitability, minimizing costs, and enhancing decision-making accuracy across all departments and processes there are some business requirements that Adventureworks company faces off.

Firstly, there is a need to focus on cost reduction initiatives to enhance profitability. This involves analyzing operational processes to identify inefficiencies and areas for optimization. By streamlining workflows, optimizing resource allocation, and negotiating favorable terms with suppliers and vendors, AdventureWorks aims to minimize overhead expenses and procurement costs. Additionally, implementing cost-effective measures for inventory management and logistics is crucial for achieving cost reduction objectives.

Secondly, AdventureWorks seeks to maximize revenue by diversifying its product offerings and enhancing marketing and sales efforts to capture additional market share. This includes developing strategies to identify new revenue streams, exploring opportunities for cross-selling and upselling to existing customers, and investing in research and development for innovative products and services.

To support informed decision-making, AdventureWorks have to establish robust reporting and analytics capabilities. This involves gathering and analyzing key performance indicators (KPIs) to provide stakeholders with real-time data and insights through intuitive dashboards and reports. By fostering a data-driven decision-making culture across all levels of the organization, AdventureWorks aims to ensure that strategic initiatives and operational decisions are informed by accurate and timely information.

2.1.1.2. Business requirements in Purchasing function

The business requirements in Purchasing function as shown in the table below:

Table 2-1. Business requirements in Purchasing (Source: Authors propose)

No.	Classified	Role	What	What for/How	Why/What for
1	Inventory	Inventory Manager	- A dashboard that shows indicators of inventory quantity/ specific product, location, product code, status. - A dashboard showing the number of available products and safety stock of that product. - A dashboard that shows a number of indicators from a product type or from a specific warehouse/inventory. - A dashboard shows the products that need to have alternative or	- To be able to know how much inventory there is of a specific product. - To calculate the total number of products, to be able to calculate the satisfaction of a certain need. - To evaluate and make purchasing decisions. - To have an overview of a product type, product in some aspects such as quantity, location, status. - Show the performance of	- Dashboard can be drill-down by level for easy analysis. - Allows Inventory manager to proactively compare indicators and search for information about the inventory of a certain product. - Allows inventory managers to rely on it to make decisions about making purchase requests. - Allows inventory managers to proactively take measures to reduce inventory storage costs. - Allow manager can identify if inventory

			additional suppliers.	managing inventory.	management is effective or not.
2	Procurement Performance Analysis	Purchasing team leader	- A dashboard that can show the average price and price fluctuations of a product in the past and the past average price of that product from vendors. - A dashboard shows the total partners that the company has. - A dashboard shows the products and their suppliers with ratings. - A dashboard that assesses the vendors in terms of price, PO's success rate, product quantity.	- Can help to determine the appropriate price for the products. - To track and identify the price trends of the products. - To evaluate the effectiveness of the purchasing decisions. - To help the purchasing team to have effective negotiation with vendors. - To compare and find the right vendor based on price and ratings. - To compare the vendors together based on the given factors and identify	- Enable the dashboard is visually intuitive, customizable, providing drill down for deeper analysis. - Enable the purchasing team to analyze the price based on time, vendors. - Help to find the best price of the product to decrease the costs and expenses. - Purchasing team is able to find the trends of the price of the product. - Enable the staff to evaluate the vendors based on price, PO's success rate. - To have better relationships with the best tier vendors. - To strengthen the relationship with potential vendors.

				the list of best vendors. - To segment the vendors for future business activities. - To make purchasing process more effective.	- To find additional or alternative vendors for the product.

2.1.2. KPI analysis

2.1.2.1. Inventory Management

- Main requirement: How can we minimize the costs and expenses of inventory management? What is the status of the specific product in terms of quality, quantity, and location? Do we need to make a purchase for the specific product? What is the status of the product in terms of current quantity, safety stock?

- Objective: Show the current quantity of products and compare with their safety stock. This comparison will help the Inventory manager determine if they need to make the purchasing decision for the product. This will help the manager can satisfy the order from the customer and also minimize the costs and expenses of the inventory management. Besides, showing the product data based on some indicators as the product ID, quantity, status, location also help the manager to easily manage the products in the inventory.

- Input: Inventory data (product quantity, safety stock quantity, product ID, location, etc.)

- Output: Comparison between the safety stock and current stock can advance in the performance of making purchasing decisions then can minimize the costs, expenses and increase the profit. By showing some important indicators of the product, the manager can capture the necessary information of the specific product and have the up-to-date information of that product. This report provides important insights and key information to

optimize cost structure, improve the profit margin and enhance the data-driven decision making.

- Business rule: data is consistent, reliable, and accurate across all sources.

- Key performance indicators:

+ *Product Quantity: present the quantity of each product.*

+ *Count of ProductID: this KPI counts all of the product ID.*

+ *Value on Hand: the monetary value of all of the stocks in the inventory.*

+ *SafetyStockLevel: Safety stock is a buffer of inventory kept to mitigate the risk of stockouts due to variability in demand or lead time.*

- MakePurchase: Gives the status of a product whether it needs to be purchased, needs to be considered for purchase or does not need to be purchased.

MakePurchase = "Urgent" if "Quantity" < "SafetyStock" *= "MakePurchase" if "SafetyStock" <= "Quantity" < "ReOrderPoint"* *= "NoNeed" if "Quantity" >= "ReOrderPoint"*	*(1)*

+ *VendorPerProduct:* The way to find do the product has more than 1 vendor or not. If they are from and under 1 vendor, it means that we are too dependent on that vendor and need to find an alternative supplier or additional supplier.

VendorPerProduct = Count(Vendor of a specific product)	*(2)*

+ *CountOfVendor by PostalCode:* present the number of vendors related to each postal code. This KPI can help define the nearest or the most suitable vendor based on the location.

2.1.2.2. Procurement Performance

- Main requirement: Who are our vendors? Which are our best partners, which are our potential vendors? What products do they supply us? Do we have to find the additional or alternative suppliers? Can we choose the vendors based on some indicators such as price,

lead time, quality or order rejected by vendor, product return rate? What is the appropriate price for the product? What is the trend in price of the product? How can we choose the right vendor in terms of pricing?

- Objective: The primary objective of vendor management is to effectively segment vendors, strengthen relationships with key partners, and identify potential alternative suppliers when necessary. By segmenting vendors into tiers based on performance metrics such as price, lead time, quality, and order rejection rates, the company can prioritize engagement with top-tier suppliers. Additionally, the aim is to identify suppliers who provide critical products or materials to establish strong and reliable partnerships.

- Input: Product data, Purchasing data (vendors data, PO data, ect),..

- Output: gaining insights into price trends, supporting negotiations with suppliers based on data-driven analysis, obtaining an overview of all suppliers, comparing suppliers based on various performance indicators, and assessing the success rate of placing purchase orders with each vendor. This analysis facilitates informed decision-making regarding supplier selection and relationship management. Furthermore, by identifying products that require additional suppliers to prevent shortages and reduce dependency on single suppliers, the company can proactively address customer needs and enhance supply chain resilience.

- Business rule: data is consistent, reliable, and accurate across all sources.

- Key performance indicators:

+ *TotalNumberOfPO:* present the number of purchase orders.

+ *TotalPurchaseAmount:* present the total purchase amount that the company spent.

+ *TotalTaxAmount:* present the total tax amount from all purchase orders.

+ *AverageTaxAmountPerPO:* present the average tax amount of each purchase order.

AverageTaxAmountPerPO = TotalTaxAmountTotalNumberOfPO	*(3)*

+ *AveragePrice:* present the average price of each product.

+ *AveragePriceByVendor:* present the average price of each product by each vendor. These KPIs can be used to find if the given price from the supplier is good or not.

+ *POSuccessRate:* show the successful rate of placing a purchase order.

POSuccessRate = Number of received Purchase orderTotal quantity Purchase orderx100	*(4)*

- *POReturnRate:* show the percentage of product was rejected and return to the supplier. Shows the quality of a order supplied by a vendor.

POReturnRate = Total number of rejected productTotal number of receivedx100	*(5)*

- *POLeadTime:* the amount of time between when a purchase order is placed to replenish products and when the order is received in the warehouse. The smaller the PO lead time is the better the supplier is.

POLeadTime = Received date - Order date	*(6)*

- *AveragePOLeadTime:* present the average of PO's lead time.

AveragePOLeadTime = POLeadTimeTotalNumberOfPO	*(7)*

- *CreditRating*: Rating whether a vendor is good or not:

CreditRating 1 = Superior, 2 = Excellent, 3 = Above average, 4 = Average, 5 = Below average	*(8)*

2.1.3. Data and data quality requirements

In terms of data, the BI system will integrate a diverse range of information, including purchase orders, supplier details, and inventory levels, offering a comprehensive overview of the purchasing operations. To facilitate timely decision-making, the system must seamlessly handle real-time or near-real-time data feeds from sources such as the company's ERP system and supplier databases. Historical data inclusion is imperative for robust trend analysis, enabling stakeholders to glean insights into past purchasing behaviors and performance.

Additionally, the system will incorporate detailed metrics related to supplier performance, encompassing delivery times, product quality, and reliability. This wealth of information is pivotal for evaluating and optimizing the efficiency of supplier relationships. Equally critical is comprehensive data on inventory levels, stockouts, and reorder points, providing essential insights into material availability and flow.

On the front of data quality, accuracy is paramount, demanding validation checks and reconciliation processes to rectify any inconsistencies. Consistency across diverse data sources will be achieved through standardized formats and reconciliation protocols. Timeliness of updates is crucial, particularly for dynamic aspects such as purchase orders and inventory levels. Ensuring data completeness across all relevant fields, from purchases to suppliers and inventory, is essential to prevent gaps and foster a comprehensive understanding of the purchasing landscape.

Furthermore, robust security measures will be implemented to safeguard sensitive purchasing data, adhering to privacy regulations and restricting access based on defined roles and responsibilities. Collectively, these meticulous data and data quality requirements aim to establish the BI solution as a reliable, comprehensive, and timely analytical platform, empowering stakeholders with accurate insights for informed decision-making within the Purchasing department at Adventure Works.

2.1.4. Functional requirements

The functional requirements for the implementation of the Business Intelligence (BI) system in Adventure Works' Purchasing department encompass a range of capabilities designed to enhance operational efficiency and decision-making processes. The system should deliver real-time data updates, enabling stakeholders to access the latest information on purchase orders, supplier performance, and inventory levels. Ad Hoc query capabilities offer users the flexibility to explore data dynamically, responding to evolving analytical needs.

Furthermore, the inclusion of forecasting and predictive analytics empowers the Purchasing department to anticipate future demand and optimize inventory levels proactively. Automated alerts and notifications ensure timely responses to critical events,

while collaborative decision-making tools foster a culture of shared insights and recommendations among stakeholders.

Automated report generation will simplify information dissemination, letting users schedule and receive regular reports on key purchasing metrics. For enhanced collaboration, a dedicated portal within the BI system will be established, fostering real-time communication, document sharing, and performance tracking with suppliers.

2.1.5. Technical requirements

Deploying a Business Intelligence (BI) solution for Adventure Works' Purchasing department requires key technical elements. Firstly, seamless integration with ERP systems, supplier databases, and inventory management is essential. Collaboration with stakeholders clarifies the scope of data for analysis.

A well-designed data model and robust data warehouse are critical, consolidating diverse data for a unified view of the Purchasing department's operations. Efficient Extract, Transform, and Load (ETL) processes ensure data quality and regular updates.

The BI tool should seamlessly connect to the data warehouse, offering interactive dashboards, reports, and visualizations tailored to Purchasing metrics. Built-in analytics must present insights in a user-friendly format for stakeholders with varying technical expertise.

Security measures, including strict access controls, data encryption, and user activity monitoring, are imperative. Compliance with data privacy regulations ensures ethical data handling.

2.2. Comparative Analysis of BI solutions

Table 2-2. Comparative Analysis of BI solutions (Source: Authors propose and summarize)

Tools	Pros	Cons
Power BI	- **Easy to use**: Great for non-technical users for data	- **Limited data transformation:** Relies on external tools for

	exploration and visualization. - **Microsoft integration**: Integrates seamlessly with Excel and SQL Server for a smooth workflow. - **Cloud-based**: Accessible from anywhere with an internet connection. - **Cost-effective**: Freemium model with paid options for additional features. - **Visually appealing**: Offers a wide range of interactive and customizable dashboards.	complex data manipulation. - **Scalability concerns:** performance can be impacted with very large datasets. - **Vendor lock-in:** May limit flexibility if you heavily rely on Microsoft products.
SSIS	- **Powerful data integration:** Efficiently automates data extraction, transformation, and loading (ETL) processes. - **Scalability:** Handles large and complex datasets effectively. - **Customization:** Offers a high degree of control over data pipelines. - **Security:** Integrates with SQL Server security features for robust data protection.	- **Steeper learning curve:** Requires technical knowledge for development and maintenance. - **Limited visualization:** Primarily focused on data movement, not data exploration. - **Cost:** Requires a SQL Server license, which can be expensive.

SSAS	- **OLAP capabilities:** Enables multidimensional data analysis for complex calculations and drill-down capabilities. - **Fast query performance:** Optimized for efficient data retrieval and analysis. - **Security:** Integrates with SQL Server security features for data access control.	- **Complexity:** Requires a high level of technical expertise for setup and management. - **Cost:** Requires a SQL Server license, adding to the cost. - **Limited visualization:** Primarily focused on data analysis, not data presentation.
Tableau	- **Advanced visualizations:** Offers exceptional customization and interactivity for creating stunning dashboards. - **Ease of use:** Relatively easy to learn for data exploration and visualization. - **Large community:** Extensive resources and support available online.	- **Cost:** Can be expensive for large deployments, especially with additional features. - **Desktop focus:** Primarily a desktop application, limiting collaboration features. - **Data transformation limitations:** Relies on external tools for complex data manipulation tasks.

2.3. Proposing BI Solution for the Project

After evaluating AdventureWorks' current operations, it's evident that while the Purchasing departments are efficiently organized and well-coordinated, there are opportunities for improvement in warehousing management, purchase order preparation, and supplier evaluation and management. Implementing a Business Intelligence (BI) solution can address these limitations effectively. With BI tools, AdventureWorks can gain

valuable insights into its operations, enabling smarter decision-making, streamlined processes, and ultimately, improved profitability.

For inventory management, a comprehensive BI solution should include dashboards tailored for the Inventory Manager. These dashboards would display crucial indicators such as inventory quantity, product location, status, and cost of goods sold (COGS). This would allow the Inventory Manager to track inventory levels, calculate total products available, evaluate purchasing decisions, and assess overall inventory performance. Additionally, the dashboards should be drill-down capable for easy analysis, empowering the Inventory Manager to proactively manage inventory and make informed decisions regarding purchasing requests and inventory storage costs.

BI dashboards designed for the Purchasing Team Leader would focus on price negotiation and vendor management. These dashboards would provide insights into historical product prices, price fluctuations, vendor pricing trends, and vendor evaluations based on factors like price, delivery time, quality, and return rate. Armed with this information, the Purchasing Team Leader can determine appropriate product prices, track price trends, negotiate effectively with vendors, and identify the best vendors based on price and other criteria. Moreover, dashboards should facilitate vendor segmentation, fostering better relationships with top-tier vendors and identifying additional or alternative suppliers when needed.

Overall, this five-part BI solution will enhance AdventureWorks' operational efficiency, drive sustainable growth, and pave the way for continued success in an increasingly competitive market landscape.

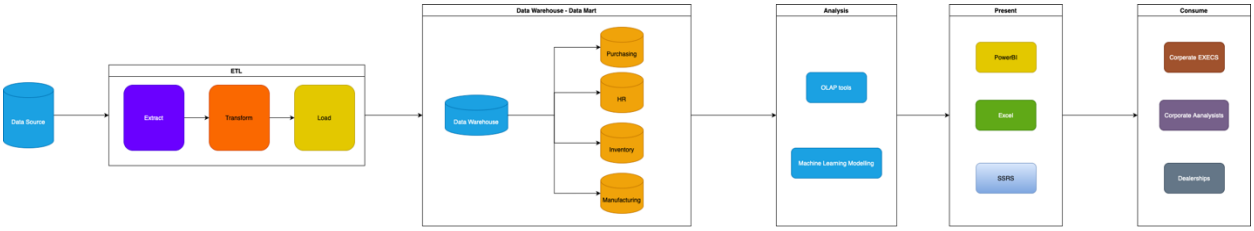


Figure 2.1. Proposed BI Solution (Source: Authors propose)

Chapter 3: Data preparation and Data modeling

In this chapter, exploratory data analysis is performed to get a comprehensive understanding of the database, while concurrently identifying the facts and dimensions that align with the analyzed requirements. Additionally, a suitable model for the data warehouse will be designed.

3.1. Data preparation

3.1.1. Overview of AdventureWorks Data

The AdventureWorks database serves as a sample database designed for Microsoft SQL Server versions 2008 through 2014. This database is documented using Dataedo and is structured to support typical online transaction processing scenarios for a fictional bicycle manufacturer known as Adventure Works Cycles. The various scenarios covered by the database include modules such as: Business Entities, People, Human Resources, Products, Manufacturing, Purchasing, Inventory, Sales and Admin. Through its comprehensive design, AdventureWorks provides a practical and illustrative platform for showcasing the functionalities of Microsoft SQL Server in a simulated business environment.

Data Structure

In the AdventureWorks database, data is organized into schemas to create a clear and manageable data structure. Each schema contains data objects related to a specific part of the company's business activities. Objects such as tables, views, and procedures are organized into schemas within the AdventureWorks database. This mechanism helps organize data in a structured and manageable way, providing flexibility in accessing and processing data. Starting from SQL Server 2005, the concept of schemas has been separated from the concept of users, where users own schemas, and schemas contain database objects.

Table 3-1. Schema in AdventureWorks database (Source: Học viện đào tạo trực tuyến, 2017)

Name of Schema	Description	Total of Tables
Human Resources	Employees of the Adventure Works Cycles company.	6
Person	Names and addresses of individual customers, suppliers, and employees.	13
Production	Products manufactured and sold by the Adventure Works Cycles company.	25
Purchasing	Suppliers of parts and other products purchased by the company.	5
Sales	Customers and data related to sales transactions.	18

Below is the comprehensive schema used in the AdventureWorks database.

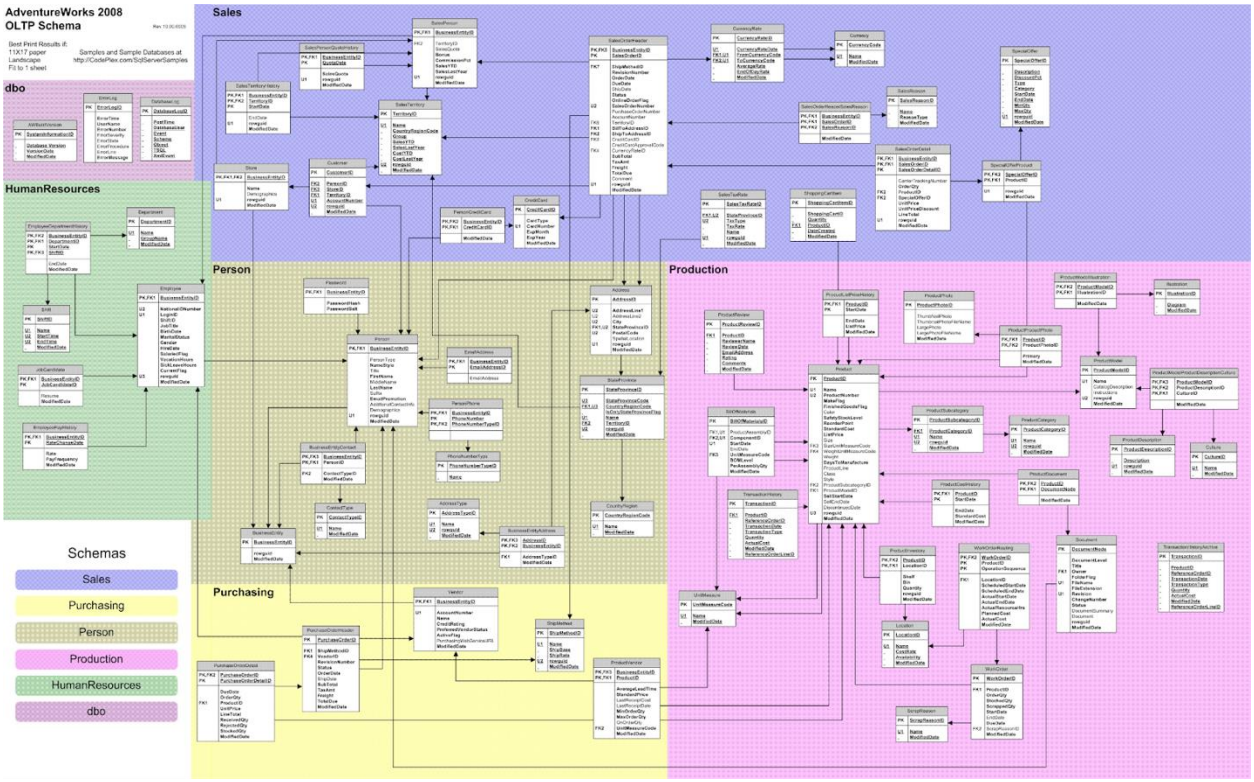


Figure 3.1. AdventureWorks ERD (Source: Học viện đào tạo trực tuyến, 2017)

3.1.2. Data Collection

During the data preparation phase, we gathered information from various sources within the AdventureWorks dataset, including:

- AdventureWorks Database: Direct queries were executed against the AdventureWorks database to retrieve relevant data. This involved accessing tables and views associated with different modules such as sales, purchasing, human resources, and manufacturing.

- Microsoft Documentation: Microsoft provides comprehensive documentation and guides related to the AdventureWorks database schema and usage. These resources were consulted to understand the database structure and relationships between different entities.

- Online Resources: Additionally, we leveraged online resources such as forums, blogs, and tutorials dedicated to AdventureWorks. These sources provided valuable insights, tips, and best practices for working with the dataset.

3.1.3. Data Description and Data Understanding

This report focuses on the primary study of the Purchasing module, an essential component of the Purchasing schema. We are discussing the structure, features, and data related to the purchasing process within the AdventureWorks system. Additionally, we also introduce the Inventory module under the Products schema, another critical part of the system. The Inventory subsystem has a crucial connection with the Purchasing subsystem, as it involves inventory management and control, which is indispensable in the purchasing process.

3.1.3.1. Module Purchasing

The Purchasing module consists of five tables.

a. Purchasing.ProductVendor:

Table 3-2. Table Purchasing.ProductVendor (Source: Dataedo, 2024)

No.	Key	Name	Data type	Description
1	PK	ProductID	int	Primary key. Foreign key to Product.ProductID.
2	PK	BusinessEntityID	int	Primary key. Foreign key to Vendor.BusinessEntityID.
3		AverageLeadTime	int	The average span of time (in days) between placing an order with the vendor and receiving the purchased product.
4		StandardPrice	money	The vendor's usual selling price.
5		LastReceiptCost	money	The selling price when last purchased.
6		LastReceiptDate	datetime	Date the product was last received by the vendor.

7		MinOrderQty	int	The maximum quantity that should be ordered.
8		MaxOrderQty	int	The minimum quantity that should be ordered.
9		OnOrderQty	int	The quantity currently on order.
10		UnitMeasureCode	nchar(3)	The product's unit of measure.
11		ModifiedDate	datetime	Date and time the record was last updated.

This table illustrates the linkage between product information and their respective suppliers within the system. It contains information regarding the relationship between specific products and the suppliers providing them. In overview, the table demonstrates the following factors:

- Product-supplier relationship: The Purchasing.ProductVendor table establishes a relationship between products and suppliers. Through the foreign keys ProductID and BusinessEntityID, this table associates specific product information with the suppliers providing them. This enables the system to track which suppliers provide which products, and vice versa.

- Transactional purchase information: This table provides detailed information about purchasing transactions with suppliers. Through fields such as AverageLeadTime, StandardPrice, LastReceiptCost, and LastReceiptDate, the system can monitor information about supply lead time, pricing, and purchase transaction history with each supplier.

- Order management: Through fields like MinOrderQty, MaxOrderQty, and OnOrderQty, this table supports the management of the ordering process with suppliers. This information ensures that orders are placed in appropriate quantities and tracks the current status of orders.

- Inventory management and unit measurement: The UnitMeasureCode field provides information about the unit of measurement for the product, which is crucial for

maintaining consistency in inventory management. This information can also be used to convert between different units of measurement as needed.

- Edit history: The ModifiedDate field records the time the record was last updated. This helps track the edit history of information in the table and ensures the integrity and reliability of the data.

b. Purchasing.PurchaseOrderDetail

Table 3-3. Table Purchasing.PurchaseOrderDetail (Source: Dataedo, 2024)

No.	Key	Name	Data type	Description
1	PK	PurchaseOrderID	int	Primary key. Foreign key to PurchaseOrderHeader.PurchaseOrderID.
2	PK	PurchaseOrderDetailID	int	Primary key. One line number per purchased product. Identity
3		DueDate	datetime	Date the product is expected to be received.
4		OrderQty	smallint	Quantity ordered.
5		ProductID	int	Product identification number. Foreign key to Product.ProductID.
6		UnitPrice	money	Vendor's selling price of a single product.
7		LineTotal	money	Per product subtotal. Computed as OrderQty * UnitPrice. Computed: isnull([OrderQty]*[UnitPrice],(0.00))

8		ReceivedQty	decimal(8, 2)	Quantity actually received from the vendor.
9		RejectedQty	decimal(8, 2)	Quantity rejected during inspection.
10		StockedQty	decimal(9, 2)	Quantity accepted into inventory. Computed as ReceivedQty - RejectedQty. Computed: isnull([ReceivedQty]-[RejectedQty],(0.00))
11		ModifiedDate	datetime	Date and time the record was last updated.

The Purchasing.PurchaseOrderDetail table serves as a comprehensive repository for detailed information regarding items included in purchase orders. In overview, the table demonstrates the following factors:

- Purchase Order Identification and Line Items: The table facilitates the management of purchase orders by providing a detailed breakdown of items included in each order. Through the PurchaseOrderID and PurchaseOrderDetailID primary keys, the table uniquely identifies each purchase order and its individual line items.

- Delivery Schedule: The DueDate field specifies the expected date for receiving each product. This information is crucial for planning and tracking the delivery schedule of ordered items.

- Quantity and Pricing: The table captures the quantity of each product ordered (OrderQty) and its corresponding unit price (UnitPrice). These fields are essential for calculating the total cost of each line item (LineTotal).

- Receipt and Inspection: The ReceivedQty field records the actual quantity of items received from the vendor, while the RejectedQty field indicates the quantity rejected during

inspection. These quantities provide insights into the accuracy and quality of received goods.

- Inventory Management: The StockedQty field represents the quantity of items accepted into inventory after considering any rejections. It helps maintain accurate inventory levels and supports efficient stock management processes.

- Modification History: The ModifiedDate field tracks the date and time of the last update to each record. This information enables the monitoring of changes made to purchase order details over time, ensuring data integrity and accountability.

c. Purchasing.PurchaseOrderHeader

Table 3-4. Table Purchasing.PurchaseOrderHeader (Source: Dataedo, 2024)

No.	Key	Name	Data type	Description
1	PK	PurchaseOrderID	int	Primary key.
2		RevisionNumber	tinyint	Incremental number to track changes to the purchase order over time.
3		Status	tinyint	Order current status. 1 = Pending; 2 = Approved; 3 = Rejected; 4 = Complete
4		EmployeeID	int	Employee who created the purchase order. Foreign key to Employee.BusinessEntityID.
5		VendorID	int	Vendor with whom the purchase order is placed. Foreign key to Vendor.BusinessEntityID.
6		ShipMethodID	int	Shipping method. Foreign key to ShipMethod.ShipMethodID.

7		OrderDate	datetime	Purchase order creation date.
8		ShipDate	datetime	Estimated shipment date from the vendor.
9		SubTotal	money	Purchase order subtotal. Computed as SUM(PurchaseOrderDetail.LineTotal)for the appropriate PurchaseOrderID.
10		TaxAmt	money	Tax amount.
11		Freight	money	Shipping cost.
12		TotalDue	money	Total due to the vendor. Computed as Subtotal + TaxAmt + Freight.
13		ModifiedDate	datetime	Date and time the record was last updated.

The Purchasing.PurchaseOrderHeader table within the Purchasing schema manages overarching information related to purchase orders. In overview, the table demonstrates the following factors:

- Purchase Order Identification and Tracking: The PurchaseOrderID field serves as the primary key, uniquely identifying each purchase order within the system. It enables efficient tracking and referencing of specific orders.

- Revision History: The RevisionNumber field tracks revisions made to the purchase order over time. This incremental number allows for the monitoring of changes and version control.

- Order Status: The Status field indicates the current status of the purchase order, providing insights into its progress. Status values include Pending, Approved, Rejected, and Complete, facilitating easy monitoring of order lifecycles.

- Employee and Vendor Information: The EmployeeID field identifies the employee who created the purchase order, linking to the Employee.BusinessEntityID table. The VendorID field specifies the vendor with whom the purchase order is placed, referencing

the Vendor.BusinessEntityID table. These fields establish relationships between the purchase order and relevant entities.

- Shipping Method and Dates: The ShipMethodID field denotes the shipping method for the order, referencing the ShipMethod.ShipMethodID table. OrderDate records the creation date of the purchase order, while ShipDate estimates the shipment date from the vendor. These fields aid in managing shipping logistics and timelines.

- Financial Details: SubTotal calculates the purchase order subtotal by summing the LineTotal values from associated PurchaseOrderDetail records. TaxAmt represents the tax amount applicable to the order, while Freight denotes the shipping cost. TotalDue computes the total amount due to the vendor, including the subtotal, tax, and freight costs. These fields provide comprehensive financial information related to the order.

- Modification History: ModifiedDate records the date and time of the last update to the purchase order record. This information facilitates tracking of changes and ensures data integrity.

d. Purchasing.ShipMethod

Table 3-5. Table Purchasing.ShipMethod (Source: Dataedo, 2024)

No.	Key	Name	Data type	Description
1	PK	ShipMethodID	int	Primary key for ShipMethod records.
2	AK	Name	nvarchar(50)	Shipping company name.
3		ShipBase	money	Minimum shipping charge.
4		ShipRate	money	Shipping charge per pound.
5	AK	rowguid	uniqueidentifier	ROWGUIDCOL number uniquely identifying the record. Used to support a merge replication sample.

6		ModifiedDate	datetime	Date and time the record was last updated.

This table serves as a centralized repository for managing shipping methods utilized by the company:

- Shipping Information Management: Through the ShipMethodID field as the primary key, meaning each shipping method is identified by a unique ShipMethodID, and the Name field, where each shipping method has a distinct and non-duplicated name, the data provides detailed information about shipping methods. This facilitates the management and selection of appropriate shipping methods for orders, allowing for easy retrieval and efficient management of shipping information.

- Shipping Charges: Understanding how shipping fees are calculated by mastering ShipBase (base fee) and ShipRate (fee per pound) makes it easier to calculate shipping costs for specific orders.

- Global Unique Identifier: Each record in the table has a rowguid, which is a globally unique identifier. This data can be useful in scenarios involving data synchronization across different databases.

- Modification Tracking: Understanding how to track changes in data through the ModifiedDate field allows for monitoring of the edit history and ensures data integrity. It provides essential information for trend analysis and decision-making.

e. Purchasing.Vendor

Table 3-6. Table Purchasing.Vendor (Source: Dataedo, 2024)

No.	Key	Name	Data type	Description
1	PK	BusinessEntityID	int	Primary key for Vendor records. Foreign key to BusinessEntity.BusinessEntityID

2	AK	AccountNumber	nvarchar(15)	Vendor account (identification) number.
3		Name	nvarchar(50)	Company name.
4		CreditRating	tinyint	1 = Superior, 2 = Excellent, 3 = Above average, 4 = Average, 5 = Below average
5		PreferredVendorStatus	bit	0 = Do not use if another vendor is available. 1 = Preferred over other vendors supplying the same product.
6		ActiveFlag	bit	0 = Vendor no longer used. 1 = Vendor is actively used.
7		PurchasingWebServiceURL	nvarchar(1024)	Vendor URL.
8		ModifiedDate	datetime	Date and time the record was last updated.

This table stores information about vendors. In overview, the table demonstrates the following factors:

- Vendor Information Management: The Purchasing.ProductVendor table provides a database for storing and managing information about vendors. Fields such as BusinessEntityID, AccountNumber, and Name allow the identification of a specific vendor, while CreditRating provides an assessment of their reliability in business transactions.

- Identifying Preferred Vendors: The PreferredVendorStatus field allows marking a vendor as preferred over others supplying the same product. This is useful for the system

to automatically select vendors when necessary, optimizing the ordering process and maximizing the benefits from vendor relationships.

- Managing Vendor Status: The ActiveFlag field determines whether a vendor is active or not. This helps in removing vendors that are no longer active or suitable for use, ensuring data cleanliness and maintenance in the system.

- Linking with Other Information in the System: The BusinessEntityID field is linked to the BusinessEntity table, connecting vendor information with other data in the system, such as business partner information or financial data.

- Tracking Changes and Updating Information: The ModifiedDate field records the last time vendor information was updated in the system, aiding in tracking and managing the change history of information.

3.1.3.2. Module Inventory

The Inventory module consists of two tables.

a. Production.Location

Table 3-7. Table Production.Location (Source: Dataedo, 2024)

No.	Key	Name	Data type	Description
1	PK	LocationID	smallint	Primary key for Location records.
2	AK	Name	nvarchar(50)	Location description.
3		CostRate	smallmoney	Standard hourly cost of the manufacturing location.
4		Availability	decimal(8, 2)	Work capacity (in hours) of the manufacturing location.
5		ModifiedDate	datetime	Date and time the record was last updated.

This table manages information about manufacturing locations within the organization. In overview, the table demonstrates the following factors:

- Location Identification: The table includes a primary key field LocationID, which uniquely identifies each location record. This key ensures efficient retrieval and organization of location-related data.

- Descriptive Information: The Name field provides a description or name for each manufacturing location. This descriptive information helps users identify and differentiate between various locations within the organization.

- Cost and Financial Data: The CostRate field stores the standard hourly cost associated with each manufacturing location. This information is crucial for financial analysis, budgeting, and cost estimation in production operations.

- Work Capacity: The Availability field represents the work capacity of each manufacturing location, measured in hours. This data quantifies the amount of time available for production activities at each location, aiding in resource planning, scheduling, and production optimization.

- Record Management: The ModifiedDate field tracks the date and time when each location record was last updated. This allows for auditing and tracking changes to location information over time, ensuring data accuracy and integrity.

b. Production.ProductInventory

Table 3-8. Table Production.ProductInventory (Source: Dataedo, 2024)

No.	Key	Name	Data type	Description
1	PK	ProductID	int	Product identification number. Foreign key to Product.ProductID.
2	PK	LocationID	smallint	Inventory location identification number. Foreign key to Location.LocationID.
3		Shelf	nvarchar(10)	Storage compartment within an inventory location.

4		Bin	tinyint	Storage container on a shelf in an inventory location.
5		Quantity	smallint	Quantity of products in the inventory location.
6		rowguid	uniqueidentifier	ROWGUIDCOL number uniquely identifying the record. Used to support a merge replication sample.
7		ModifiedDate	datetime	Date and time the record was last updated.

This table is integral to the management of inventory within the organization. In overview, the table demonstrates the following factors:

- Identification of Products and Locations: The table facilitates the association of products with specific inventory locations. Through the ProductID and LocationID fields, it establishes links to the Product and Location tables, respectively. This relational structure enables efficient tracking and management of inventory across various locations.

- Storage Details: The Shelf and Bin fields provide detailed information about the physical storage of products within inventory locations. This granular data allows for precise organization and retrieval of items, optimizing warehouse operations and reducing handling errors.

- Inventory Quantities: The Quantity field indicates the quantity of products stored at each inventory location. This information is crucial for inventory management, enabling organizations to monitor stock levels, forecast demand, and plan replenishment activities effectively.

- Support for Replication: The presence of the rowguid field, containing globally unique identifiers (GUIDs), indicates support for replication scenarios. This feature facilitates data synchronization across distributed databases, ensuring consistency and reliability in multi-location environments.

- Record Maintenance: The ModifiedDate field tracks the timestamp of the last update to each inventory record. This allows for the monitoring of inventory changes over time, facilitating auditing, version control, and data integrity maintenance.

3.1.4. EDA and data cleaning

After gaining a comprehensive understanding of the data to be used, a more detailed analysis is conducted on the columns of these datasets. Exploratory Data Analysis has been performed on the ProductVendor, PurchaseOrderDetail, PurchaseOrderHeader, ShipMethod, Vendor, Location, and ProductInventory tables.

Null Ratio

The results of checking for null values show that in many tables, the proportion of null values exceeds 25%. Therefore, based on the null ratio and the importance of the data column for the Purchasing processes as well as the requirements set forth, columns with more than 30% empty values and are not important for the given problem will not be used, as presented in the table below.

Table 3-9. Null Ratio ((Source: Authors propose)

No.	Table's Name	Feature	Null Ratio
1	Product.Product	Color	49.20%
2	Product.Product	Size	57.34%
3	Product.Product	SizeUnitMeasureCode	58.33%
4	Product.Product	WeightUnitMeasureCode	59.33%
5	Product.Product	Weight	59.33%
6	Product.Product	DaysToManufacture	7.54%
7	Product.Product	ProductLine	44.84%
8	Product.Product	Class	43.45%
9	Product.Product	Style	64.88%
10	Product.Product	ProductSubcategoryID	42.26%
11	Product.Product	ProductModelID	41.47%
12	Product.Product	SellEndDate	52.38%

13	Product.Product	DiscontinuedDate	85.71%
14	Purchasing.Vendor	PurchasingWebServiceURL	96.15%
15	Person.CountryRegion	AddressID	25.47%
16	Person.CountryRegion	AddressLine1	25.16%
17	Person.CountryRegion	ModifiedDate	25.16%

Additionally, a table related to Purchasing processes, Product.Product, was found to have many columns containing numerous empty values. Figure 3.2. illustrates that many columns have a null value ratio exceeding 50%. In this table, features with many empty values are removed, except for ProductSubcategoryID, which is retained due to its importance for calculating KPIs, despite having more than 50% null values.

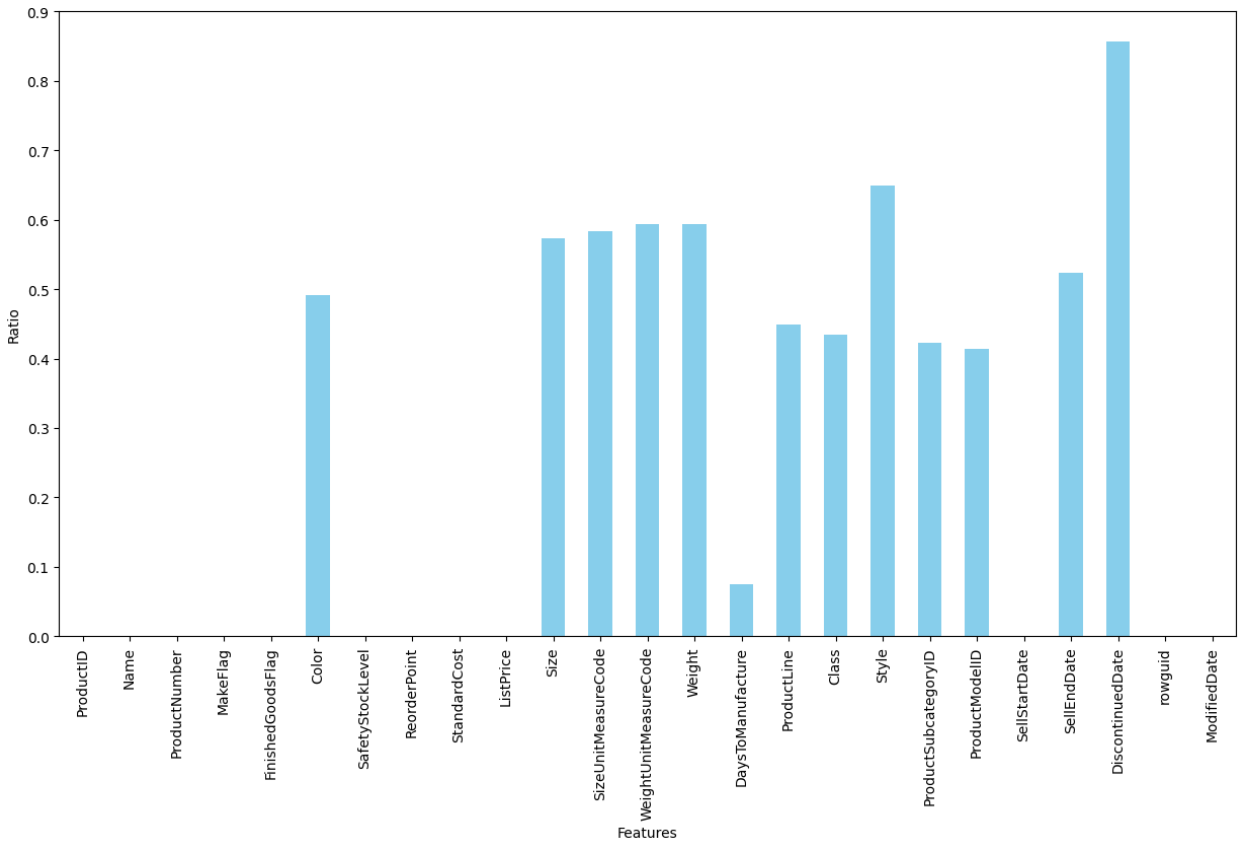


Figure 3.2. Ratio of null values (Source: Authors propose)

Outlier

Most columns in the tables have a short range of values, thus containing few outliers. However, the Product.Product table has 2 columns containing many outliers, namely ListPrice and DaysToManufacture. These large outliers are handled by setting them to the maximum value (Upper Limit) using the IQR method.

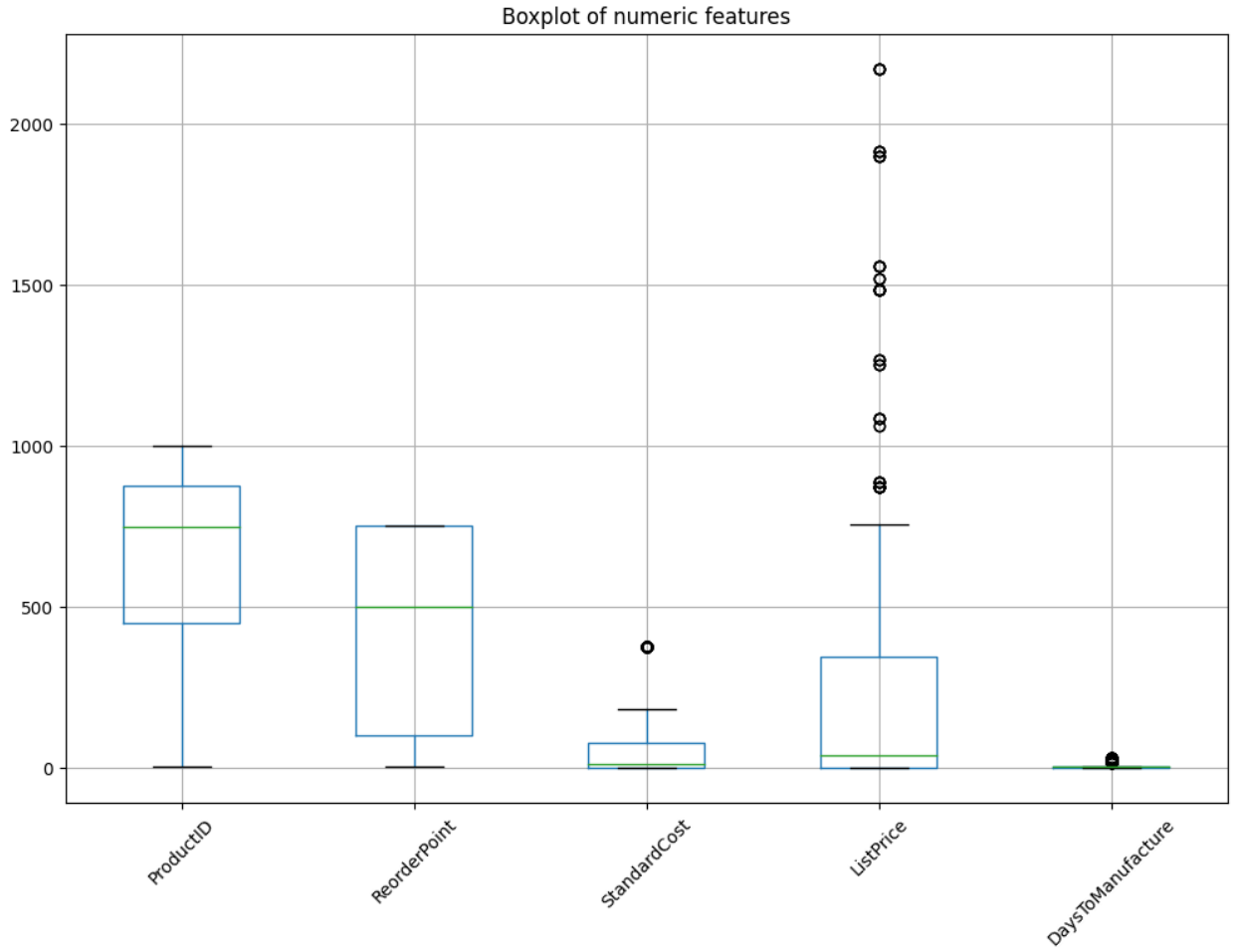


Figure 3.3. Boxplot of numeric features (Source: Authors propose)

3.2. Data Warehouse Design

3.2.1. Bus Matrix

To build up a data warehouse, Bus Matrix is a common technique to clearly define the project's business processes (business requirements) and the related dimensions for each requirement. Based on the business requirements mentioned in section 2.1., our team has defined a bus matrix consisting of 3 facts and 5 dimensions that could be included in the solution's Data Mart:

- Facts: FactInventoryManagement, FactProcurementPerformance.
- Dimensions: DimTime, DimProduct, DimVendor, DimGeography, DimLocation.

Table 3-10. Bus Matrix (Source: Authors propose)

BUSINESS PROCESSES	**COMMON DIMENSIONS**				
	DIM Time	**DIM Product**	**DIM Vendor**	**DIM Geography**	**DIM Location**
Inventory Management	X	X			X
Procurement Performance	X	X	X	X	

For the business requirement of Inventory Management, our team has built relevant dimensions including DIM Time, DIM Product, and DIM Location. By applying time, product, and location dimensions, businesses can gain insights into inventory trends over time and for specific product types. Moreover, utilizing the location dimension helps optimize warehouse management and efficient goods distribution, ensuring effective storage and management of products.

In the case of Procurement Performance analysis, the application of dimensions such as DIM Time, DIM Product, DIM Vendor, and DIM Geography demonstrates close interaction. Time is utilized to monitor price trends and assess price fluctuations over time, while product and vendor dimensions aid in comparing prices from various sources. Combined with the geographical dimension, businesses can search for the best-priced vendors for each product in different regions, this dimension can bring a clearer view of the vendor's information.

3.2.2. Master Data

In the context of the Purchasing module, the master data typically refers to the core data elements that provide foundational information about entities involved in the purchasing process.

Master data serves as the backbone of a data mart, providing the necessary context and reference points for analyzing transactional data related to purchasing activities. The following *Table 3-2* gives some sights about what master data will be used in terms of this project.

Table 3-11. Master Data (Source: Dataedo, 2024)

Object	Description
Product	Products that were purchased from vendors (Materials & Trading Goods) or generated through manufacturing.
Vendor	Companies from whom Adventure Works Cycles purchases parts or other goods.
Geography	Describe the vendors' addresses using the view: Purchasing.vVendorWithAddresses.
Location	Describe the locations of product storage warehouses.

3.2.3. Transactional Data

Aside from the master data, a data mart also depends on data generated through records of individual purchasing transactions which is called transactional data. This type of data captures the specific actions taken within the purchasing process and reflects the financial and business impact of them. The project's transactional data is visible in *Table 3-12* below.

Table 3-12. Transactional Data (Source: Dataedo, 2024)

Object	Description

PurchaseOrderHeader	General purchase order information such as the order's vendor, shipping method, and total cost due to vendor, etc.
PurchaseOrderDetail	Individual products associated with a specific purchase order, represented by the received quantity, stocked quantity, and per product subtotal, etc.
ProductInventory	Describe the storage quantity of each product.

3.3. Fact and Dimention tables

3.3.1. Data mapping

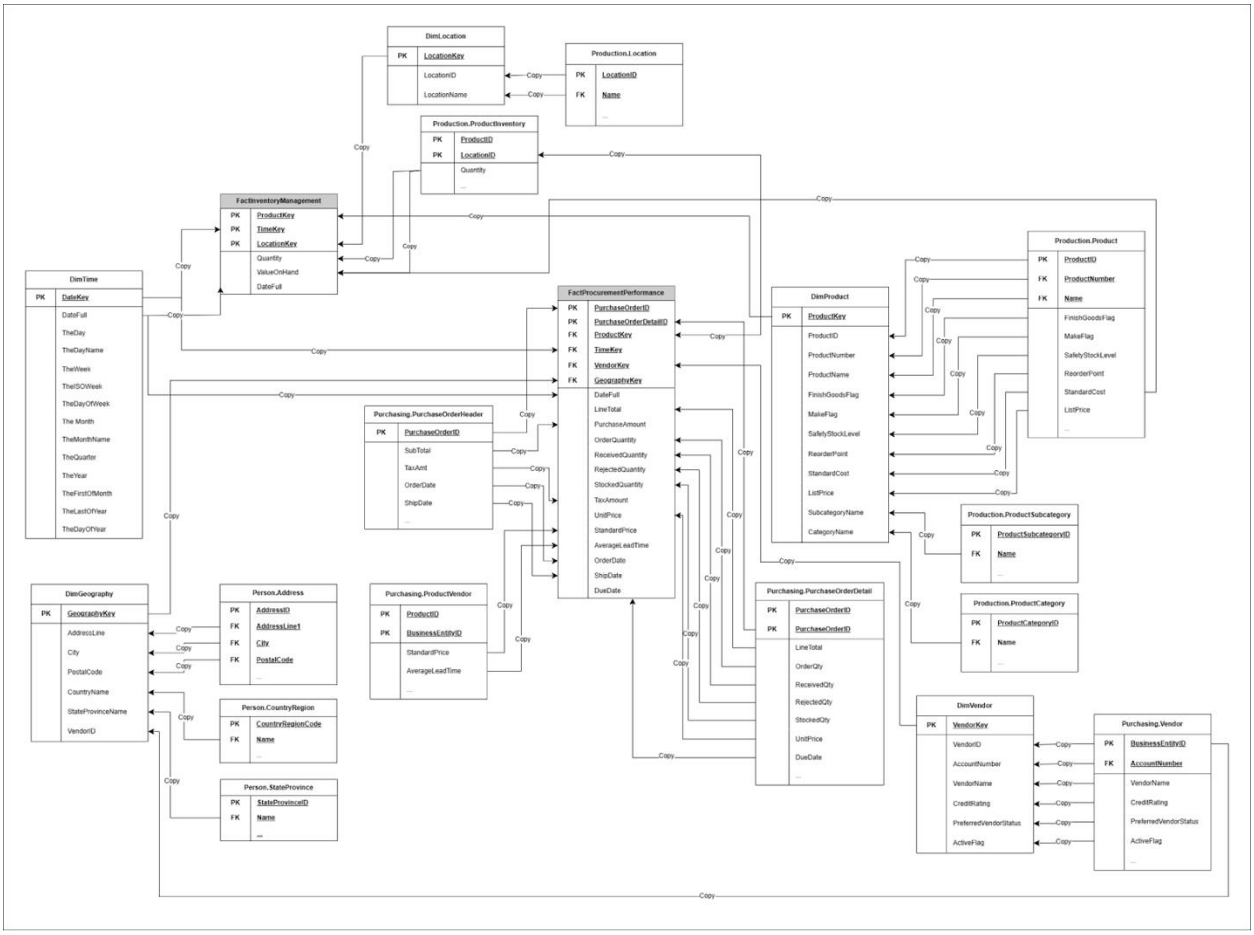


Figure 3.4. Data mapping (Source: Authors propose)

The presented table illustrates the process of data mapping and describes the transformation of data from its source into Fact and Dimension tables within the context of constructing a Purchasing data mart in this project. Data mapping stands as an important component in this project, serving purposes that are integral to the efficiency and integrity of the data warehouse architecture.

Primarily, data mapping ensures that the transition of data from various sources to the data warehouse is smooth and organized. It helps us see how different pieces of data relate to each other and where they should be placed in the warehouse.

Moreover, this step is important for maintaining the quality and consistency of the data. By carefully mapping out the data, we can identify any errors, duplications, or inconsistencies in the source data and correct them before they affect the integrity of our analyses.

Furthermore, data mapping helps align the technical aspects of the data warehouse with the goals and needs of the organization. By mapping the data to specific tables that reflect the relevant business processes and metrics, we ensure that the data warehouse is set up to support the analytical needs of the stakeholders.

In conclusion, data mapping is essential for organizing, cleaning, and aligning the data in a way that supports effective analysis and decision-making within an organization.

3.3.2. Dimensions tables

As analyzed in *Table 3-10. Bus Matrix*, there are five dimensions should be focused in case of this project's Business Intelligence solution as follows:

AdventureWorks' Purchasing module heavily relies on the DIMProduct dimension table, part of our robust BI solution utilizing the Adventureworks 2019 dataset. This table contains intricate qualitative details about the vast array of products offered by AdventureWorks. With attributes such as ProductID, ProductNumber, ProductName, and CategoryName, etc., we can intricately analyze and calculate essential key performance indicators (KPIs) for our Inventory Management dashboard. These indicators include inventory quantity, product location, COGS (Cost of goods sold), and Stock ReturnRate, etc., enabling Inventory Managers to obtain invaluable insights into the operations and

make informed decisions, consequently driving improvements in inventory management and overall profitability.

The smooth functioning of Adventure Work Cycles' procurement activities depends significantly on the DIMVendor and DIMGeography dimension tables. In the Procurement Performance dashboard, the DIMVendor table offers valuable quantitative insights into vendors' performance. Attributes like VendorName, CreditRating, and PreferredVendorStatus allow evaluating vendor's score, as well as assess procurement performance based on the associated DIMGeography table. Analyzing the vendor's geography, including attributes such as City, PostalCode, CountryName, and StateProvinceName, helps making informed decisions regarding vendor selection, facilitating improved procurement strategies and streamlined supplier management.

To incorporate the time element into this solution, the DIMTime dimension table enhances our ability to analyze historical data. By leveraging the Adventureworks 2019 dataset, this table enriches the Procurement Performance dashboard. The DateKey, set as the primary key, allows for easy time-based analysis. The attributes in this table, such as DateFull, TheDay, TheDayName, TheDayOfWeek, TheMonth, TheMonthName, TheQuarter, TheYear, TheFirstOfMonth, TheLastOfYear, and TheDayOfYear, enable us to evaluate procurement performance based on historical trends. This information empowers us to make data-driven decisions, compare procurement performance over time, and identify patterns and fluctuations to optimize procurement strategies.

Another necessary dimension is the DIMLocation table which complements our analysis within the Purchasing module by providing information about warehouse locations. This table is a primary component in our Inventory Management dashboard. The LocationKey, a surrogate key, uniquely identifies each warehouse location. The attribute, LocationName, represents the warehouse name. Leveraging this table, we can analyze and visualize warehouse-related information such as inventory quantity, product location, and other key indicators necessary for effective inventory management. This data empowers Inventory Managers and Directors to evaluate warehouse performance, make informed decisions, and optimize inventory storage and logistics strategies.

The integration of the DIMProduct, DIMVendor, DIMGeography, DIMTime, and DIMLocation dimension tables revolutionizes procurement-related decision-making. These dimension tables enable informed decision-making, enhanced inventory performance, and improved profitability. By leveraging qualitative and quantitative information, Adventure Work Cycles can optimize inventory management, streamline procurement activities, and achieve cost savings. The comprehensive dashboards, driven by these dimension tables, provide valuable insights into inventory performance, vendor evaluation, and future market trends, positioning the company for sustained growth.

To better see, the *Table 3-13.* is a data-detail-description version of the five dimension tables above:

Table 3-13. Data-detail Dimension Tables (Source: Authors propose)

Mapping Dimension tables					**Data sources (Data mapping)**	**Attribute mapping**
No.	**Name**	**Data type**	**Null**	**Description**		
DIMProduct - Qualitative Information of Product						
1	ProductKey	int		Surrogate key assigned to each product in the dimensional model. (PK)	SSIS Auto generate	
2	ProductID	int		Identifier for the product within the AdventureWorks database.	Production.Product	ProductID
3	ProductNumber	nvarchar(25)		Describe the natural characteristics of the product.	Production.Product	ProductNumber

4	ProductName	nvarchar(50)		Name of the product.	Production.Product	Name
5	FinishGoodsFlag	bit		0 = Product is not a salable item. 1 = Product is salable. Default: 1	Production.Product	FinishGoodsFlag
6	MakeFlag	bit		0 = Product is purchased, 1 = Product is manufactured in-house. Default: 1	Production.Product	MakeFlag
7	SafetyStockLevel	smallint		Minimum inventory quantity	Production.Product	SafetyStockLevel
8	ReorderPoint	smallint		Inventory level that triggers a purchase order or work order	Production.Product	ReorderPoint
9	StandardCost	money		Standard cost of the product	Production.Product	StandardCost
10	ListPrice	money		Selling price	Production.Product	ListPrice
11	SubcategoryName	nvarchar(50)		Subcategory description.	Production.Subcategory	Name
12	CategoryName	nvarchar(50)		Category name.	Production.Category	Name
DIMVendor - Quatitative Information of Vendor						

1	VendorKey	int		Surrogate key uniquely identifying each vendor. (PK)	SSIS auto-generated	
2	GeographyKey	int		Primary key of DIMGeography	DIMGeography	GeographyKey
3	AccountNumber	varchar(15)				
4	VendorID	int		Primary for vendor records	Purchasing.Vendor	BusinessEntityID
5	VendorName	nvarchar(50)		Vendor name	Purchasing.Vendor	VendorName
6	CreditRating	tinyint		1 = Superior, 2 = Excellent, 3 = Above average, 4 = Average, 5 = Below average	Purchasing.Vendor	CreditRating
7	PreferredVendorStatus	bit		0 = Do not use if another vendor is available. 1 = Preferred over other vendors supplying the same product. Default: 1	Purchasing.Vendor	PreferredVendorStatus
8	ActiveFlag	bit		0 = Vendor no longer used. 1 = Vendor is actively used. Default:	Purchasing.Vendor	ActiveFlag

				1		
DIMGeography - Quatitative Information about Geography of Vendor						
1	GeographyKey	int		Surrogate key uniquely identifying each geography. (PK)	SSIS auto-generated	
2	City	nvarchar(30)		Name of the city	Person.Address	City
3	PostalCode	nvarchar(15)		Postal code of the street address	Person.Address	PostalCode
4	CountryName	nvarchar(30)		Country or region name	Person.CountryRegion	Name
5	StateProvinceName	nvarchar(50)		State or Province description	Person.StateProvince	Name
6	AddressLine	nvarchar(60)		Address of Vendor	[Person].[Address]	[AddressLine1]
7	VendorID	int		Primary key for Person records	[Purchasing].[Vendor]	[BusinessEntityID]
DIMTime - Qualitative Information of Time						
1	DateKey	int		Primary key for DIMTime	SQL Server COMMAND	

2	DateFull	datetime	x	Detailed time information		
3	TheDay	int	x			
4	TheDayName	varchar(10)	x			
5	TheWeek					
6	TheISOWeek					
7	TheDayOfWeek	tinyint	x			
8	TheMonth					
9	TheMonthName					
10	TheQuarter					
11	TheYear	smallint	x			
12	TheFirstOfMonth					
13	TheLastOfYear					
14	TheDayOfYear					
DIMLocation - Qualitative Information of Warehouse location						
1	LocationKey	int		Surrogate key uniquely identifying	SSIS auto-generated	

				each location. (PK)		
2	LocationID	smallint		Primary key for location records	Production.Location	LocationID
3	LocationName	nvarchar(50)		Name of the location	Production.Location	Name

3.3.3. Fact tables

Similar to DIM tables, the two FACT tables mentioned in the bus matrix will also be clarified as below:

FACTInventoryManagement:

The purpose of the FACTInventoryManagement table is to capture and store data related to product inventory management within the AdventureWorks BI solution. This fact table serves as a central repository of information regarding the quantity and value of products in inventory, allowing for efficient analysis and reporting on inventory-related metrics.

The FACTInventoryManagement table is designed to support the calculation of key performance indicators (KPIs) that help evaluate and monitor the state of the inventory. By consolidating data from various sources, such as the DIMProduct, DIMTime, and DIMLocation dimension tables, the FACTInventoryManagement table enables the generation of insights into inventory levels, product locations, and inventory performance.

The table includes columns such as ProductKey, TimeKey, and LocationKey, which serve as foreign keys referencing the primary keys of their respective dimensions. These keys establish relationships between the FACTInventoryManagement table and the corresponding dimension tables, enabling multidimensional analysis of inventory data. Additionally, columns like Quantity and ValueOnHand provide specific information about the inventory's quantity and monetary value.

FACTProcurementPerformance:

The FACTProcurementPerformance table within this project serves as a main

repository for the systematic categorization and analysis of suitable data related to procurement activities and their subsequent performance. By capturing and consolidating data from DIMProduct, DIMTime, DIMVendor, and DIMGeography, it enables a thorough exploration and evaluation of procurement operations from critical perspectives.

The associative relationships established via foreign keys such as ProductKey, TimeKey, VendorKey, and GeographyKey connect the FACTProcurementPerformance table with the corresponding dimension tables, fostering comprehensive and insightful analyses. Central to the FACTProcurementPerformance table are critical attributes such as PurchaseOrderID, PurchaseOrderDetailID, PurchaseAmount, OrderQuantity, ReceivedQuantity, RejectQuantity, StockedQuantity, TaxAmount, UnitPrice, StandardPrice, AverageLeadTime, OrderDate, ShipDate, and DueDate. These attributes collectively contribute to a comprehensive understanding of purchasing patterns, financial aspects, and temporal considerations associated with procurement activities.

Through meticulous data examination and analysis derived from the FACTProcurementPerformance table, stakeholders can gain meaningful insights into procurement performance. Quantitative measures, such as the total monetary value of purchase orders, the quantity of items ordered, received, and rejected, along with pertinent details regarding vendors, provide critical information for performance assessment. Furthermore, the exploration of key temporal facets, including average lead time and adherence to prescribed delivery dates, further enhances the evaluation of procurement efficiency.

Ultimately, leveraging the insights garnered from the FACTProcurementPerformance table empowers stakeholders to make informed decisions, optimize vendor relationships, and refine procurement strategies. By understanding the underlying patterns and dynamics of procurement performance, organizations can fine-tune their operations, ensuring efficacy, and cultivating cost-effectiveness in the procurement domain.

Table 3-14 will provide detailed information on those two FACT tables.

Table 3-14. Fact tables (Source: Authors propose)

Mapping Fact tables					Data sources (Data Mapping)	Attribute mapping
No.	**Name**	**Data type**	**Null**	**Description**		
FACTInventoryManagement						
1	ProductKey	int		DIMProduct primary key	DIMProduct	ProductKey
2	TimeKey	int		DIMTime primary key	DIMTime	TimeKey
3	LocationKey	int		DIMLocation primary key	DIMLocation	LocationKey
4	Quantity	smallint		Product inventory	Production.ProductInventory	Quantity
5	ValueOnHand	int		Quantity * Standard Cost		
FACTProcurementPerformance						
1	ProductKey	int		DIMProduct primary key	DIMProduct	ProductKey
2	TimeKey	int		DIMTime primary key	DIMTime	TimeKey
3	VendorKey	int		DIMVendor primary key	DIMVendor	VendorKey

4	GeographyKey	int		DIMGeography primary key	DIMGeography	GeographyKey
5	PurchaseOrderID	INT			Purchasing.PurchaseOrderHeader	PurchaseOrderID
6	PurchaseOrderDetailID	INT			Purchasing.PurchaseOrderDetail	PurchaseOrderDetailID
7	PurchaseAmount	Money			Purchasing.PurchaseOrderHeader	LineTotal
8	OrderQuantity	smallint			Purchasing.PurchaseOrderDetail	OrderQty
9	ReceivedQuantity	decimal (8,2)			Purchasing.PurchaseOrderDetail	ReceivedQty
10	RejectQuantity	decimal (8,2)			Purchasing.PurchaseOrderDetail	RejectedQty
11	StockedQuantity	decimal (9,2)			Purchasing.PurchaseOrderDetail	StockedQty
12	TaxAmount	Money			Purchasing.PurchaseOrderHeader	TaxAmt
13	UnitPrice	Money			Purchasing.PurchaseOrderDetail	UnitPrice
14	StandardPrice	Money			Purchasing.ProductVendor	StandardPrice
15	AverageLeadTime	Int			Purchasing.Product	AverageLeadTi

					Vendor	me
16	OrderDate	Datetime			Purchasing.PurchaseOrderHeader	OrderDate
17	ShipDate	Datetime			Purchasing.PurchaseOrderHeader	ShipDate
18	DueDate	Datetime			Purchasing.PurchaseOrderDetail	DueDate

3.4. Data warehouse model

After having detailed information about the DIM and FACT tables and their relationships, the next step is to build the data warehouse model. Because of the relationship between the DIM and FACT tables, it is determined that the Data warehouse model is in Galaxy Schema type.

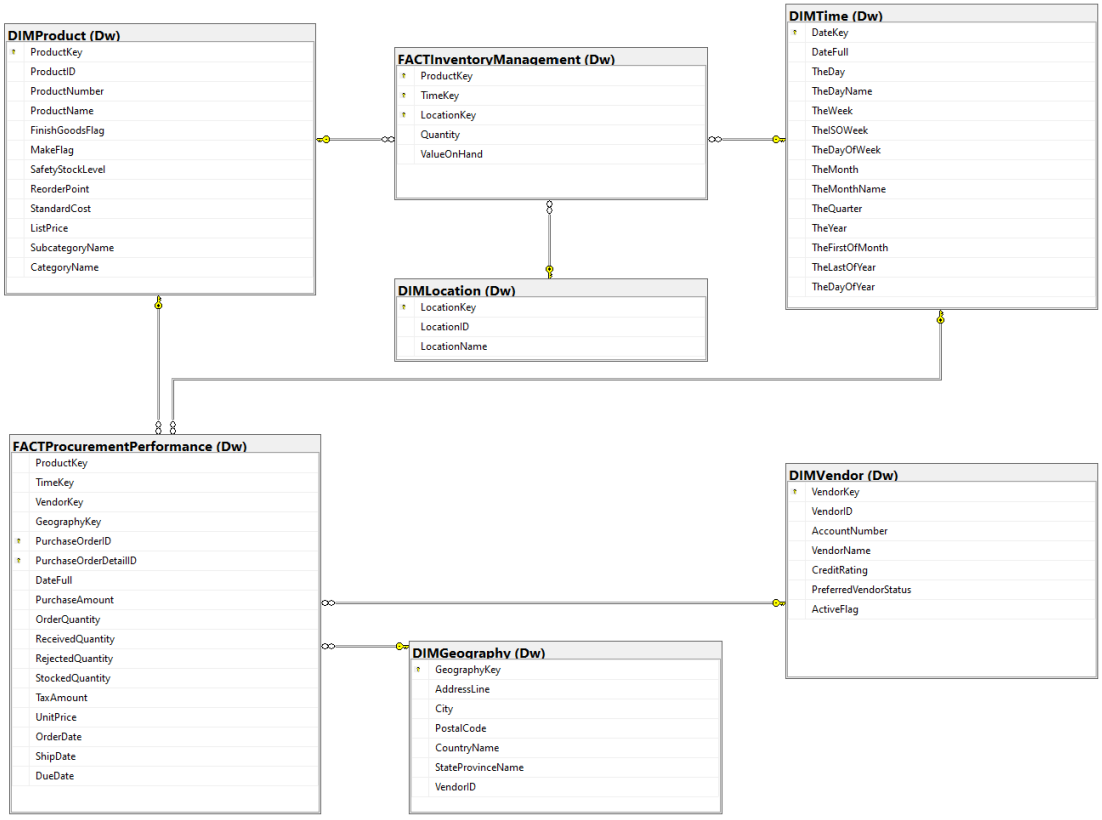


Figure 3.5. The galaxy schema (Source: Authors propose)

Table 3-15. DIM and FACT tables relationships (Source: Authors propose)

No	Table	Relationship	Table	Description
1	DIMProduct	one-many	FACTInventoryMangement	A product can have one or many inventory management records but an inventory management record belongs to only one product.
2	DIMLocation	one-many	FACTInventoryMangement	A location can have one or many inventory management records but an inventory management record belongs to only one

				location.
3	DIMTime	one-many	FACTInventoryMangement	A time (date) can have one or many inventory management records but an inventory management record belongs to only one time (date).
4	DIMProduct	one-many	FACTProcurementPerformance	A product can have one or many purchasing price analysis records but a purchasing price analysis record belongs to only one product.
5	DIMTime	one-many	FACTProcurementPerformance	A time (date) can have one or many purchasing price analysis records but a purchasing price analysis record belongs to only one time (date).
6	DIMVendor	one-many	FACTProcurementPerformance	A vendor can have one or many purchasing price analysis records but a purchasing price analysis record belongs to only one vendor.
7	DIMGeography	one-many	FACTProcurementPerformance	A geography can have one or many purchasing price analysis records but a purchasing price analysis record belongs to only one geography.

The table provides a concise overview of the one-to-many relationships between various dimension and fact tables, highlighting the cardinality and associations within the data model.

Chapter 4: Data Integration

This chapter focuses on data integration, a crucial phase in the extract, transform, load (ETL) process. In this chapter, we delve into the seamless merging of data from various sources into a unified format suitable for analysis and reporting.

4.1. ETL Process

Step 1: Start by creating a new SSIS Project in Visual Studio.

Step 2: Add and set up an OLE DB connection manager to establish a connection with the destination database.

4.1.1. Dimension table's ETL Process

4.1.1.1. ETL Process of DIMProduct

Step 1: Create DimProduct Package: Create a new SSIS package named DimProduct. In the Control Flow, drag a new Data flow task and name it DIMProduct.

Figure 4.1. Create Data Flow Task for DIMProduct (Source: Authors propose)

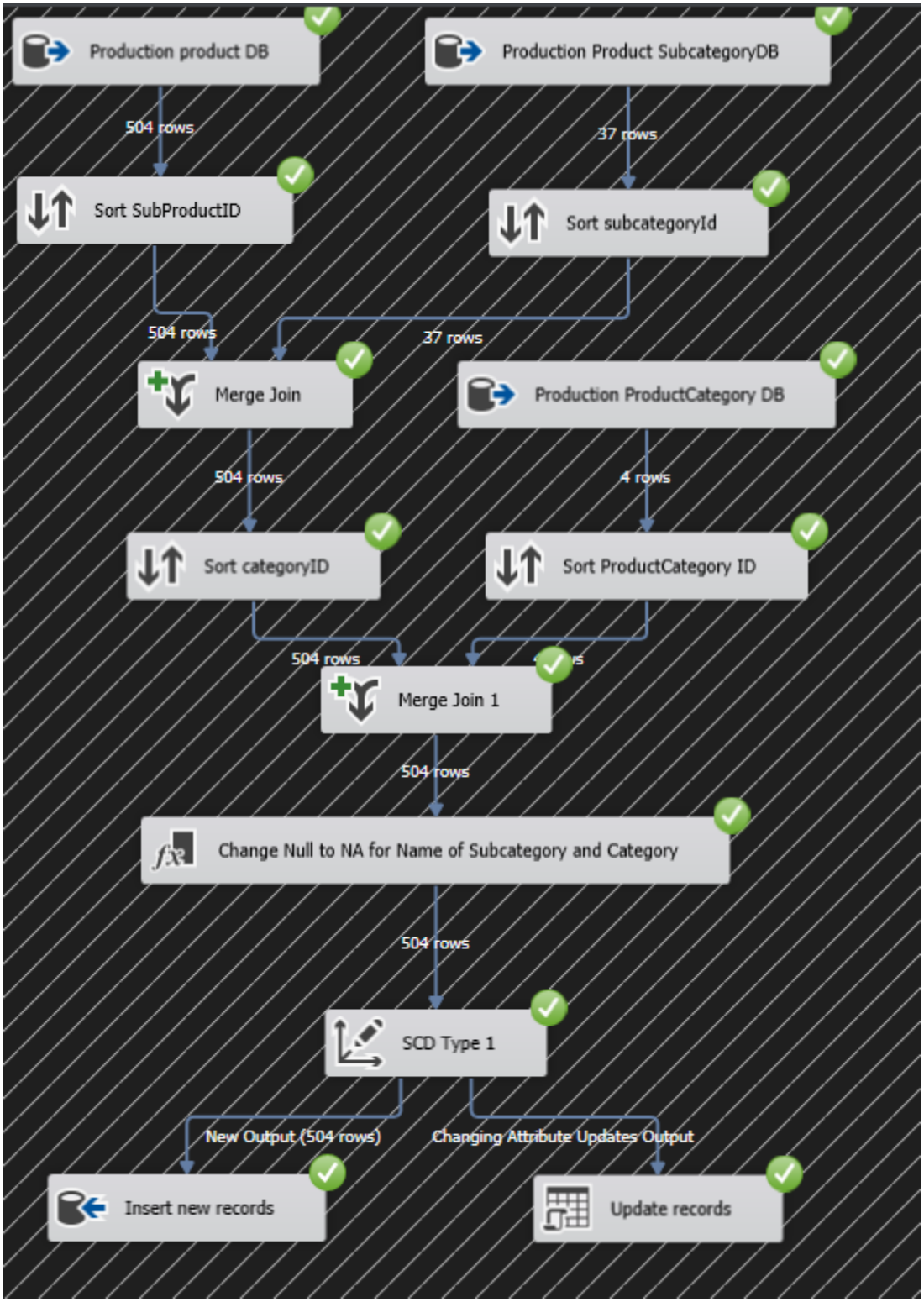


Figure 4.2. ETL Process of DIMProduct (Source: Authors propose)

Step 2: Data Flow Task - Integrate and Configure: Within the Data Flow Task, integrate data from the Production.Product, Production.SubProductCategory, and

Production.ProductCategory in AW2022 databases. Configure appropriate data sources and destinations to extract data from these databases and load it into the DimProduct table.

Step 3: Add Sort transformations after each OLE DB Source component to sort the data based on the respective columns.

- For the Product DB, sort the data based on ProductSubcategoryID.

- For the SubCategory DB, sort the data based on ProductSubcategoryID.

- For the Category DB, sort the data based on ProductCategoryID.

Step 4: Merge Join SubCategory: Connect the sorted outputs of both OLE DB Source components to the Merge Join transformation. Configure the Merge Join to join and sort the data based on the ProductSubCategoryID.

Step 4: Merge Join SubCategory: Connect the sorted outputs of the ProductSubcategoryID and ProductCategoryID to the Merge Join transformation. Configure the Merge Join to join the data based on the ProductCategoryID.

Step 5: Handle NULL Values: Utilize a Derived Column Transformation to convert NULL values in the subcategory and category columns to 'NA' before loading them into the DimProduct table.

Step 6: Slowly Changing Dimension (SCD): Implement Slowly Changing Dimension components to handle changes in dimension data. Identify the appropriate type of SCD Type 1 based on the requirements and nature of changes in the data. Configure the SCD component to update existing records and handle new records or changes appropriately in the DimProduct table.

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

Results Messages

	ProductKey	ProductID	ProductNumber	ProductName	FinishGoodsFlag	MakeFlag	SafetyStockLevel	ReorderPoint	StandardCost	ListPrice	SubcategoryName	CategoryName
1	1	1	AR-5381	Adjustable Race	0	0	1000	750	0.00	0.00	NA	NA
2	2	323	CR-9981	Crown Race	0	0	1000	750	0.00	0.00	NA	NA
3	3	322	CR-7833	Chainring	0	0	1000	750	0.00	0.00	NA	NA
4	4	2	BA-8327	Bearing Ball	0	0	1000	750	0.00	0.00	NA	NA
5	5	3	BE-2349	BB Ball Bearing	0	1	800	600	0.00	0.00	NA	NA
6	6	4	BE-2908	Headset Ball Bearings	0	0	800	600	0.00	0.00	NA	NA
7	7	316	BL-2036	Blade	0	1	800	600	0.00	0.00	NA	NA
8	8	317	CA-5965	LL Crankarm	0	0	500	375	0.00	0.00	NA	NA
9	9	318	CA-6738	ML Crankarm	0	0	500	375	0.00	0.00	NA	NA
10	10	319	CA-7457	HL Crankarm	0	0	500	375	0.00	0.00	NA	NA
11	11	324	CS-2812	Chain Stays	0	1	1000	750	0.00	0.00	NA	NA
12	12	320	CB-2903	Chainring Bolts	0	0	1000	750	0.00	0.00	NA	NA
13	13	325	DC-8732	Decal 1	0	0	1000	750	0.00	0.00	NA	NA
14	14	342	FW-1200	Flat Washer 6	0	0	1000	750	0.00	0.00	NA	NA
15	15	330	EC-T209	Touring End Caps	0	1	1000	750	0.00	0.00	NA	NA
16	16	326	DC-9824	Decal 2	0	0	1000	750	0.00	0.00	NA	NA
17	17	327	DT-2377	Down Tube	0	1	800	600	0.00	0.00	NA	NA
18	18	328	EC-M092	Mountain End Caps	0	1	1000	750	0.00	0.00	NA	NA
19	19	329	EC-R098	Road End Caps	0	1	1000	750	0.00	0.00	NA	NA
20	20	331	FE-3760	Fork End	0	1	800	600	0.00	0.00	NA	NA
21	21	372	HJ-7162	Thin-Jam Hex Nut 8	0	0	1000	750	0.00	0.00	NA	NA
22	22	332	FH-2981	Freewheel	0	0	500	375	0.00	0.00	NA	NA
23	23	341	FW-1000	Flat Washer 1	0	0	1000	750	0.00	0.00	NA	NA
24	24	343	FW-1400	Flat Washer 2	0	0	1000	750	0.00	0.00	NA	NA
25	25	373	HJ-9080	Thin-Jam Hex Nut 12	0	0	1000	750	0.00	0.00	NA	NA
26	26	321	CN-6137	Chainring Nut	0	0	1000	750	0.00	0.00	NA	NA
27	27	345	FW-3800	Flat Washer 4	0	0	1000	750	0.00	0.00	NA	NA
28	28	359	HJ-1213	Thin-Jam Hex Nut 9	0	0	1000	750	0.00	0.00	NA	NA
29	29	348	FW-7160	Flat Washer 5	0	0	1000	750	0.00	0.00	NA	NA
30	30	349	FW-9160	Flat Washer 7	0	0	1000	750	0.00	0.00	NA	NA
31	31	350	FC-3654	Fork Crown	0	1	800	600	0.00	0.00	NA	NA
32	32	351	FC-3982	Front Derailleur Cage	0	0	800	600	0.00	0.00	NA	NA
33	33	352	FL-2301	Front Derailleur Link...	0	0	800	600	0.00	0.00	NA	NA
34	34	355	GP-0982	Guide Pulley	0	0	800	600	0.00	0.00	NA	NA
35	35	356	GT-0820	LL Grip Tape	0	0	800	600	0.00	0.00	NA	NA
36	36	357	GT-1209	ML Grip Tape	0	0	800	600	0.00	0.00	NA	NA
37	37	347	FW-5800	Flat Washer 8	0	0	1000	750	0.00	0.00	NA	NA
38	38	426	LJ-3410	Thin-Jam Lock Nut 15	0	0	1000	750	0.00	0.00	NA	NA
39	39	360	HJ-1220	Thin-Jam Hex Nut 10	0	0	1000	750	0.00	0.00	NA	NA
40	40	371	HJ-7161	Thin-Jam Hex Nut 7	0	0	1000	750	0.00	0.00	NA	NA

Query executed successfully. DESKTOP-PTHVTES\MSSQLSERVER... DESKTOP-PTHVTES\duydu ... DW_PURCHASING 00:00:00 504 rows

Figure 4.3. SQL Query result of ETL DIMProduct's process (Source: Authors propose)

4.1.1.2. SQL Query result of DIMTime

- Drop Existing Table: Checks if the table DimTime already exists in the dw schema and drops it if it does.

- Define Variables: Sets up variables including @StartDate (initial date), @Year (number of years to generate data for), and calculates @CutoffDate as the end date for data generation.

- Generate Date Sequence: Utilizes two common table expressions (seq and d) to generate a sequence of dates from @StartDate to @CutoffDate.

- Extract Date Attributes: Extracts various date attributes such as day, day name, week, month, quarter, year, etc., from each date in the sequence and stores them in the src common table expression.

- Insert Data into DimTime: Inserts the data from the src common table expression into the DimTime table in the dw schema. It orders the records by the DateKey.

- Recursion Control: Uses OPTION (MAXRECURSION 0) to ensure that the recursive common table expression can generate dates for the entire specified period without hitting the default recursion limit.

Results Messages

	DateKey	DateFull	TheDay	TheDayName	TheWeek	TheISOWeek	TheDayOfWeek	TheMonth	TheMonthName	TheQuarter	TheYear	TheFirstOfMonth	TheLastOfYear	TheDayOfYear
1	20100101	2010-01-01 00:00:00.000	1	Friday	1	53	6	1	January	1	2010	2010-01-01	2010-12-31	1
2	20100102	2010-01-02 00:00:00.000	2	Saturday	1	53	7	1	January	1	2010	2010-01-01	2010-12-31	2
3	20100103	2010-01-03 00:00:00.000	3	Sunday	2	53	1	1	January	1	2010	2010-01-01	2010-12-31	3
4	20100104	2010-01-04 00:00:00.000	4	Monday	2	1	2	1	January	1	2010	2010-01-01	2010-12-31	4
5	20100105	2010-01-05 00:00:00.000	5	Tuesday	2	1	3	1	January	1	2010	2010-01-01	2010-12-31	5
6	20100106	2010-01-06 00:00:00.000	6	Wednesday	2	1	4	1	January	1	2010	2010-01-01	2010-12-31	6
7	20100107	2010-01-07 00:00:00.000	7	Thursday	2	1	5	1	January	1	2010	2010-01-01	2010-12-31	7
8	20100108	2010-01-08 00:00:00.000	8	Friday	2	1	6	1	January	1	2010	2010-01-01	2010-12-31	8
9	20100109	2010-01-09 00:00:00.000	9	Saturday	2	1	7	1	January	1	2010	2010-01-01	2010-12-31	9
10	20100110	2010-01-10 00:00:00.000	10	Sunday	3	1	1	1	January	1	2010	2010-01-01	2010-12-31	10
11	20100111	2010-01-11 00:00:00.000	11	Monday	3	2	2	1	January	1	2010	2010-01-01	2010-12-31	11
12	20100112	2010-01-12 00:00:00.000	12	Tuesday	3	2	3	1	January	1	2010	2010-01-01	2010-12-31	12
13	20100113	2010-01-13 00:00:00.000	13	Wednesday	3	2	4	1	January	1	2010	2010-01-01	2010-12-31	13
14	20100114	2010-01-14 00:00:00.000	14	Thursday	3	2	5	1	January	1	2010	2010-01-01	2010-12-31	14
15	20100115	2010-01-15 00:00:00.000	15	Friday	3	2	6	1	January	1	2010	2010-01-01	2010-12-31	15
16	20100116	2010-01-16 00:00:00.000	16	Saturday	3	2	7	1	January	1	2010	2010-01-01	2010-12-31	16
17	20100117	2010-01-17 00:00:00.000	17	Sunday	4	2	1	1	January	1	2010	2010-01-01	2010-12-31	17
18	20100118	2010-01-18 00:00:00.000	18	Monday	4	3	2	1	January	1	2010	2010-01-01	2010-12-31	18
19	20100119	2010-01-19 00:00:00.000	19	Tuesday	4	3	3	1	January	1	2010	2010-01-01	2010-12-31	19
20	20100120	2010-01-20 00:00:00.000	20	Wednesday	4	3	4	1	January	1	2010	2010-01-01	2010-12-31	20
21	20100121	2010-01-21 00:00:00.000	21	Thursday	4	3	5	1	January	1	2010	2010-01-01	2010-12-31	21
22	20100122	2010-01-22 00:00:00.000	22	Friday	4	3	6	1	January	1	2010	2010-01-01	2010-12-31	22
23	20100123	2010-01-23 00:00:00.000	23	Saturday	4	3	7	1	January	1	2010	2010-01-01	2010-12-31	23
24	20100124	2010-01-24 00:00:00.000	24	Sunday	5	3	1	1	January	1	2010	2010-01-01	2010-12-31	24
25	20100125	2010-01-25 00:00:00.000	25	Monday	5	4	2	1	January	1	2010	2010-01-01	2010-12-31	25
26	20100126	2010-01-26 00:00:00.000	26	Tuesday	5	4	3	1	January	1	2010	2010-01-01	2010-12-31	26
27	20100127	2010-01-27 00:00:00.000	27	Wednesday	5	4	4	1	January	1	2010	2010-01-01	2010-12-31	27
28	20100128	2010-01-28 00:00:00.000	28	Thursday	5	4	5	1	January	1	2010	2010-01-01	2010-12-31	28
29	20100129	2010-01-29 00:00:00.000	29	Friday	5	4	6	1	January	1	2010	2010-01-01	2010-12-31	29
30	20100130	2010-01-30 00:00:00.000	30	Saturday	5	4	7	1	January	1	2010	2010-01-01	2010-12-31	30
31	20100131	2010-01-31 00:00:00.000	31	Sunday	6	4	1	1	January	1	2010	2010-01-01	2010-12-31	31
32	20100201	2010-02-01 00:00:00.000	1	Monday	6	5	2	2	February	1	2010	2010-02-01	2010-12-31	32
33	20100202	2010-02-02 00:00:00.000	2	Tuesday	6	5	3	2	February	1	2010	2010-02-01	2010-12-31	33
34	20100203	2010-02-03 00:00:00.000	3	Wednesday	6	5	4	2	February	1	2010	2010-02-01	2010-12-31	34
35	20100204	2010-02-04 00:00:00.000	4	Thursday	6	5	5	2	February	1	2010	2010-02-01	2010-12-31	35
36	20100205	2010-02-05 00:00:00.000	5	Friday	6	5	6	2	February	1	2010	2010-02-01	2010-12-31	36
37	20100206	2010-02-06 00:00:00.000	6	Saturday	6	5	7	2	February	1	2010	2010-02-01	2010-12-31	37
38	20100207	2010-02-07 00:00:00.000	7	Sunday	7	5	1	2	February	1	2010	2010-02-01	2010-12-31	38
39	20100208	2010-02-08 00:00:00.000	8	Monday	7	6	2	2	February	1	2010	2010-02-01	2010-12-31	39
40	20100209	2010-02-09 00:00:00.000	9	Tuesday	7	6	3	2	February	1	2010	2010-02-01	2010-12-31	40

Query executed successfully. DESKTOP-PTHVTES\MSSQLSERVER... DESKTOP-PTHVTES\duydu ... DW_PURCHASING 00:00:00 1,826 rows

Figure 4.4. SQL Query result of ETL DIMTime's process (Source: Authors propose)

4.1.1.3. ETL Process of DIMLocation

Step 1: Create DimLocation Package: Create a new SSIS package named DimLocation. In the Control Flow, drag a Data flow task and name it DIMLocation.

Figure 4.5. Create Data Flow Task for DIMLocation (Source: Authors propose)

Step 2: Data Flow Task - Integrate and Configure: Within the Data Flow Task, integrate data from the Production.Location from AW2022 databases. Configure appropriate data sources and destinations to extract data from these databases and load it into the DimLocationt table.

Step 3: Slowly Changing Dimension (SCD): Implement Slowly Changing Dimension components to handle changes in dimension data. Identify the appropriate type of SCD Type 1 based on the requirements and nature of changes in the data. Configure the SCD component to update existing records and handle new records or changes appropriately in the DimLocationtable.

Step 4: Include and configure an OLE DB Destination to insert the transformed data into a table in the destination database.

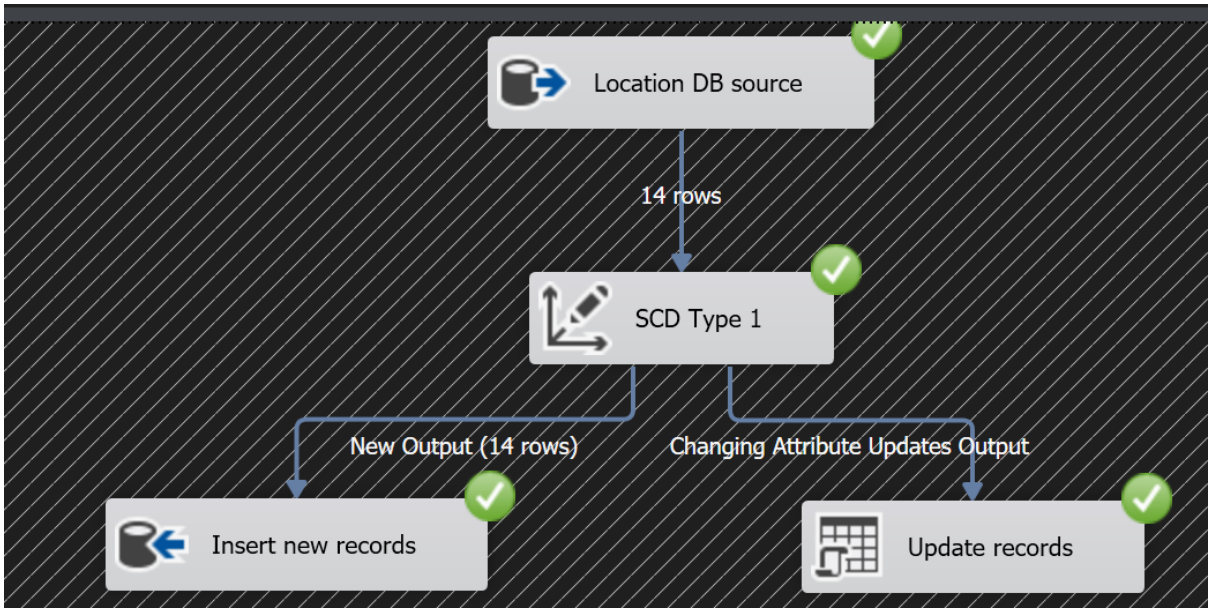


Figure 4.6. ETL Process of DIMLocation (Source: Authors propose)

Results Messages

	LocationKey	LocationID	LocationName
1	1	1	Tool Crib
2	2	2	Sheet Metal Racks
3	3	3	Paint Shop
4	4	4	Paint Storage
5	5	5	Metal Storage
6	6	6	Miscellaneous Storage
7	7	7	Finished Goods Storage
8	8	10	Frame Forming
9	9	20	Frame Welding
10	10	30	Debur and Polish
11	11	40	Paint
12	12	45	Specialized Paint
13	13	50	Subassembly
14	14	60	Final Assembly

Query executed successfully. DESKTOP-PTHVTES\MSSQLSERVER... DESKTOP-PTHVTES\duydu ... DW_PURCHASING 00:00:00 14 rows

Figure 4.7. SQL Query result of ETL DIMLocation process (Source: Authors propose)

4.1.1.4. ETL Process of DIMGeography

Step 1: Create DIMGeography Package: Create a new SSIS package named DimProduct. In the Control Flow, drag a Data flow task and name it DIMGeography.

Figure 4.8. Create Data Flow Task for DIMGeography (Source: Authors propose)

Step 2: Data Flow Task - Integrate and Configure: Within the Data Flow Task, integrate data from the view Purchasing.VendorWithAddress from AW2022 databases. Configure appropriate data sources and destinations to extract data from these databases and load it into the DimGeography table.

Step 3: Slowly Changing Dimension (SCD): Implement Slowly Changing Dimension components to handle changes in dimension data. Identify the appropriate type of SCD Type 1 based on the requirements and nature of changes in the data. Configure the SCD component to update existing records and handle new records or changes appropriately in the DimLocationtable.

Step 4: Include and configure an OLE DB Destination to insert the transformed data into a table in the destination database.

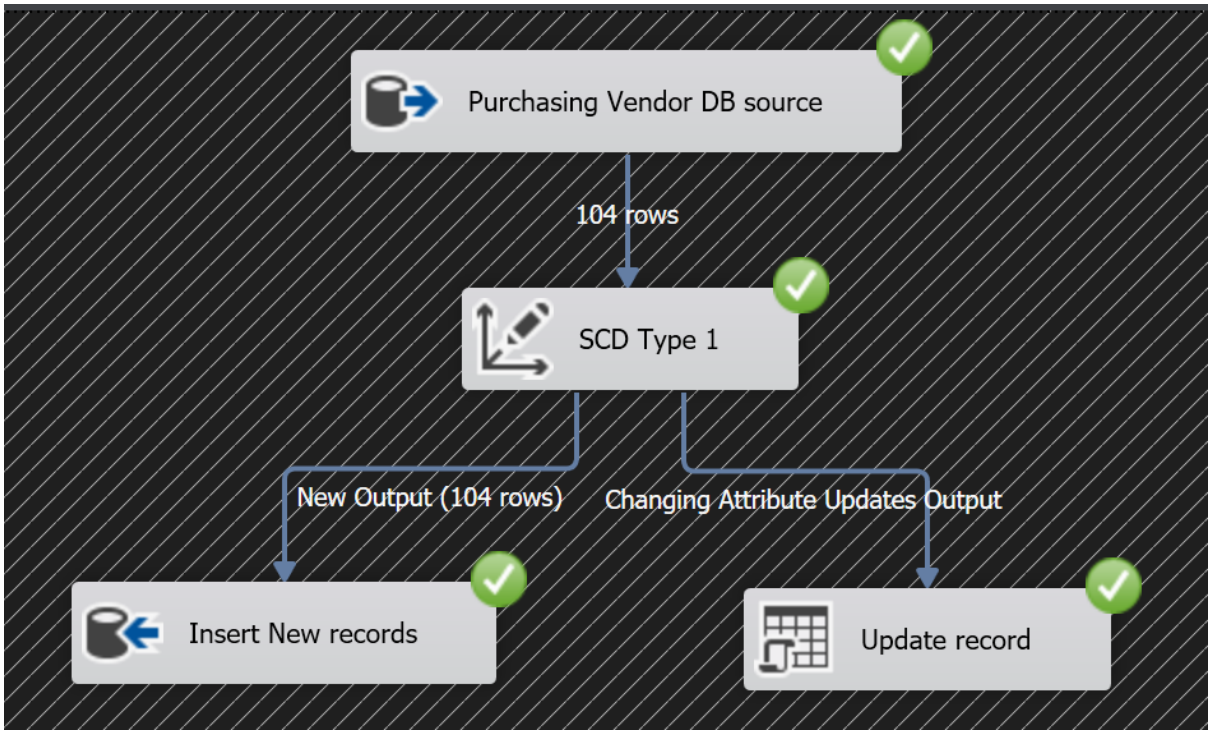


Figure 4.9. ETL Process of DIMGeography (Source: Authors propose)

Results Messages

	GeographyKey	AddressLine	City	PostalCode	CountryName	StateProvinceName	VendorID
1	1	28 San Marino Ct.	Bellingham	98225	United States	Washington	1492
2	2	4659 Montoya	Altadena	91001	United States	California	1494
3	3	7995 Edwards Ave.	Lynnwood	98036	United States	Washington	1496
4	4	90 Sunny Ave	Berkeley	94704	United States	California	1498
5	5	9098 Story Lane	Albany	12210	United States	New York	1500
6	6	4823 Stonewood Ct.	Walla Walla	99362	United States	Washington	1502
7	7	15 Pear Dr.	Newport Beach	92625	United States	California	1504
8	8	6441 Co Road	Lemon Grove	85252	United States	Arizona	1506
9	9	50 Via Del Sol	Lynnwood	98036	United States	Washington	1508
10	10	683 Larch Ct.	Salt Lake City	84101	United States	Utah	1510
11	11	298 Sunnybrook Drive	Spring Valley	91977	United States	California	1512
12	12	6 Dancing Road	Burien	98168	United States	Washington	1514
13	13	8513 Hurlstone Ct.	Altadena	91001	United States	California	1516
14	14	8844 Garcia	West Covina	91791	United States	California	1518
15	15	8981 Carmel Drive	W. Linn	89701	United States	Nevada	1520
16	16	8127 Otter Dr.	Boise	83702	United States	Idaho	1522
17	17	207 Concerto Circle	Salem	97301	United States	Oregon	1524
18	18	20 Rambling Rose Ave.	West Covina	91791	United States	California	1526
19	19	3195 RiverRock Dr.	Burlingame	94010	United States	California	1528
20	20	7651 Smiling Tree Court	Los Angeles	90012	United States	California	1530
21	21	3253 La Jolla	Salem	97301	United States	Oregon	1532
22	22	35 Buckthorn Court	Kirkland	98033	United States	Washington	1534
23	23	67 Monetary Way	Berkeley	94704	United States	California	1536
24	24	8197 Hermosa	Salem	97301	United States	Oregon	1538
25	25	9830 May Way	Mill Valley	59715	United States	Montana	1540
26	26	3664 Colt Ct.	Richmond	94801	United States	California	1542
27	27	1 Mt. Dell Drive	Portland	97205	United States	Oregon	1544
28	28	2342 Peachwillow	Denver	80203	United States	Colorado	1546
29	29	3 Chablis Court	Torrance	90505	United States	California	1548
30	30	7 Mayda Way	Bremerton	98312	United States	Washington	1550
31	31	2014 Delta Road	Burien	98168	United States	Washington	1552
32	32	8040 Erie Dr	Houston	77003	United States	Texas	1554
33	33	7786 Olive St	Olympia	98501	United States	Washington	1556
34	34	7 B Way	Woodburn	97071	United States	Oregon	1558
35	35	628 Muir Road	Los Angeles	90012	United States	California	1560
36	36	68 Filling Ave.	Beaverton	97005	United States	Oregon	1562
37	37	1192 Parkway Drive	Mill Valley	68601	United States	Nebraska	1564
38	38	5807 Churchill Dr.	Corvallis	97330	United States	Oregon	1566
39	39	419 River Ash Court	Lakewood	63301	United States	Missouri	1568
40	40	7053 Laurel Dr.	La Mesa	88044	United States	New Mexico	1570

Query executed successfully. DESKTOP-PTHVTES\MSSQLSERVER... DESKTOP-PTHVTES\duydu ... DW_PURCHASING 00:00:00 104 rows

Figure 4.10. SQL Query result of ETL DIMGeography process (Source: Authors propose)

4.1.1.5. ETL Process of DIMVendor

Step 1: Create DIMVendor Package: Create a new SSIS package named DIMVendor.

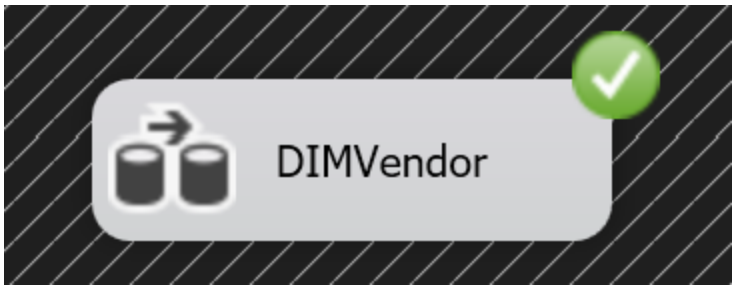


Figure 4.11. Create Data Flow Task for DIMVendor (Source: Authors propose)

Step 2: Data Flow Task - Integrate and Configure: Within the Data Flow Task, integrate data from the table Purchasing.Vendor from the AW2022 database. Configure appropriate data sources and destinations to extract data from these databases and load it into the DIMVendor table.

Step 3: Slowly Changing Dimension (SCD): Implement Slowly Changing Dimension components to handle changes in dimension data. Identify the appropriate type of SCD Type 1 based on the requirements and nature of changes in the data. Configure the SCD component to update existing records and handle new records or changes appropriately in the DimVendor table.

Step 4: Include and configure an OLE DB Destination to insert the transformed data into a table in the destination database.

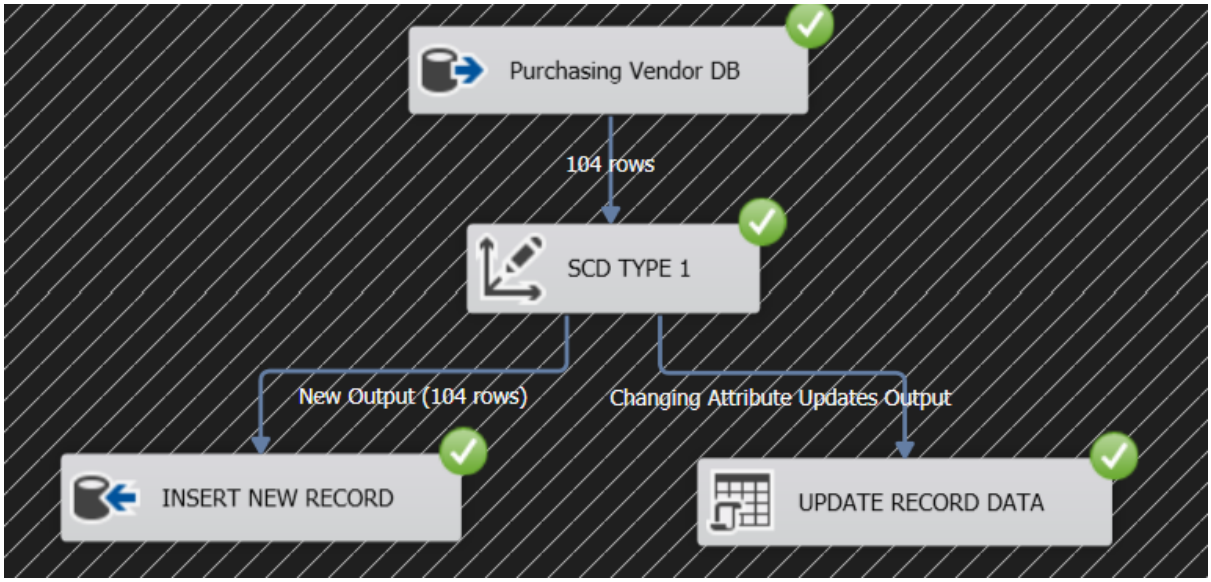
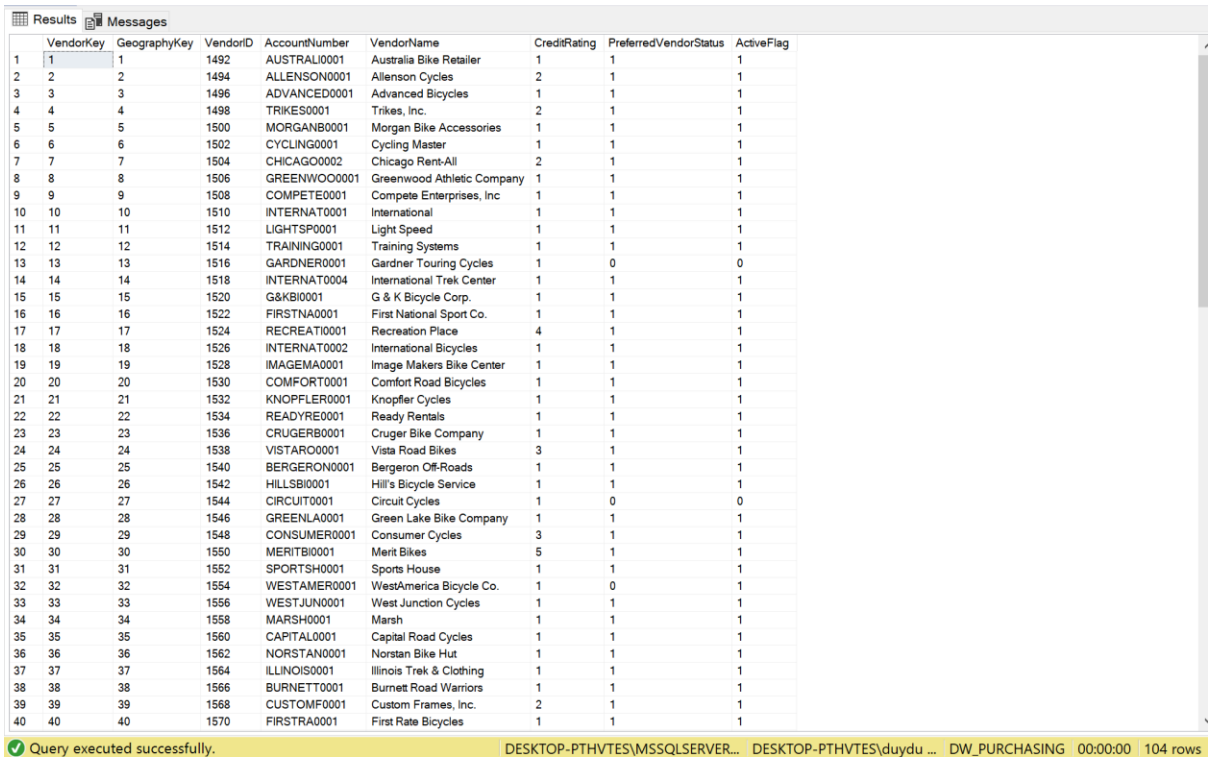


Figure 4.12. ETL Process of DIMVendor (Source: Authors propose)



Results Messages

	VendorKey	GeographyKey	VendorID	AccountNumber	VendorName	CreditRating	PreferredVendorStatus	ActiveFlag
1	1	1	1492	AUSTRALI0001	Australia Bike Retailer	1	1	1
2	2	2	1494	ALLENSON0001	Allenson Cycles	2	1	1
3	3	3	1496	ADVANCED0001	Advanced Bicycles	1	1	1
4	4	4	1498	TRIKES0001	Trikes, Inc.	2	1	1
5	5	5	1500	MORGANB0001	Morgan Bike Accessories	1	1	1
6	6	6	1502	CYCLING0001	Cycling Master	1	1	1
7	7	7	1504	CHICAGO0002	Chicago Rent-All	2	1	1
8	8	8	1506	GREENWOO0001	Greenwood Athletic Company	1	1	1
9	9	9	1508	COMPETE0001	Compete Enterprises, Inc	1	1	1
10	10	10	1510	INTERNAT0001	International	1	1	1
11	11	11	1512	LIGHTSP0001	Light Speed	1	1	1
12	12	12	1514	TRAINING0001	Training Systems	1	1	1
13	13	13	1516	GARDNER0001	Gardner Touring Cycles	1	0	0
14	14	14	1518	INTERNAT0004	International Trek Center	1	1	1
15	15	15	1520	G&KBI0001	G & K Bicycle Corp.	1	1	1
16	16	16	1522	FIRSTNA0001	First National Sport Co.	1	1	1
17	17	17	1524	RECREATI0001	Recreation Place	4	1	1
18	18	18	1526	INTERNAT0002	International Bicycles	1	1	1
19	19	19	1528	IMAGEMA0001	Image Makers Bike Center	1	1	1
20	20	20	1530	COMFORT0001	Comfort Road Bicycles	1	1	1
21	21	21	1532	KNOPFLER0001	Knopfler Cycles	1	1	1
22	22	22	1534	READYRE0001	Ready Rentals	1	1	1
23	23	23	1536	CRUGERB0001	Cruger Bike Company	1	1	1
24	24	24	1538	VISTARO0001	Vista Road Bikes	3	1	1
25	25	25	1540	BERGERON0001	Bergeron Off-Roads	1	1	1
26	26	26	1542	HILLSBI0001	Hill's Bicycle Service	1	1	1
27	27	27	1544	CIRCUIT0001	Circuit Cycles	1	0	0
28	28	28	1546	GREENLA0001	Green Lake Bike Company	1	1	1
29	29	29	1548	CONSUMER0001	Consumer Cycles	3	1	1
30	30	30	1550	MERITBI0001	Merit Bikes	5	1	1
31	31	31	1552	SPORTSH0001	Sports House	1	1	1
32	32	32	1554	WESTAMER0001	WestAmerica Bicycle Co.	1	0	1
33	33	33	1556	WESTJUN0001	West Junction Cycles	1	1	1
34	34	34	1558	MARSH0001	Marsh	1	1	1
35	35	35	1560	CAPITAL0001	Capital Road Cycles	1	1	1
36	36	36	1562	NORSTAN0001	Norstan Bike Hut	1	1	1
37	37	37	1564	ILLINOIS0001	Illinois Trek & Clothing	1	1	1
38	38	38	1566	BURNETT0001	Burnett Road Warriors	1	1	1
39	39	39	1568	CUSTOMF0001	Custom Frames, Inc.	2	1	1
40	40	40	1570	FIRSTRA0001	First Rate Bicycles	1	1	1

Query executed successfully. DESKTOP-PTHVTES\MSSQLSERVER... DESKTOP-PTHVTES\duydu ... DW_PURCHASING 00:00:00 104 rows

Figure 4.13. SQL Query result of ETL DIMVendor process (Source: Authors propose)

4.1.2. Fact table's ETL Process

4.1.2.1. ETL Process of FACTInventoryManagement

Step 1: Create FactInventoryManagement Package: Create a new SSIS package named FactInventoryManagement. In the Control Flow, implement logic to create and drop the FactInventoryManagement table if it already exists.

Step 2: Add Constraint with SQL Statement.

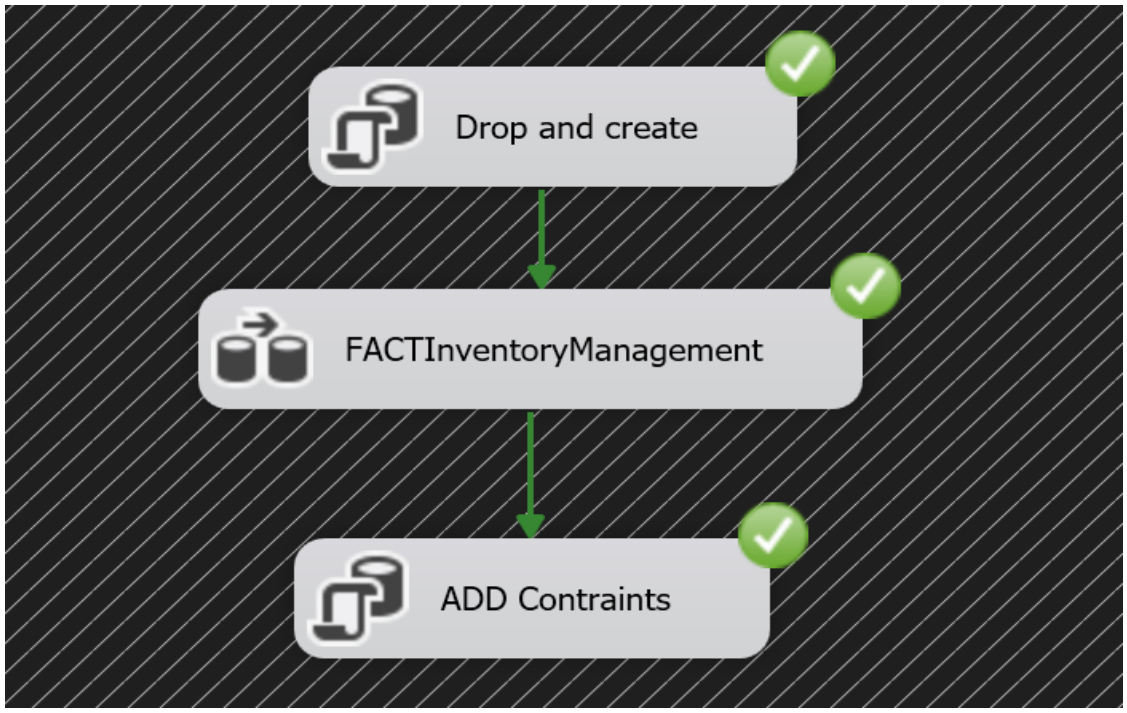


Figure 4.14. Create Data Flow Task for FactInventoryManagement (Source: Authors propose)

Step 3: Data Flow Task - Integrate and Configure: Within the Data Flow Task, integrate data from the Production.ProductInventory in AW2022 databases.

Step 4: Lookup with DimTime:

- Add a Lookup Transformation after the OLE DB Source.

- Configure the Lookup Transformation to join with the DimTime table using appropriate matching columns.

- Redirect rows with no match to a No Match Output.

Step 5: Lookup with DimLocation:

- Add another Lookup Transformation after the No Match Output of the previous Lookup.

- Configure this Lookup Transformation to join with the DimLocation table using appropriate matching columns.

- Redirect rows with no match to a No Match Output.

Step 6: Lookup with DimProduct:

- Add another Lookup Transformation after the No Match Output of the previous Lookup.

- Configure this Lookup Transformation to join with the DimProduct table using appropriate matching columns.

- Redirect rows with no match to a No Match Output.

Step 8: Derived Column Transformation:

- Add a Derived Column Transformation after the No Match Output of the last Lookup.

- Create a new column named ValueOnHand with the expression Quantity * Standard Cost.

Step 9: Destination: FactInventoryManagement:

- Add an OLE DB Destination component to the Data Flow Task.

- Configure the OLE DB Destination to connect to your destination database and specify the FactInventoryManagement table as the destination.

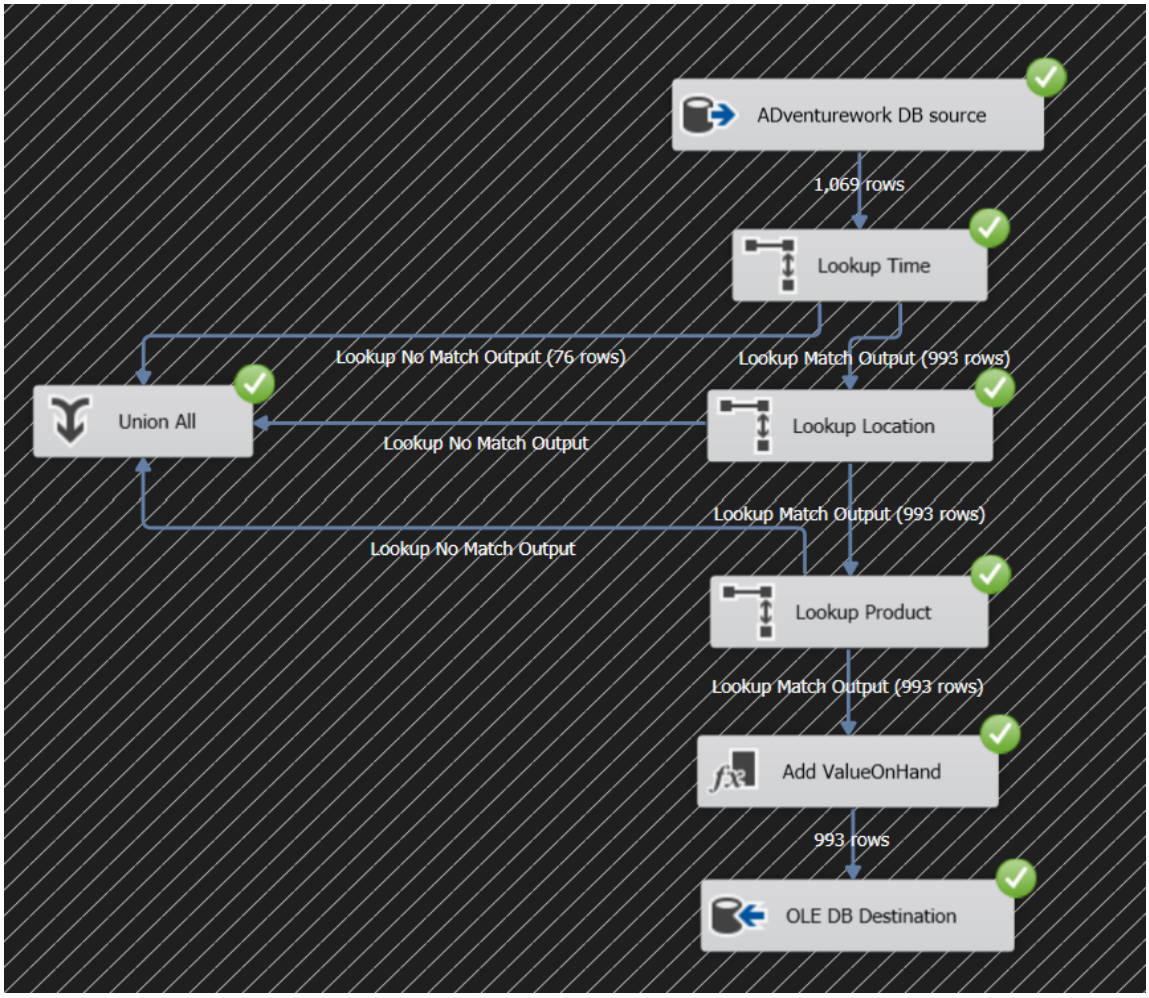


Figure 4.15. ETL Process of FactInventoryManagement (Source: Authors propose)

4.1.2.2. ETL Process of FACTProcurmentPerformance

Step 1: Create FactProcurmentPerformance Package: Create a new SSIS package named FactProcurmentPerformance. In the Control Flow, implement logic to create and drop the FactProcurmentPerformance table if it already exists.

Step 2: Add Constraint with SQL Statement.

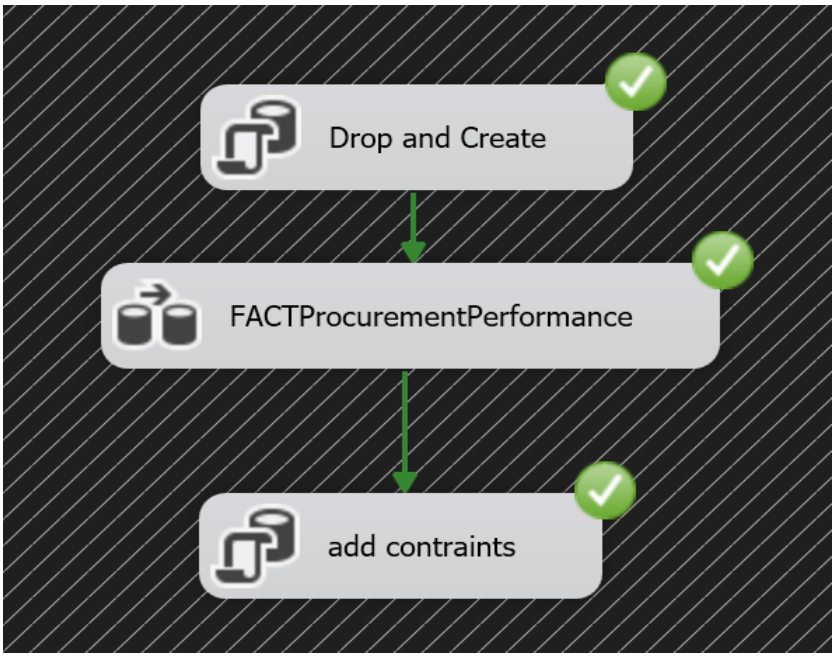


Figure 4.16. Create Data Flow Task for FACTProcurementManagement (Source: Authors propose)

Step 3: Data Flow Task - Integrate and Configure: Within the Data Flow Task, integrate data from the Purchasing in AW2022 databases.

Step 4: Lookup with DimTime:

- Add a Lookup Transformation after the OLE DB Source.

- Configure the Lookup Transformation to join with the DimTime table using appropriate matching columns.

- Redirect rows with no match to a No Match Output.

Step 5: Lookup with DimProduct:

- Add another Lookup Transformation after the No Match Output of the previous Lookup.

- Configure this Lookup Transformation to join with the DimProduct table using appropriate matching columns.

- Redirect rows with no match to a No Match Output.

Step 6: Lookup with DimVendor:

- Add another Lookup Transformation after the No Match Output of the previous Lookup.

- Configure this Lookup Transformation to join with the DimVendor table using appropriate matching columns.

- Redirect rows with no match to a No Match Output.

Step 7: Lookup with DimGeography:

- Add another Lookup Transformation after the No Match Output of the previous Lookup.

- Configure this Lookup Transformation to join with the DimGeography table using appropriate matching columns.

- Redirect rows with no match to a No Match Output.

Step 8: Derived Column Transformation:

- Add a Derived Column Transformation after the No Match Output of the last Lookup.

- Change datatype of 3 columns (OrderDate, ShipDate and DueDate) into an integer value representing the date in the format ‘YYYYMMDD’.

Step 9: Destination: FactProcurmentPerformance:

- Add an OLE DB Destination component to the Data Flow Task.

- Configure the OLE DB Destination to connect to your destination database and specify the FactProcurmentPerformance table as the destination.

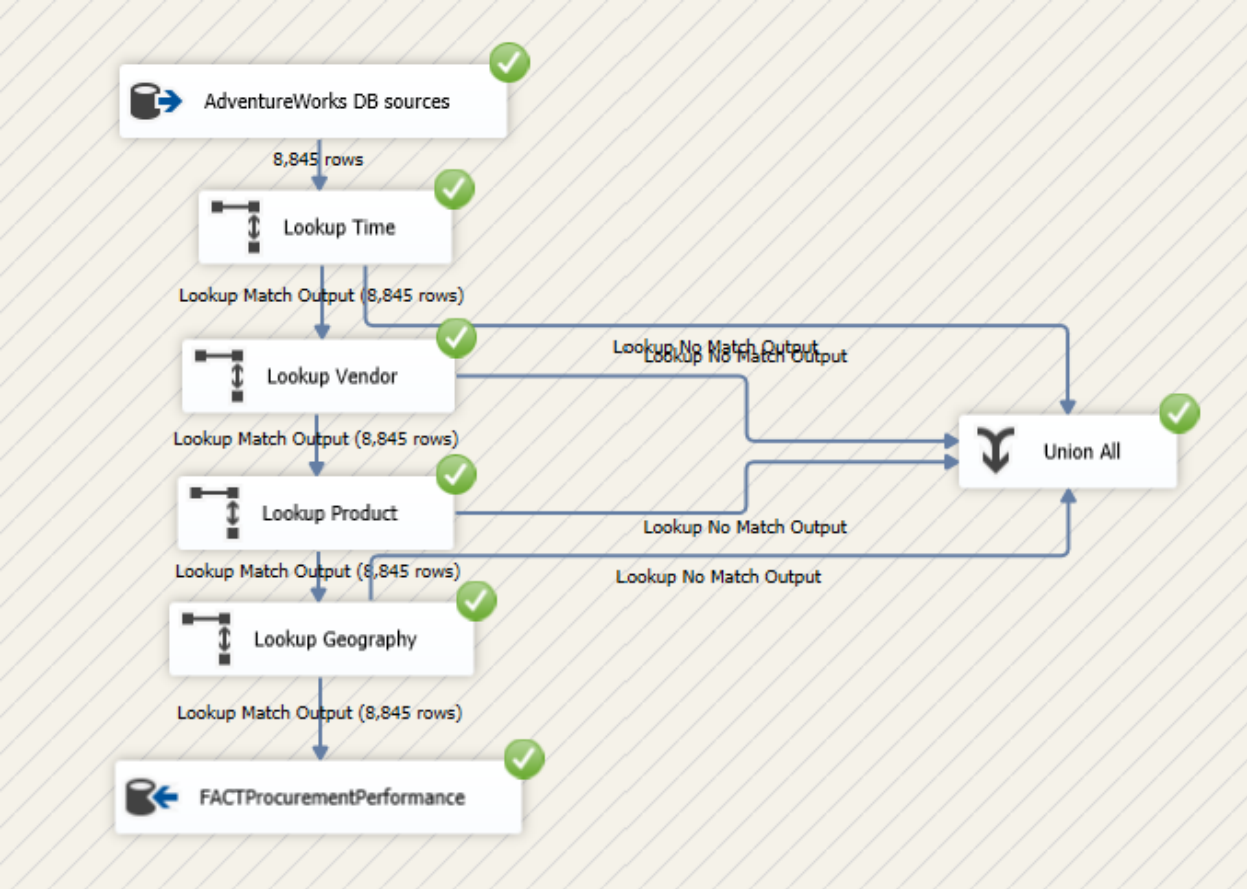
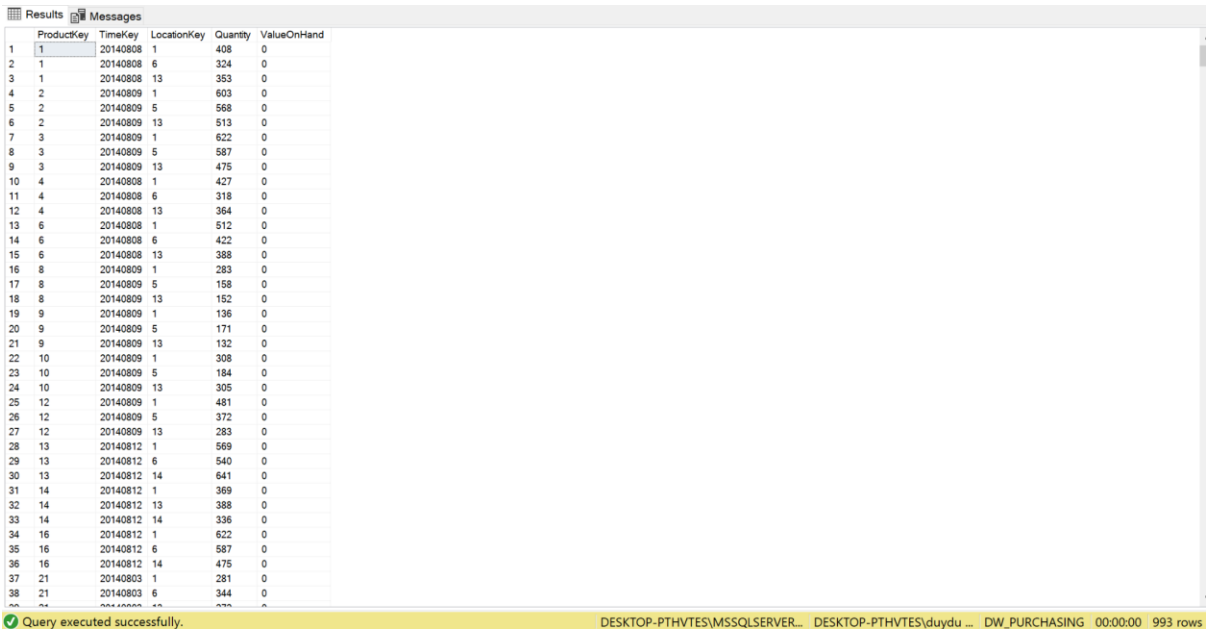


Figure 4.17. ETL Process of FACTProcurementManagement (Source: Authors propose)

4.2. Integration result



Results Messages

	ProductKey	TimeKey	LocationKey	Quantity	ValueOnHand
1	1	20140808	1	408	0
2	1	20140808	6	324	0
3	1	20140808	13	353	0
4	2	20140809	1	603	0
5	2	20140809	5	568	0
6	2	20140809	13	513	0
7	3	20140809	1	622	0
8	3	20140809	5	587	0
9	3	20140809	13	475	0
10	4	20140808	1	427	0
11	4	20140808	6	318	0
12	4	20140808	13	364	0
13	6	20140808	1	512	0
14	6	20140808	6	422	0
15	6	20140808	13	388	0
16	8	20140809	1	283	0
17	8	20140809	5	158	0
18	8	20140809	13	152	0
19	9	20140809	1	136	0
20	9	20140809	5	171	0
21	9	20140809	13	132	0
22	10	20140809	1	308	0
23	10	20140809	5	184	0
24	10	20140809	13	305	0
25	12	20140809	1	481	0
26	12	20140809	5	372	0
27	12	20140809	13	283	0
28	13	20140812	1	569	0
29	13	20140812	6	540	0
30	13	20140812	14	641	0
31	14	20140812	1	369	0
32	14	20140812	13	388	0
33	14	20140812	14	336	0
34	16	20140812	1	622	0
35	16	20140812	6	587	0
36	16	20140812	14	475	0
37	21	20140803	1	281	0
38	21	20140803	6	344	0

Query executed successfully. DESKTOP-PTHVTES\MSSQLSERVER... DESKTOP-PTHVTES\duydu ... DW_PURCHASING 00:00:00 993 rows

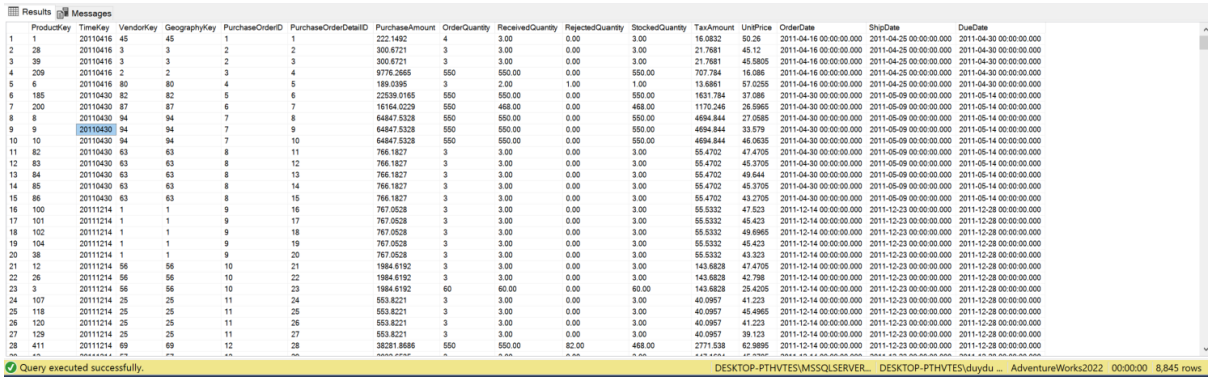
Figure 4.18. SQL Query result of ETL FACTInventoryManagement process (Source: Authors propose)

The FACTInventoryManagement table is a crucial component of the data warehouse, capturing and organizing vital information for inventory management in the business ecosystem. It consists of five columns: ProductKey, TimeKey, LocationKey, Quantity, and ValueOnHand. Each row represents a unique combination of products, times, locations, quantities, and values on hand. The ProductKey allows for seamless integration, while the TimeKey enables temporal analysis. The Quantity column represents the volume of inventory, and the ValueOnHand column represents the monetary value. However, some records in the ValueOnHand field might exhibit zero values, which could be due to factors such as stock depletion, delayed valuation, or data entry errors. Analyzing these zero values is essential to improve inventory management practices and optimize supply chain operations.

In the FACTInventoryManagement table, the integration and organization of information regarding inventory management are critical for effective business operations. This table contains five columns: ProductKey, TimeKey, LocationKey, Quantity, and ValueOnHand. Each row represents a unique combination of a product, time, location, quantity, and value on hand. The ProductKey facilitates seamless integration, while the

TimeKey enables temporal analysis. The Quantity column showcases the volume of inventory, while the ValueOnHand column represents its financial value. However, it is noteworthy that certain records in the ValueOnHand field might display zero values. This occurrence could be attributed to various factors, including stock depletion, delayed valuation processes, or data entry errors. Understanding the causes behind these zero values is crucial in addressing inventory discrepancies and optimizing supply chain management.

The FACTInventoryManagement table plays a vital role in capturing and organizing information relevant to efficient inventory management. This table encompasses five distinct columns: ProductKey, TimeKey, LocationKey, Quantity, and ValueOnHand. Each row represents a unique combination of product, time, location, quantity, and value. The ProductKey ensures seamless integration of data, while the TimeKey allows for temporal analysis. The Quantity column showcases the inventory volume, and the ValueOnHand column depicts the monetary value. It is important to note that some records within the ValueOnHand column might have zero values. This could be due to factors such as stock depletion, delays in valuation processes, or errors in data entry. Understanding the reasons behind these zero values is essential in improving inventory management practices and optimizing the efficiency of the overall supply chain.



	ProductKey	TimeKey	VendorKey	GeographyKey	PurchaseOrderID	PurchaseOrderDetailID	PurchaseAmount	OrderQuantity	ReceivedQuantity	RejectedQuantity	StockedQuantity	TaxAmount	UnitPrice	OrderDate	ShipDate	DueDate
1	1	20110416	45	45	1	1	222.1492	4	3.00	0.00	3.00	16.0832	50.26	2011-04-16 00:00:00.000	2011-04-25 00:00:00.000	2011-04-30 00:00:00.000
2	28	20110416	3	3	2	2	300.6721	3	3.00	0.00	3.00	21.7681	45.12	2011-04-16 00:00:00.000	2011-04-25 00:00:00.000	2011-04-30 00:00:00.000
3	39	20110416	3	3	2	3	300.6721	3	3.00	0.00	3.00	21.7681	45.5805	2011-04-16 00:00:00.000	2011-04-25 00:00:00.000	2011-04-30 00:00:00.000
4	209	20110416	2	2	3	4	9776.2665	550	550.00	0.00	550.00	707.784	16.086	2011-04-16 00:00:00.000	2011-04-25 00:00:00.000	2011-04-30 00:00:00.000
5	6	20110416	80	80	4	5	189.0395	3	2.00	1.00	1.00	13.6861	57.0255	2011-04-16 00:00:00.000	2011-04-25 00:00:00.000	2011-04-30 00:00:00.000
6	185	20110430	82	82	5	6	22539.0165	550	550.00	0.00	550.00	1631.784	37.086	2011-04-30 00:00:00.000	2011-05-09 00:00:00.000	2011-05-14 00:00:00.000
7	200	20110430	87	87	6	7	16164.0229	550	468.00	0.00	468.00	1170.246	26.5965	2011-04-30 00:00:00.000	2011-05-09 00:00:00.000	2011-05-14 00:00:00.000
8	8	20110430	94	94	7	8	64847.5328	550	550.00	0.00	550.00	4694.844	27.0585	2011-04-30 00:00:00.000	2011-05-09 00:00:00.000	2011-05-14 00:00:00.000
9	9	20110430	94	94	7	9	64847.5328	550	550.00	0.00	550.00	4694.844	33.579	2011-04-30 00:00:00.000	2011-05-09 00:00:00.000	2011-05-14 00:00:00.000
10	10	20110430	94	94	7	10	64847.5328	550	550.00	0.00	550.00	4694.844	46.0635	2011-04-30 00:00:00.000	2011-05-09 00:00:00.000	2011-05-14 00:00:00.000
11	82	20110430	63	63	8	11	766.1827	3	3.00	0.00	3.00	55.4702	47.4705	2011-04-30 00:00:00.000	2011-05-09 00:00:00.000	2011-05-14 00:00:00.000
12	83	20110430	63	63	8	12	766.1827	3	3.00	0.00	3.00	55.4702	45.3705	2011-04-30 00:00:00.000	2011-05-09 00:00:00.000	2011-05-14 00:00:00.000
13	84	20110430	63	63	8	13	766.1827	3	3.00	0.00	3.00	55.4702	49.644	2011-04-30 00:00:00.000	2011-05-09 00:00:00.000	2011-05-14 00:00:00.000
14	85	20110430	63	63	8	14	766.1827	3	3.00	0.00	3.00	55.4702	45.3705	2011-04-30 00:00:00.000	2011-05-09 00:00:00.000	2011-05-14 00:00:00.000
15	86	20110430	63	63	8	15	766.1827	3	3.00	0.00	3.00	55.4702	43.2705	2011-04-30 00:00:00.000	2011-05-09 00:00:00.000	2011-05-14 00:00:00.000
16	100	20111214	1	1	9	16	767.0528	3	3.00	0.00	3.00	55.5332	47.523	2011-12-14 00:00:00.000	2011-12-23 00:00:00.000	2011-12-28 00:00:00.000
17	101	20111214	1	1	9	17	767.0528	3	3.00	0.00	3.00	55.5332	45.423	2011-12-14 00:00:00.000	2011-12-23 00:00:00.000	2011-12-28 00:00:00.000
18	102	20111214	1	1	9	18	767.0528	3	3.00	0.00	3.00	55.5332	49.6965	2011-12-14 00:00:00.000	2011-12-23 00:00:00.000	2011-12-28 00:00:00.000
19	104	20111214	1	1	9	19	767.0528	3	3.00	0.00	3.00	55.5332	45.423	2011-12-14 00:00:00.000	2011-12-23 00:00:00.000	2011-12-28 00:00:00.000
20	38	20111214	1	1	9	20	767.0528	3	3.00	0.00	3.00	55.5332	43.323	2011-12-14 00:00:00.000	2011-12-23 00:00:00.000	2011-12-28 00:00:00.000
21	12	20111214	56	56	10	21	1984.6192	3	3.00	0.00	3.00	143.6828	47.4705	2011-12-14 00:00:00.000	2011-12-23 00:00:00.000	2011-12-28 00:00:00.000
22	26	20111214	56	56	10	22	1984.6192	3	3.00	0.00	3.00	143.6828	42.798	2011-12-14 00:00:00.000	2011-12-23 00:00:00.000	2011-12-28 00:00:00.000
23	3	20111214	56	56	10	23	1984.6192	60	60.00	0.00	60.00	143.6828	25.4205	2011-12-14 00:00:00.000	2011-12-23 00:00:00.000	2011-12-28 00:00:00.000
24	107	20111214	25	25	11	24	553.8221	3	3.00	0.00	3.00	40.0957	41.223	2011-12-14 00:00:00.000	2011-12-23 00:00:00.000	2011-12-28 00:00:00.000
25	118	20111214	25	25	11	25	553.8221	3	3.00	0.00	3.00	40.0957	45.4965	2011-12-14 00:00:00.000	2011-12-23 00:00:00.000	2011-12-28 00:00:00.000
26	120	20111214	25	25	11	26	553.8221	3	3.00	0.00	3.00	40.0957	41.223	2011-12-14 00:00:00.000	2011-12-23 00:00:00.000	2011-12-28 00:00:00.000
27	129	20111214	25	25	11	27	553.8221	3	3.00	0.00	3.00	40.0957	39.123	2011-12-14 00:00:00.000	2011-12-23 00:00:00.000	2011-12-28 00:00:00.000
28	411	20111214	69	69	12	28	38281.8686	550	550.00	82.00	468.00	2771.538	62.9895	2011-12-14 00:00:00.000	2011-12-23 00:00:00.000	2011-12-28 00:00:00.000

Figure 4.19. SQL Query result of ETL FACTProcurementPerformance process (Source: Authors propose)

The FACTProcurementPerformance table, housed within the SQL server, is a comprehensive collection of columns capturing crucial data related to procurement activities. The table includes information such as ProductKey, VendorKey,

GeographyKey, PurchaseOrderID, PurchaseOrderDetailID, PurchaseAmount, OrderQuantity, ReceivedQuantity, RejectedQuantity, StockedQuantity, TaxAmount, UnitPrice, OrderDate, ShipDate, and DueDate. By analyzing this data, organizations can gain valuable insights to drive procurement performance analysis and support effective vendor selection.

The table's key columns provide multidimensional perspectives for evaluating procurement performance. ProductKey and VendorKey allow organizations to assess performance from product and vendor standpoints, while GeographyKey provides visibility into geographical aspects. The inclusion of time-related columns, like OrderDate, ShipDate, and DueDate, enables temporal analysis and trend identification. Additionally, the table’s financial columns, such as PurchaseAmount and TaxAmount, enable cost analysis while OrderQuantity, ReceivedQuantity, RejectedQuantity, and StockedQuantity offer insights into efficiency and inventory management.

With its diverse data columns, the FACTProcurementPerformance table serves as a valuable resource for organizations seeking to optimize procurement performance and make informed vendor selection decisions. By leveraging the captured data, businesses can enhance procurement operations, identify areas for improvement, and achieve cost optimization. The table’s comprehensive nature and the depth of information it provides enable organizations to implement data-driven strategies to drive success in their procurement endeavors.

Chapter 5: Multi-dimensional data analysis

In this chapter, specific analytical methods will be presented to assist in generating the most useful insights or information, along with reports and queries that will be executed using specific tools. The process of obtaining and visualizing the data on a broader scale will be outlined to ensure that administrators can understand and readily make informed decisions for their business.

5.1. Multi-dimensional analysis strategy

Following the integration of data into the Data Warehouse, the selection of OLAP multidimensional database is a strategic decision aimed at enhancing comprehensiveness and efficiency in purchasing data analysis. The comprehensiveness of OLAP allows us to analyze data from various perspectives, including time, product, supplier, and geography. This not only helps understand the relationships and trends in purchasing data but also provides a comprehensive and detailed view of the enterprise's purchasing activities. The main objective of using OLAP is to strengthen strategic decision-making and optimize the purchasing process. By organizing data in a structured and understandable manner, OLAP supports strategic decision-making by providing accurate and detailed information. It enables managers and decision-makers to make smarter and more strategic decisions, from supplier selection to inventory management and marketing strategy.

Multi-dimensional data in the Purchasing module

The application of a multi-dimensional database to the Purchasing module of Adventureworks not only optimizes the procurement management process but also brings significant strategic benefits. One of the most important advantages of using a multi-dimensional database is the ability to analyze data from various perspectives. By organizing data into dimensions such as time, product, supplier, and geography, we can gain a deeper understanding of the factors influencing the company's purchasing activities, enabling us to evaluate overall purchasing trends, cost factors, and supplier relationships.

Furthermore, multi-dimensional access supports the identification of key performance indicators (KPIs) related to the Purchasing module. By defining and

monitoring KPIs across multiple data dimensions such as the number of processed orders, average processing time, supplier delivery performance, and cost fluctuations, the module can track its performance and assess progress towards strategic objectives.

Moreover, the application of multi-dimensional databases supports advanced analysis techniques such as data mining, forecasting, and trend analysis. These techniques enable the Purchasing module to uncover hidden patterns, predict future demand, and forecast market fluctuations, thereby supporting proactive decision-making and risk mitigation strategies.

5.2. Building Cube

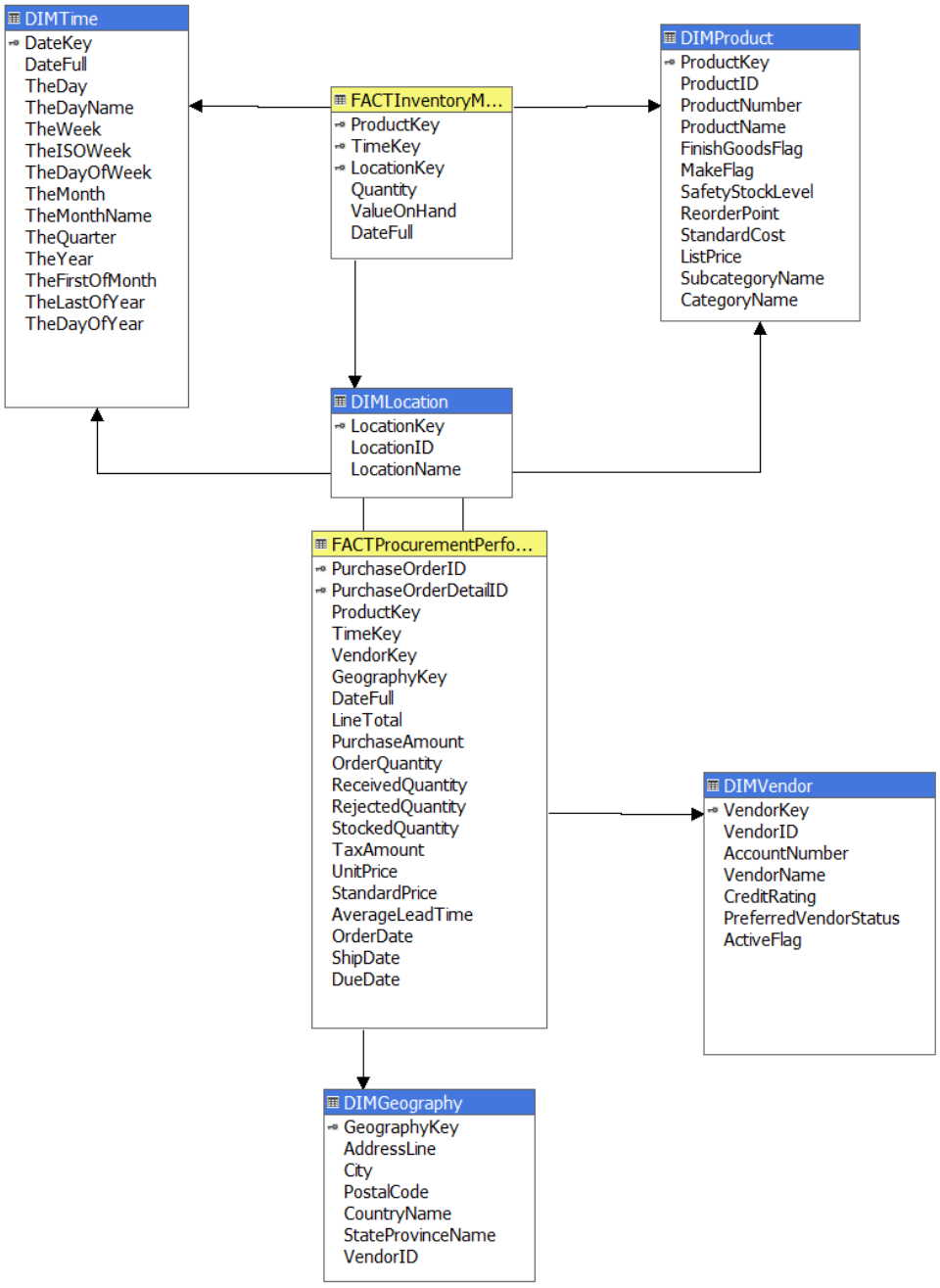


Figure 5.1. Buiding Cube (Source: Authors propose)

In our process to set up a data warehouse using Visual Studio, creating a data cube is an important milestone. Using this cube lets us examine data from various angles, and helps us to understand complex data patterns, aiding strategic business decisions.

Using the tools available in Visual Studio - particularly the SQL Server Analysis Services (SSIS) - we can make a data cube. This cube provides a three-dimensional view of our data, which greatly improves our ability to analyze and report on the data in our warehouse.

Our data cube is made up of several different tables: FACTProcurementPerformance and FACTInventoryPerformance are the central tables, holding the core information about our business processes. These tables connect to DIMProduct, DIMVendor, DIMGeography, DIMTime, and DIMLocation, which provide the additional details that bring our data to life.

The FACTProcurementPerformance table, for instance, contains details such as the PurchaseOrderID, which links to DIMProduct and DIMVendor. This connection allows our data cube to maintain a clear and comprehensive picture of our procurement operations.

Likewise, the FACTInventoryPerformance is linked to DIMProduct and DIMLocation. The resultant linkage allows our data cube to present a detailed picture of the inventory performance in various locations.

To connect these tables together, we use key identifiers- unique codes or identifiers for each record. In FACTProcurementPerformance, for example, we use a specific ProductKey and VendorKey to link to the DIMProduct and DIMVendor tables respectively. Similarly, FACTInventoryPerformance is connected to DIMProduct and DIMLocation using a ProductKey and LocationKey. These relationships are established in Visual Studio via a graphical representation tool that showcases the data tables and their connections.

In our process to set up a data warehouse using Visual Studio, creating a data cube is an important milestone. Using this cube lets us examine data from various angles, and helps us to understand complex data patterns, aiding strategic business decisions.

Using the tools available in Visual Studio - particularly the SQL Server Analysis Services (SSIS) - we can make a data cube. This cube provides a three-dimensional view of our data, which greatly improves our ability to analyze and report on the data in our warehouse.

Our data cube is made up of several different tables: FACTProcurementPerformance and FACTInventoryPerformance are the central tables, holding the core information about our business processes. These tables connect to DIMProduct, DIMVendor, DIMGeography, DIMTime, and DIMLocation, which provide the additional details that bring our data to life.

The FACTProcurementPerformance table, for instance, contains details such as the PurchaseOrderID, which links to DIMProduct and DIMVendor. This connection allows our data cube to maintain a clear and comprehensive picture of our procurement operations.

Likewise, the FACTInventoryPerformance is linked to DIMProduct and DIMLocation. The resultant linkage allows our data cube to present a detailed picture of the inventory performance in various locations.

To connect these tables together, we use key identifiers- unique codes or identifiers for each record. In FACTProcurementPerformance, for example, we use a specific ProductKey and VendorKey to link to the DIMProduct and DIMVendor tables respectively. Similarly, FACTInventoryPerformance is connected to DIMProduct and DIMLocation using a ProductKey and LocationKey. These relationships are established in Visual Studio via a graphical representation tool that showcases the data tables and their connections.

In short, building a data cube using a set of interconnected tables gives us a view of our data that is multi-dimensional. This allows us to gain deeper insights into our business operations, aiding in effective decision making.

5.3. Analysis with SSAS

The thesis employs SSAS to examine data from the data warehouse, generating helpful reports for decision-making. In this phase, the defined KPIs will be created in SSAS using Multidimensional Expressions (MDX) query language. To create them, the first step is to define the name and purpose of the KPI. The formulas to calculate them should be mentioned in Part 2.1.2. KPI analysis The construction involves the following steps:

- Step 1: Begin SQL Server Management Studio and connect to the Analysis Services instance for Cube creation.

- Step 2: Create a new Data Source.
- Step 3: Define the type of Data Source.
- Step 4: Design a new Cube.
- Step 5: Improve the Cube by adding Measures and Dimensions.
- Step 6: Create Calculations and, if necessary, Key Performance Indicators (KPIs).

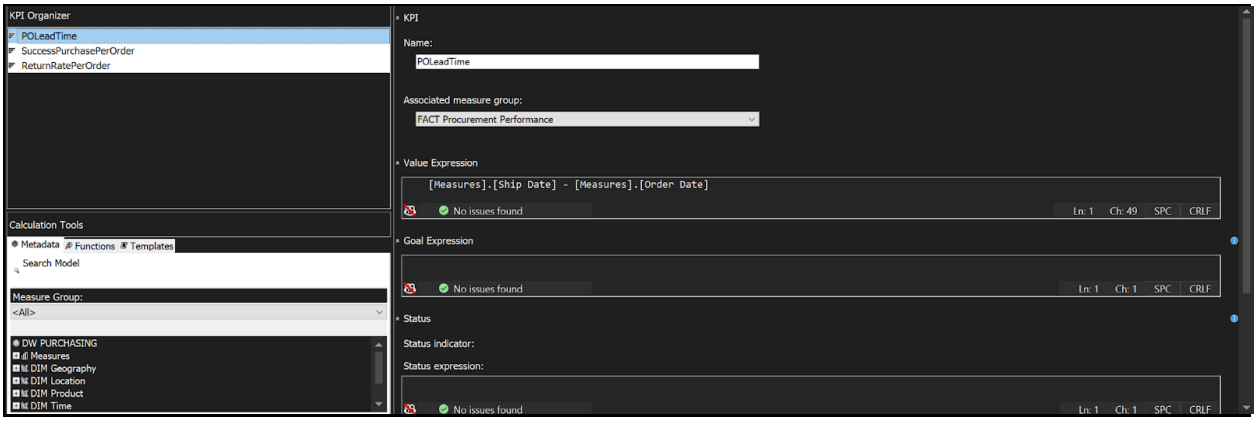


Figure 5.2. POLeadTime Generation

The KPI "POLeadTime" is a measure that enables the tracking and analysis of the purchase order lead time in manufacturing processes. It measures the number of days it takes for items to arrive after an order is placed. This KPI is particularly relevant in manufacturing industries where project managers need to keep up with production needs by closely monitoring the shipping times required for various parts. By having visibility into the lead time, project managers can make informed decisions to optimize production schedules and ensure timely delivery of materials.

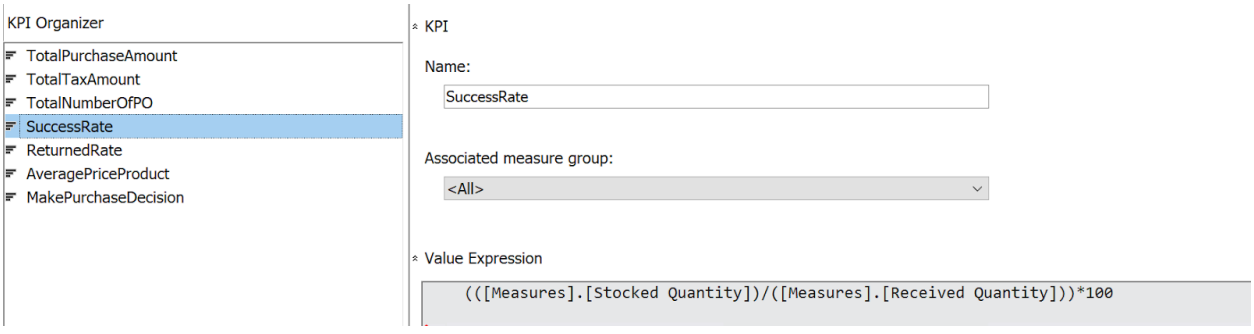


Figure 5.3. SuccessRate Generation

The "*SuccessRate*" KPI is a valuable metric that measures the percentage of successful purchases completed per order. This KPI provides insights into the efficiency of the purchase process and the organization's ability to fulfill orders without any issues. By tracking the success rate of purchases, businesses can evaluate the effectiveness of their procurement operations, identify any bottlenecks or challenges, and implement necessary improvements to enhance order fulfillment and customer satisfaction.

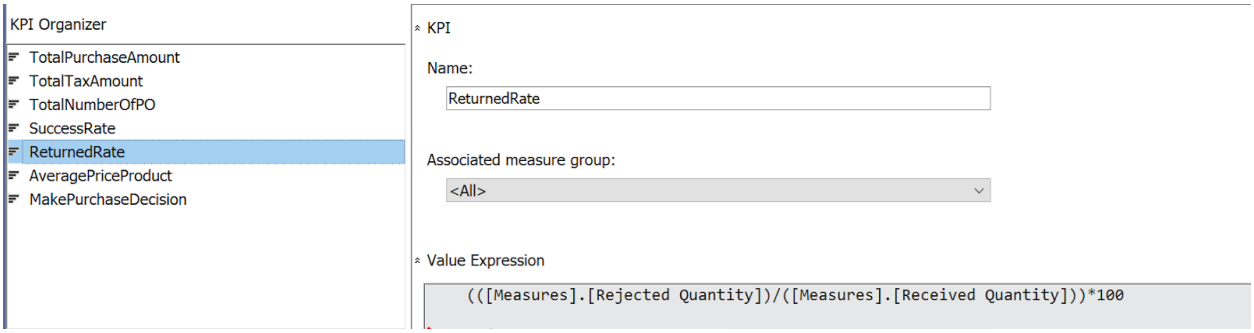


Figure 5.4. ReturnRatePerOrder Generation

The "*ReturnRatePerOrder*" KPI measures the percentage of returned items per order. This KPI provides insights into the quality of products, customer satisfaction, and the effectiveness of the return process. By monitoring the return rate, organizations can identify patterns and trends in returns, assess the reasons behind them, and take corrective actions to enhance product quality, streamline the return process, and minimize customer dissatisfaction.

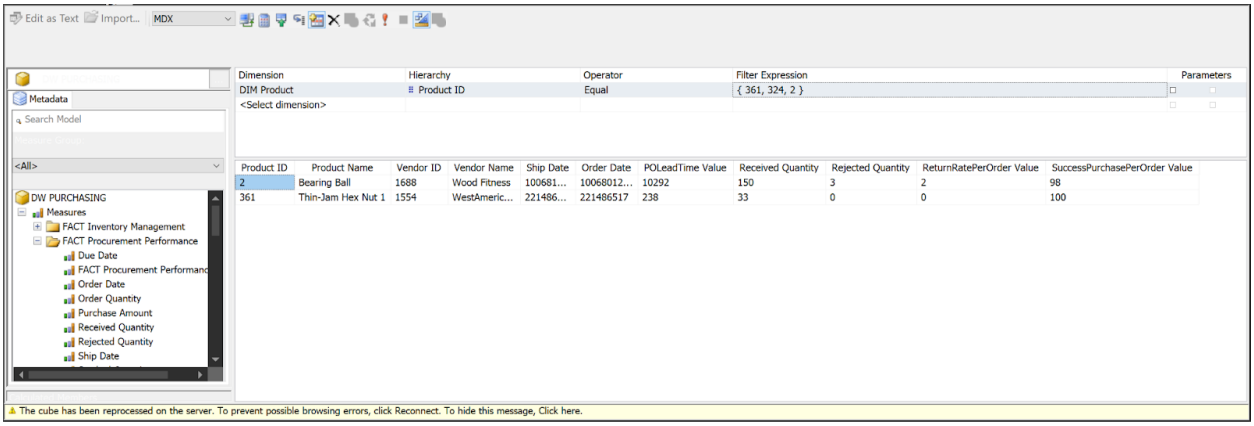


Figure 5.5. The three KPIs generated

The *Figure 5.5.* show results of executed query with the 3 KPIs above: POLeadTime, SuccessPurchasePerOrder, ReturnRatePerOrder using MDX language.

Product ID	Product Name	Vendor ID	AveragePriceProduct Value
1	Adjustable R...	1580	55.5410662337662
2	Bearing Ball	1688	46.3171666666667
317	LL Crankarm	1578	112.837213181818
317	LL Crankarm	1678	103.907058990909
318	ML Crankarm	1578	112.837213181818
318	ML Crankarm	1678	103.907058990909
319	HL Crankarm	1556	51.2946525454545
319	HL Crankarm	1578	108.114995681818
319	HL Crankarm	1678	99.4783847409091
320	Chainring Bolts	1514	602.125004761905
320	Chainring Bolts	1602	507.04664
320	Chainring Bolts	1604	569.146044186047
321	Chainring Nut	1514	602.125004761905
321	Chainring Nut	1602	507.04664
321	Chainring Nut	1604	569.146044186047
322	Chainring	1514	35.8881885772358

Figure 5.6. The "AveragePriceProduct Value" KPI

The "AveragePriceProduct Value" KPI refers to the calculated average price of a product. This value is derived by summing up the prices of all instances of a particular product and then dividing the total by the number of instances. It provides insight into the typical cost of a specific product, which can be useful for various purposes such as pricing strategies, cost analysis, and financial forecasting.

Dimension	Hierarchy	Operator	Filter Expression	Param...
DIM Time	The Year	Equal	{ 2013, 2014, 2015 }	
<Select dimension>				

Product ID	Product Name	The Year	AveragePriceProduct V...
1	Adjustable R...	2013	55.5411666666667
1	Adjustable R...	2014	55.5411666666667
2	Bearing Ball	2013	46.3171666666667
2	Bearing Ball	2014	46.3171666666667
317	LL Crankarm	2013	108.377613804196
317	LL Crankarm	2014	109.008388181818
318	ML Crankarm	2013	108.377613804196
318	ML Crankarm	2014	109.008388181818
319	HL Crankarm	2013	82.2646256735537
319	HL Crankarm	2014	84.1190175524476
320	Chainring Bolts	2013	571.056420930233
320	Chainring Bolts	2014	544.149505555556
321	Chainring Nut	2013	571.056420930233
321	Chainring Nut	2014	544.149505555556
322	Chainring	2013	33.7028036507936
322	Chainring	2014	33.37502625

Figure 5.7. The "AveragePriceProduct Value" KPI

The insights on "AveragePriceProduct Value" by specific year can be obtained by the integration with DIM Time. It involves calculating the average price of a specific product across all transactions or instances within that year. This analysis allows for a more granular understanding of how the price of the product fluctuates over time, which can be valuable for identifying trends, assessing market dynamics, and making informed decisions related to pricing strategies and financial planning.

Product ID	Product Name	Safety Stock Level	Reorder Point	Quantity
1	Adjustable R...	1000	750	1085
2	Bearing Ball	1000	750	1109
3	BB Ball Beari...	800	600	1352
316	Blade	800	600	1361
317	LL Crankarm	500	375	593
318	ML Crankarm	500	375	439
319	HL Crankarm	500	375	797
320	Chainring Bolts	1000	750	1136
321	Chainring Nut	1000	750	1750
322	Chainring	1000	750	1684
323	Crown Race	1000	750	1684
324	Chain Stays	1000	750	1629
325	Decal 1	1000	750	1750
326	Decal 2	1000	750	1684
327	Down Tube	800	600	1364
328	Mountain En...	1000	750	1629

Figure 5.8. The "Safety Stock Level" KPI

Safety Stock Level refers to the extra inventory held by a company as a buffer against fluctuations in demand or supply. It acts as a safeguard to ensure that the company does not run out of stock when unexpected events occur, such as sudden increases in demand or delays in replenishment. The safety stock level is typically set above the average expected demand during the lead time it takes to replenish inventory. By maintaining a safety stock, a company can reduce the risk of stockouts and avoid potential disruptions to its operations.

Reorder Point is the inventory level at which a company needs to reorder a particular product to replenish its stock. It is calculated based on factors such as the lead time required to receive new inventory, the average demand during that lead time, and the desired level of safety stock. When the actual inventory level reaches or falls below the reorder point, it triggers the replenishment process, ensuring that the company maintains sufficient stock to meet customer demand without experiencing stockouts. The reorder point helps optimize inventory management by ensuring that inventory is replenished at the right time, minimizing the risk of overstocking or understocking.

Category Name	Value On Hand
Accessories	88212
Bikes	14623653
Clothing	138133
Components	4375836
NA	866847

Figure 5.9. The "Value On Hand" KPI

Another insightful KPI is the "Value On Hand", which demonstrates the monetary worth of the inventory or assets currently held by a business at a specific point in time. It represents the total value of all products, materials, or goods available in inventory, usually calculated by multiplying the quantity of each item by its unit price. Understanding the value on hand is crucial for businesses to assess their financial health, manage cash flow, and make informed decisions about inventory levels, pricing strategies, and investment priorities. It provides insight into the value of the assets that the business can potentially convert into revenue through sales or other means.

TotalTaxAmount Value
13091014.2

Figure 5.10. The "TotalAmountValue" KPI

The "TotalAmountValue" KPI is the combined monetary value of all items or transactions in a given context, calculated by adding up their individual values. It offers a holistic perspective on the financial worth of the items under consideration, serving various purposes like financial reporting and decision-making.

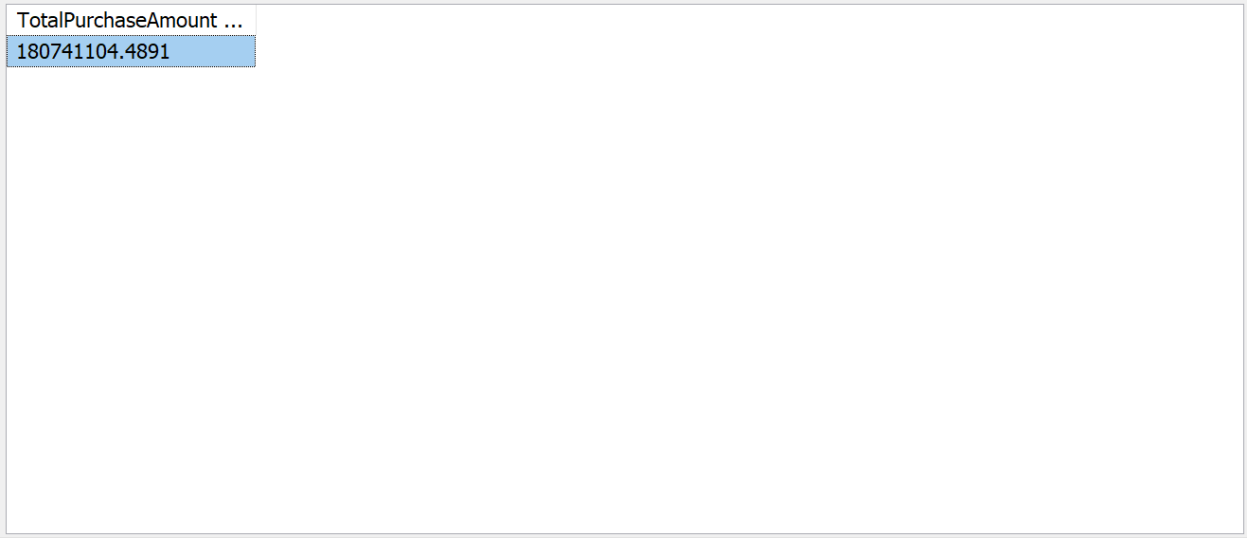


Figure 5.11. The "TotalPurchaseAmount" KPI

The KPI "TotalPurchaseAmount" represents the overall monetary value of all purchases made within a specific timeframe or across a certain set of transactions. It is calculated by summing up the purchase amounts for all transactions during the defined period. "TotalPurchaseAmount" provides insight into the total expenditure incurred by a business or entity on acquiring goods, services, or assets.

Chapter 6: Data Analytics and Visualization

This chapter focuses on the development and analysis of two main dashboards in AdventureWorks' Business Intelligence (BI) project: the Inventory Management Dashboard and the Procurement Performance Dashboard. it outlines the objectives and scope of these dashboards, highlighting their importance in providing actionable insights for decision-making within the Purchasing department.

6.1. Report and dashboard systems

In the realm of Business Intelligence (BI) implementation, the efficacy of data-driven decision-making hinges significantly on the clarity and accessibility of insights conveyed through reports and dashboards. These visualizations serve as the conduit through which complex data is distilled into actionable intelligence, facilitating informed choices that drive business success. Within the context of Adventure Works' BI initiative, the development of intuitive and insightful dashboards is paramount to empowering stakeholders within the Purchasing department to extract maximum value from their data assets. In this report, we delve into the development of three distinct dashboards – the "Inventory Management" Dashboard and the Procurement Performance Dashboard.

- ***Scope***: based on the Purchasing department of the dataset from 2011 to 2014 of AdventureWorks Company.

- ***Objective***: By transforming raw data into intuitive visualizations, the primary objective of the dashboards is to visually represent key metrics and data points to address the specified requirements outlined for supporting decision-making processes within AdventureWorks' Purchasing department. The overarching goal is to provide stakeholders with actionable insights that facilitate informed decisions regarding inventory management, procurement performance, and supplier relationships. Through interactive features, customizable views, and real-time updates, the dashboards aim to empower users to analyze trends, identify opportunities for improvement, and drive strategic initiatives for enhancing operational efficiency and cost-effectiveness.

6.2. Data analysis with Power BI

6.2.1. The Inventory Management Dashboard

We have built the Inventory Management Dashboard as a comprehensive tool, providing real-time information on key inventory metrics. This dashboard enables warehouse managers to monitor inventory levels, turnover rates, and identify potential stock outs or excess stock situations. Through dynamic visualizations such as bar charts, line graphs, and KPI widgets, stakeholders can observe inventory metrics including quantity, location, and specific product details such as product codes and current value. Moreover, the dashboard supports tracking of available products and safety stock levels, empowering managers to make proactive decisions regarding purchase requisitions and inventory optimization. With customizable views, drill-down analysis capabilities, and proactive alerts, the dashboard serves as a crucial tool for evaluating inventory performance and identifying cost-saving opportunities.

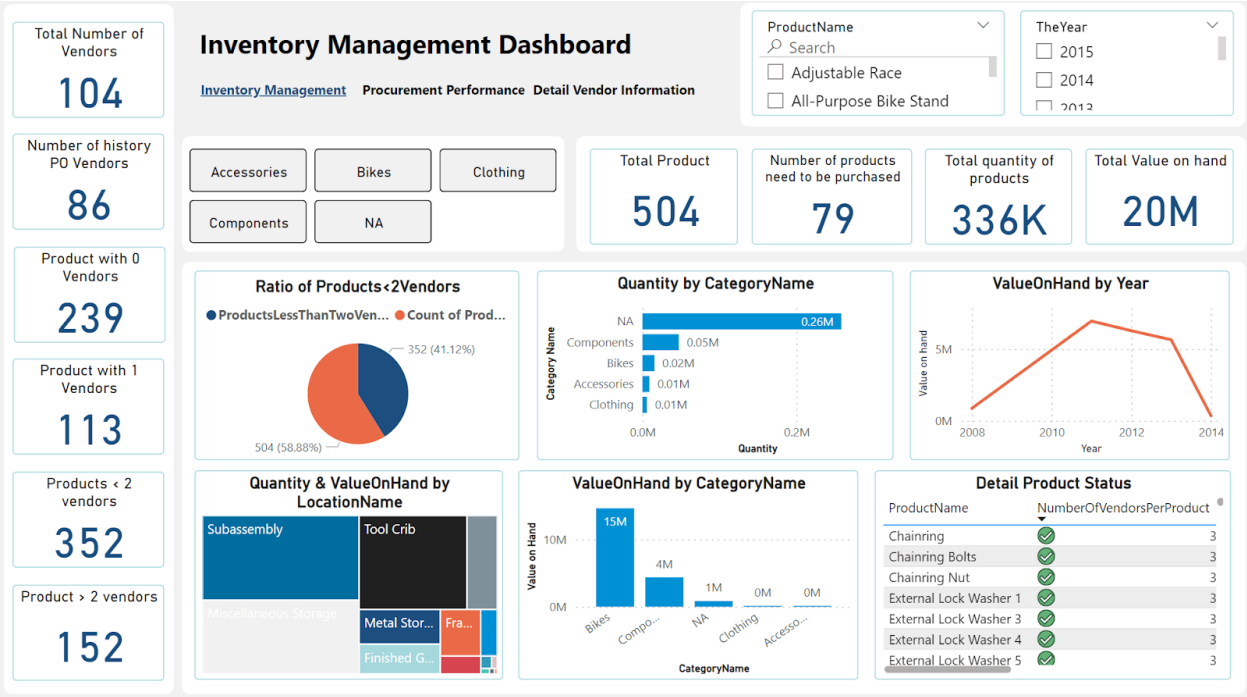


Figure 6.1. The Inventory Management Dashboard

Figure 6.2. Insight Cards

The BI dashboard in the image shows detailed information about supplier quantities, products, and purchase orders (POs). Data is presented as information cards, each card displaying a specific metric.

Key information cards:

- Total number of suppliers: 104
- Number of historical PO vendors: 86
- This product has 1 supplier: 113
- The product has 2 suppliers: 352
- This product has 0 suppliers: 239
- Products have more than 2 suppliers: 152

Based on the data shown, the following observations can be made:

- The number of suppliers is relatively large (104).
- The ratio of historical PO suppliers (86) to the total number of suppliers (104) is quite high, indicating that the business has long-term relationships with many suppliers.
- The number of products with 1 supplier (113) is greater than the number of products with 2 suppliers (352), showing that businesses tend to focus on cooperating with a few suppliers for products. specifically.
- There are a significant number of products with no suppliers (239) and products with more than 2 suppliers (152). This may be because the business is in the process of finding new suppliers for these products or because these products are manufactured in-house.

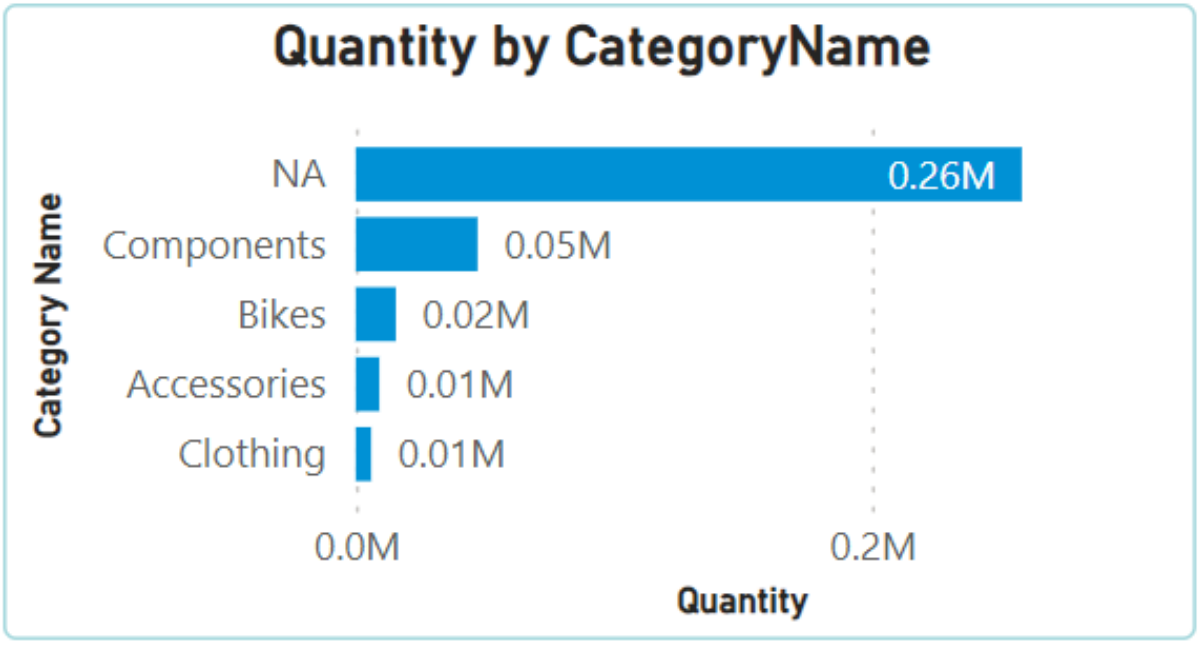


Figure 6.3. Bar chart "Sum of Quantity by CategoryName"

To have a clear view of the quantity of each product category, a bar chart should be used. Adventure Works Cycles had 4 categories that are components, bikes, accessories, clothing and others products that were not be set a category for. According to the figure, "Components" is the category that accounts for the most goods with the number of 47214, followed closely by "Bikes" with 15536 stocks.

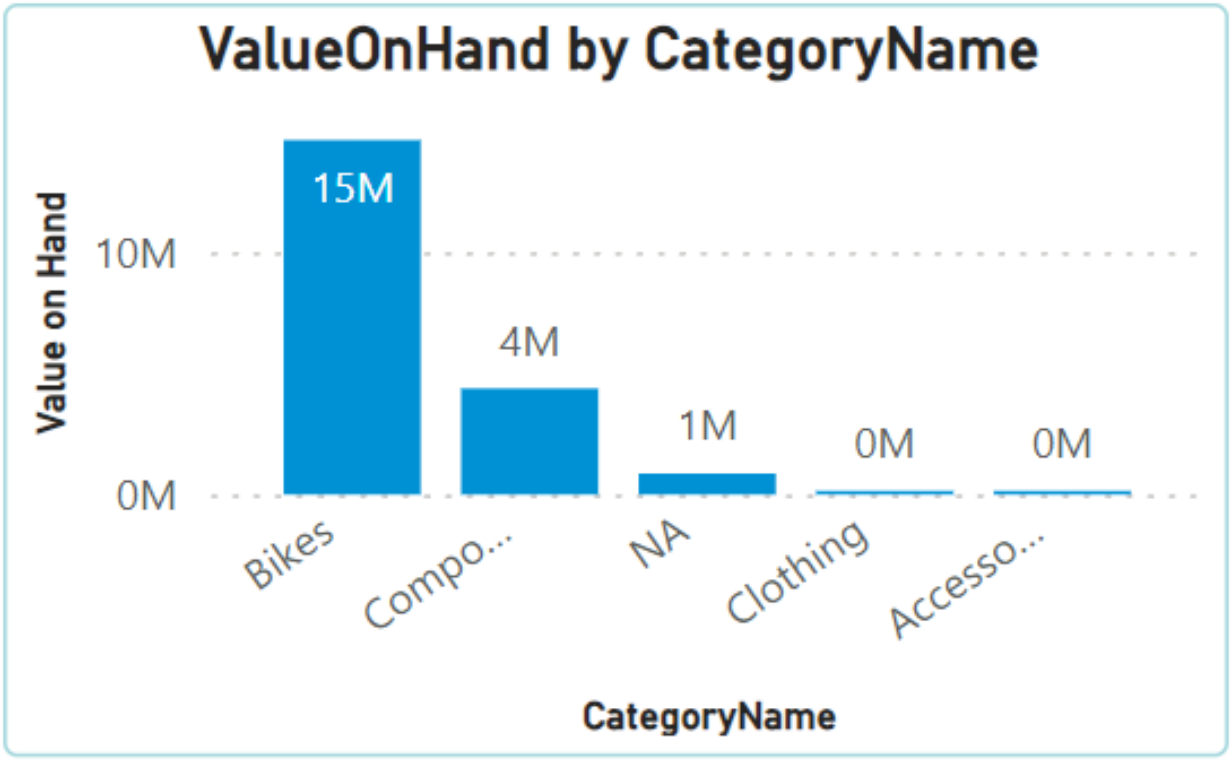


Figure 6.4. Bar chart "Sum of ValueOnHand by CategoryName"

In addition to the ValueOnHand by Category chart, we can see that the "Bikes", the main category of Adventure Works Cycles, accounts for the largest value of the company with more than 14.6 million dollars.

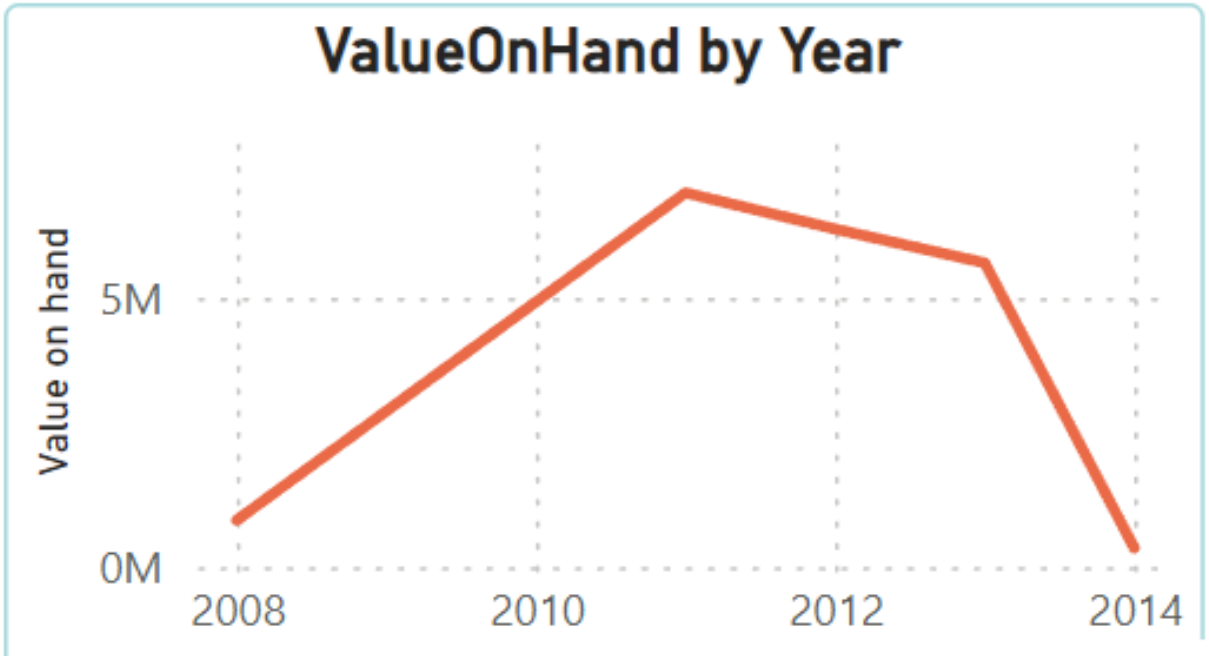


Figure 6.5. Line chart "ValueOnHand by Year"

The line chart shows the trend of value on hand of all product by the period of year from 2008 to 2014

ProductName	NumberOfVendorsPerProduct	NeedMakePurchase	SafetyStockLevel	Sum of Quantity	ReorderPoint	CategoryName
Chainring	3		1000	1684	750	NA
Chainring Bolts	3		1000	1136	750	NA
Chainring Nut	3		1000	1750	750	NA
External Lock Washer 1	3		1000	1046	750	NA
External Lock Washer 3	3		1000	981	750	NA
External Lock Washer 4	3		1000	972	750	NA
External Lock Washer 5	3		1000	952	750	NA
External Lock Washer 6	3		1000	1057	750	NA
External Lock Washer 7	3		1000	1672	750	NA
External Lock Washer 9	3		1000	963	750	NA
Flat Washer 1	3		1000	1103	750	NA
Flat Washer 2	3		1000	1673	750	NA
Flat Washer 3	3		1000	1702	750	NA
Flat Washer 4	3		1000	936	750	NA
Flat Washer 5	3		1000	911	750	NA
Flat Washer 6	3		1000	1093	750	NA
Flat Washer 7	3		1000	926	750	NA
Flat Washer 8	3		1000	896	750	NA
Flat Washer 9	3		1000	1072	750	NA
Hex Nut 16	3		1000	941	750	NA
Hex Nut 17	3		1000	926	750	NA
Hex Nut 5	3		1000	925	750	NA
Hex Nut 6	3		1000	1781	750	NA
Hex Nut 7	3		1000	1911	750	NA
Hex Nut 8	3		1000	770	750	NA
Hex Nut 9	3		1000	789	750	NA
HL Crankarm	3		500	797	375	NA
Thin-Jam Lock Nut 1	3		1000	1623	750	NA
Thin-Jam Lock Nut 10	3		1000	1628	750	NA

Figure 6.6. Make Purchase Action by Product

The table shows whether the product needs to be Purchased or not or considered. The “Green” symbol means “No need”, yellow means “Need to make purchase” and “Red” means “Urgent to make purchase”.

The BI dashboard in the image shows detailed information about inventory quantity and value by location name. Data is presented as column charts and tables.

Column chart:

Column chart displays inventory quantity by location name. The location with the highest inventory is "Subassembly", with 1,200 units. The location with the lowest inventory is "Finished Goods", with 200 units.

Board:

The table displays inventory quantities and values by location name. The location with the highest inventory value is "Miscellaneous Storage", with a value of $10,000. The location with the lowest inventory value is "Finished Goods", with a value of $2,000.

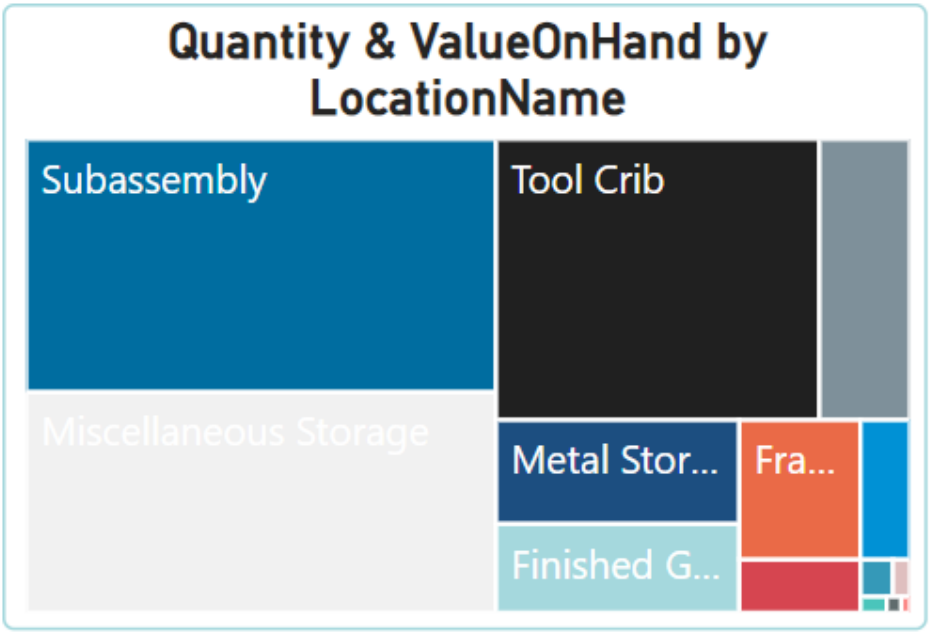


Figure 6.7. Quality & value on hand by LocationName

Based on the data shown, the following observations can be made:

- The highest amount of inventory is stored in the "Subassembly" location. This may be because this location stores the parts needed to produce the finished product.

- The highest inventory value is stored in the "Miscellaneous Storage" location. This may be because this location stores high-value materials.

- The amount of inventory in the "Finished Goods" position is relatively low. This may be because this location only stores finished products that are ready to ship.

Figure 6.8. Slicer with ProductName and Year

This is an area that helps managers easily filter the information they find necessary. Here, the dashboard provides them with a time filter in years to easily choose the appropriate time period as desired. In addition, the filter also allows filtering by product name, tick or search product name to help users easily select the product they want to see insights.

6.2.2. The Procurement Performance Dashboard

For purchasing team leaders, the Procurement Performance Analysis Dashboard offers a comprehensive overview of vendor relationships and procurement effectiveness. Through visualizations of average product prices, price fluctuations, and historical data, leaders can assess vendor performance and make informed decisions regarding price negotiations. Additionally, the dashboard provides insights into vendor ratings, purchase order success rates, and product quantity, enabling the identification of top-performing vendors and segmentation for future business activities. With customizable segmentation options, intuitive interfaces, and detailed price trend analysis, the dashboard facilitates strategic vendor selection, effective negotiation, and overall improvement in procurement processes.

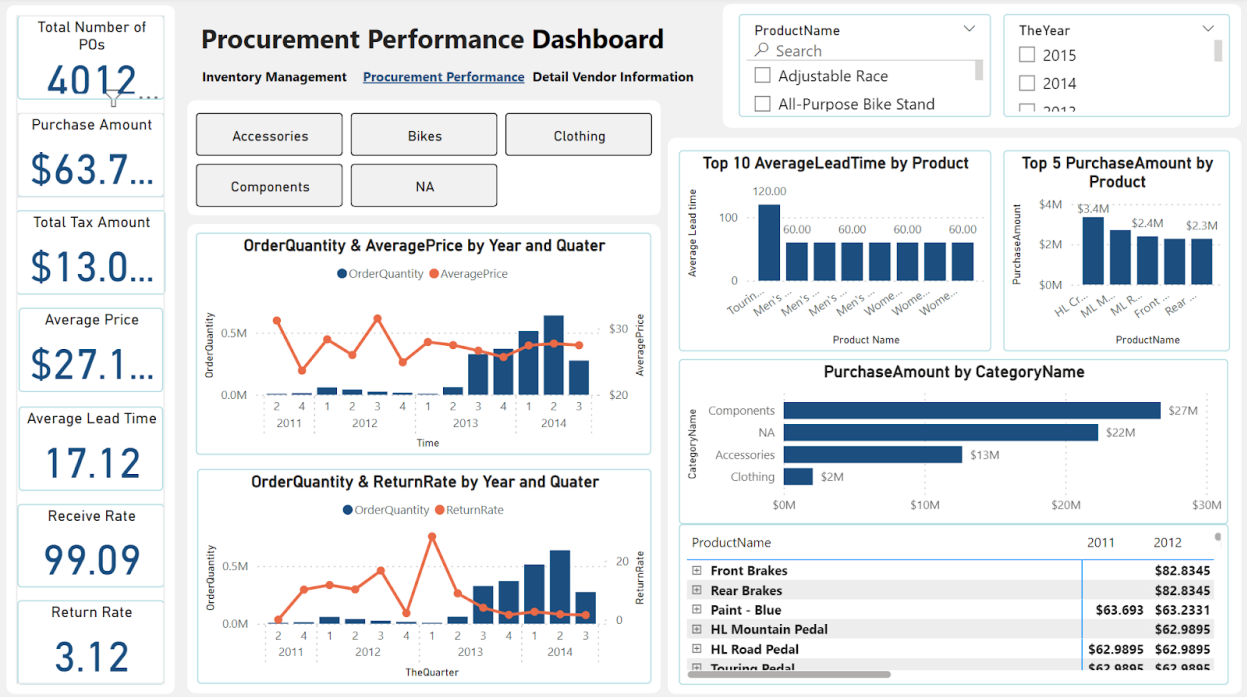


Figure 6.9. Dashboard Procurement Performance

6.2.2.1. Dashboard Name

Figure 6.10. Name of the dashboard "Procurement Performance"

The name of the chart is shown in the left corner of the dashboard and is shown with outstanding colors and large, clear font so that readers can easily understand the meaning of the dashboard right from looking at this dashboard name.

The name of the second dashboard is "The Procurement Performance Dashboard". The meaning of the dashboard is to allow managers to gain insights about the performance of the purchasing process, the trend of the average price of the product, the total cost spent on the purchasing process, etc. This helps managers promptly grasp the performance of the purchasing process as well as support them in choosing the appropriate price.

6.2.2.2. Visual insight cards

Figure 6.11. Visual insight cards

Visual insight cards are cards that provide general numerical insights about the performance of the purchasing process.

The first is the "Total Number of POs" card, this card shows the total number of Purchase Orders that have been made up to the present time of the data set. It shows that

the business has made 4012 Purchase Orders between 2011 and 2014. Showing that the business's purchasing performance is very high.

Second, “Purchase Amount” represents the total amount of money the business has spent to buy goods during this period. $63.8 million is the amount of money the business spent to buy goods for its production and sales. But it should be noted that this amount does not include the "Tax Amount", which will be mentioned next.

Third, “Total Tax Amount” shows the tax amount that the business spends on the Purchase Order. With $13.1 million, we see that the amount of tax spent on the total number of Purchase Orders is very large. It is necessary to make effective purchasing decisions to reduce taxes.

Fourth, “Average Price” represents the average number of days a process takes from placing an order to receiving the goods. 17.1 days is not a high number, but there still needs to be measures to reduce this index to increase production and sales efficiency.

Fifth, "Received Rate" shows the ratio of the number of goods received on an order compared to the number of goods when the business placed an order on Purchase Order. The rate of 99.09% is an acceptable rate, but there should still be measures to make the supplier deliver the exact quantity agreed upon by the business and supplier.

Finally, "Return Rate" is an index that shows the number of goods inventoried by the business but not of the desired quality so returned to the supplier. With 3.12%, the "Return Rate" is a quite small number, but it is necessary to consider which suppliers have a high "Return rate" to avoid wasting time and costs that reduce the efficiency of purchasing goods. .

6.2.2.3. Slicer Time(Year), Product Name

This is an area that helps managers easily filter the information they find necessary. Here, the dashboard provides them with a time filter in years to easily choose the appropriate time period as desired. In addition, the filter also allows filtering by product name, tick or search product name to help users easily select the product they want to see insights.

6.2.2.4. Slicer Product Category

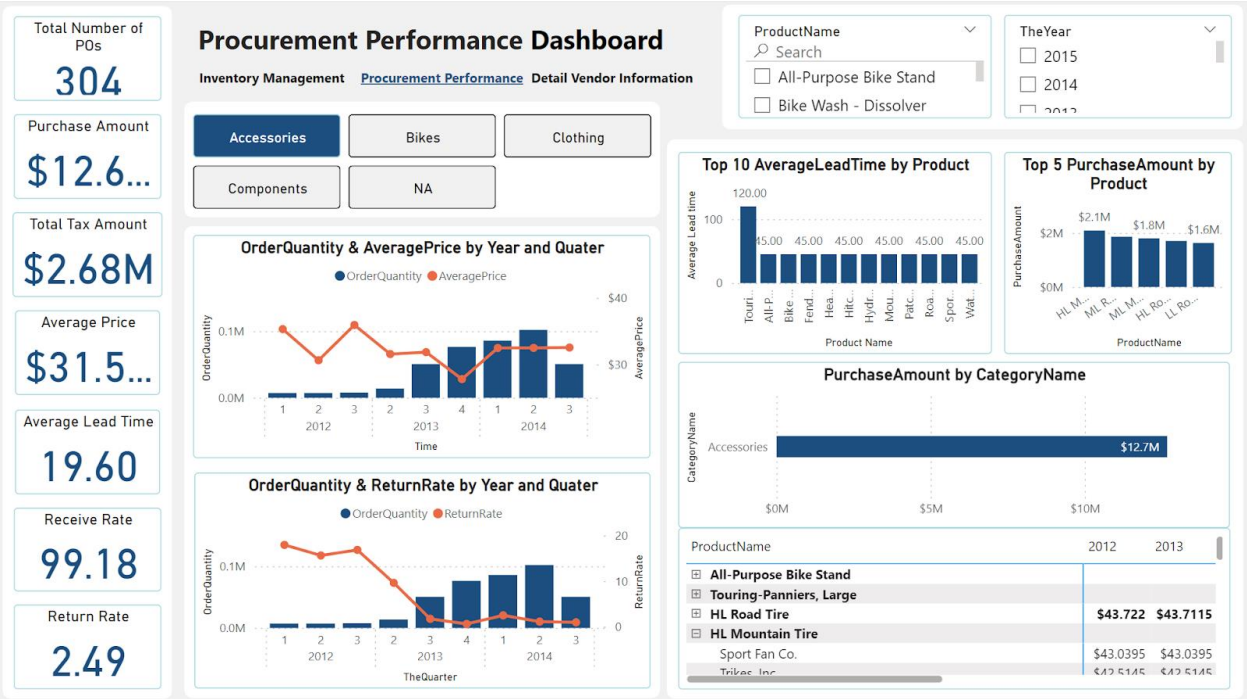


Figure 6.12. Dashboard Procurement Performance with Accessories filter

This is a filter about "Product Category", it helps users to immediately select large categories and get immediate insight about the Category. Large buttons help users make selections immediately without the hassle and difficulty of searching. This filter is placed like this to meet the need for immediate conversion and comparison with other categories.

6.2.2.5. Page navigation

Figure 6.13. Page Navigation bar

This is a filter about "Product Category", it helps users to immediately select large categories and get immediate insight about the Category. Large buttons help users make selections immediately without the hassle and difficulty of searching. This filter is placed like this to meet the need for immediate conversion and comparison with other categories.

6.2.2.6. Line and column chart shows "Order Quantity" and "Average Price" by "Time (Year and Quater)"

The chart shows the total number of Orders (column) over months and years, while the line chart shows price fluctuations over time. With this chart, users can easily know what the price fluctuations are over the years and what the current trend is in order to know the upcoming price trend of a product; In addition, the chart provides fluctuations in the number of Purchase Orders over the past period of time.

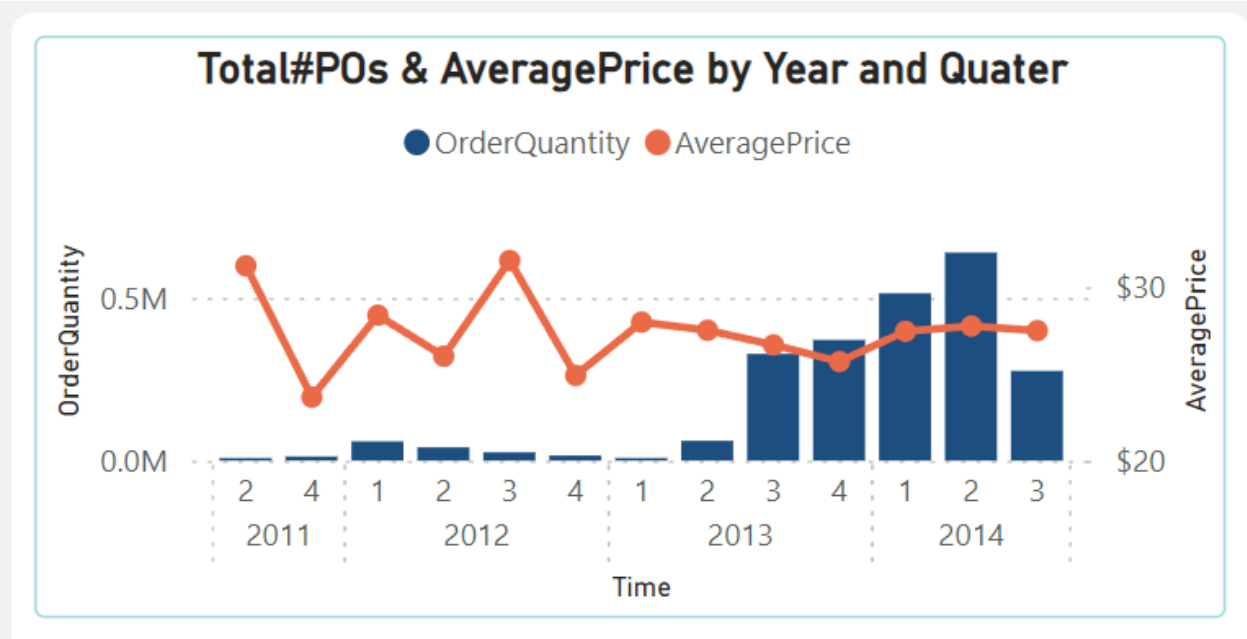


Figure 6.14. Total#POs & AveragePrice by Year and Quater

The image above with the product filter being the product "ML Mountain tire" shows the fluctuations in price and quantity of Purchase orders placed over the period from 2012 to 2014, in addition we can see the price range. The average price of the product tends to be stable from the end of the fourth quarter of 2013 to the end of the third quarter of 2014, fluctuating at a price of $36.

6. Line and column chart shows “Order Quantity” and “Return Rate” by “Time (Year and Quater)”

The chart "Order Quantity & Return Rate by Year and Quater" shows the fluctuations in the quantity of goods (Order quantity) over the past period of time of the business, with the line chart showing the fluctuations in the exchange rate. Rate of goods returned to suppliers over the years. The combined column and line chart also shows the correlation between "Return rate" and the quantity of goods (Order Quantity) over time.

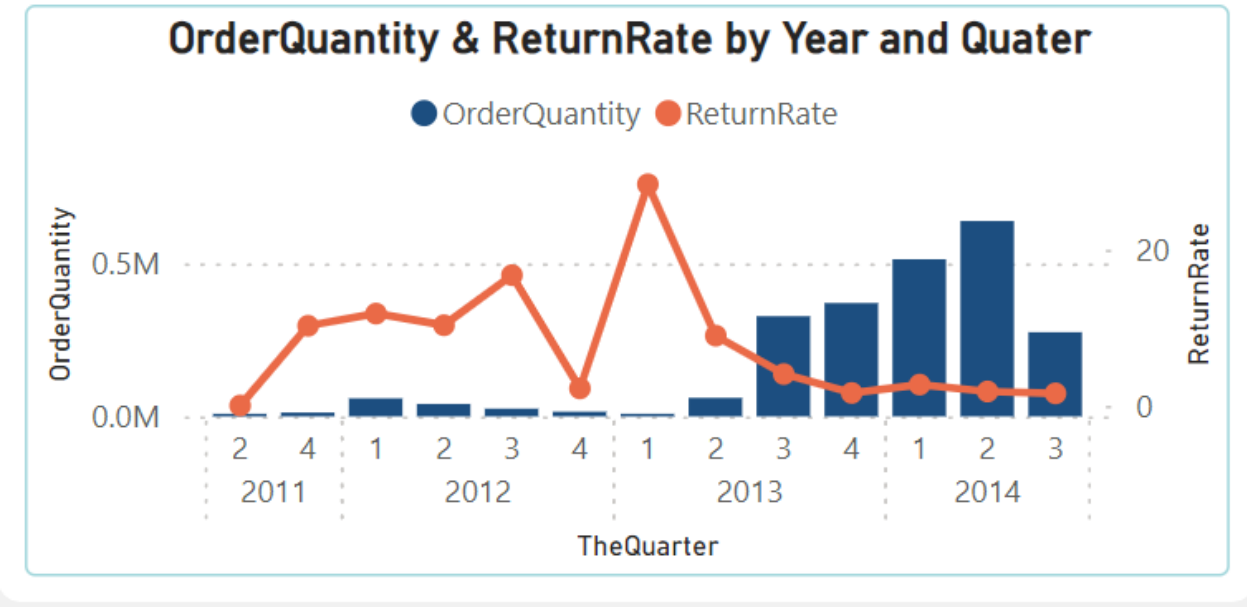


Figure 6.15. Order Quantity and ReturnRate by Year and Quater

The image above with the product filter is the product "ML Mountain tire" showing the fluctuations in price and Order quantity of this product over the period from 2012 to 2014. With the period from 2012 to the end of In the second quarter of 2014, the number of ordered goods for this product gradually increased, but then in the fourth quarter of 2014, the "Order quantity" of the product decreased to only half compared to the previous quarter. The "Return rate" of this product increased dramatically in the third quarter of 2012 with 51.26% but gradually decreased and by the third quarter of 2014 decreased to 2% and maintained until the end of the third quarter of

2014.

Figure 6.16. Order Quantity and ReturnRate by Year and Quater with Accessories Filter

Similar to the previous image, the image above shows two types of charts but is filtered by the value Category = Accessories.

Price fluctuations were quite erratic but stabilized in 2014 at a price of $32.55 dollars, showing that businesses have controlled the input prices of this type of goods very well.

During the period from 2012 to the end of the second quarter of 2014, the number of goods ordered from Accessories gradually increased, but then in the fourth quarter of 2014, the number of "Order quantity" of Accessories decreased to only half compared to the previous year. previous quarter.

6.2.2.7. "AveragePrice forecasting next 2 years"

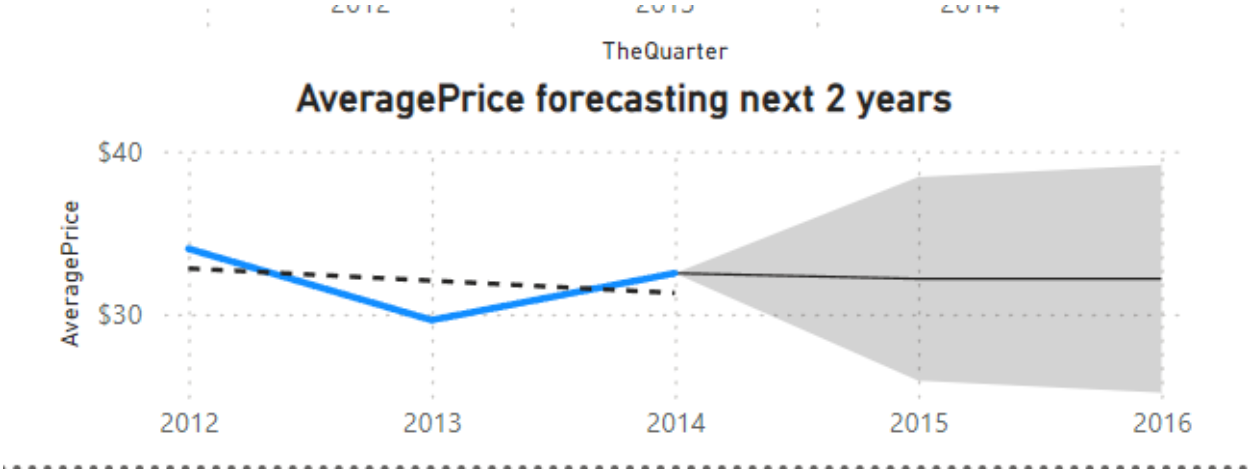


Figure 6.17. Average Price forecasting in the next 2 years

The line chart above shows the price trend of the product "ML Mountain tire" from 2012 to 2014. This price trend decreases and then increases but does not decrease or increase too much. In addition, the chart also provides price predictions for this product in the next 2 years. Helps managers plan to buy goods if their price is within a reasonable price range.

6.2.2.8. Bar chart shows "Total Purchase Amount" by "Category Name"

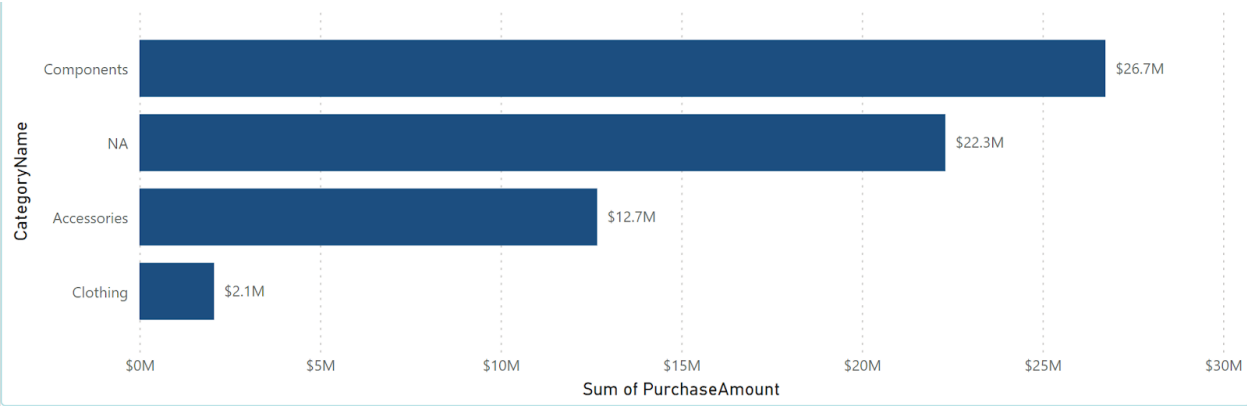


Figure 6.18. Total Purchase Amount by Category Name

The chart "Purchase Amount by Category Name" shows the total amount of money spent to buy goods for each Category throughout the years. The chart shows that Category "Component" accounts for the largest "Total Purchase Amount" with $26.7 million dollars, followed by "NA" with $22.3 million dollars, "Accessories" with $12.7 million dollars and lastly "Clothing" " vpows $2.1 million dollars.

6.2.2.9. Column chart shows Top 5 "Product" by the highest "Total Purchase Amount"

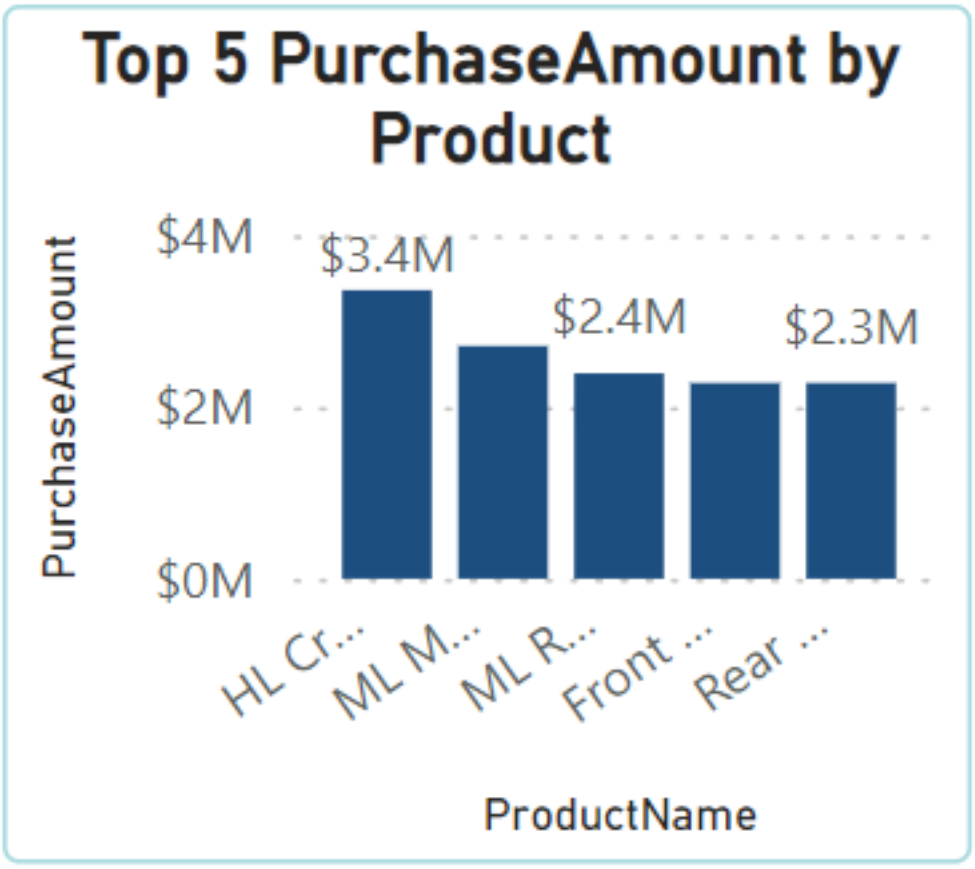


Figure 6.19. Top 5 Purchase Amount by Product

The column chart above shows the Top 5 products with the largest Purchase Amount. With "HL Cramkerm" at the top with $3.4 million dollars. The chart provides which products are frequently purchased so that you can have an appropriate purchasing plan to help optimize purchasing performance.

6.2.2.10. Column chart shows Top 10 "Product" by the highest "Average Lead time"

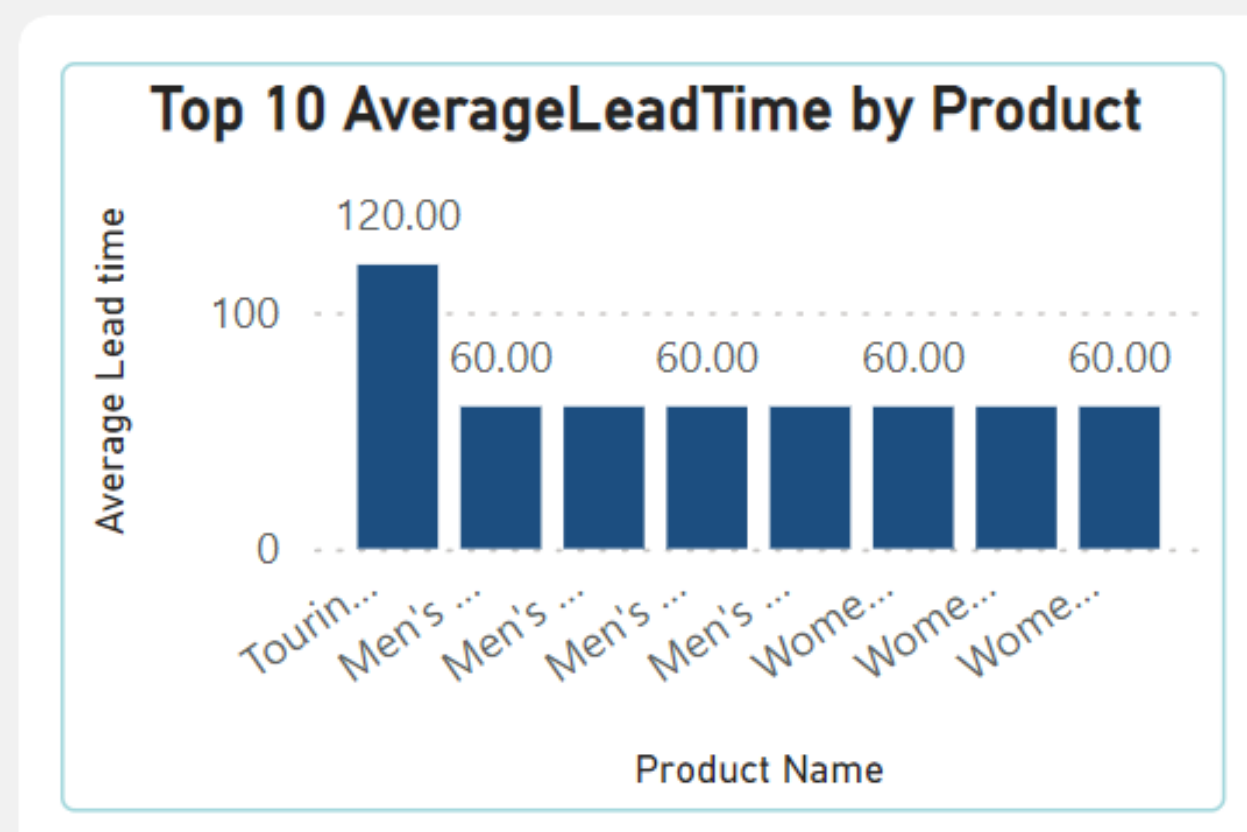


Figure 6.20. Top 10 Average Lead Time by Product

The chart above shows the top 10 products with the highest PO LeadTime. At the top is the product "Touring-panniers, Large" which has the largest POLeadTime with 120 days. Care must be taken when ordering this product because it may be late or wait too long, causing delays in the supply chain for the production and sales of goods, thereby reducing the overall performance of the business.

6.2.2.11. Table shows the "Average Price" of product and "Average Price" of product by Vendors in recent years.

ProductName	2011	2012	2013	2014	Total
Front Brakes		$82.8345	$82.8345	$82.8345	$82.8345
Rear Brakes		$82.8345	$82.8345	$82.8345	$82.8345
Paint - Blue	$63.693	$63.2331	$63.3098	$63.2994	$63.3037
HL Mountain Pedal		$62.9895	$62.9895	$62.9895	$62.9895
HL Road Pedal	$62.9895	$62.9895	$62.9895	$62.9895	$62.9895
Touring Pedal	$62.9895	$62.9895	$62.9895	$62.9895	$62.9895
All-Purpose Bike Stand				$59.46	$59.46
Paint - Red	$57.1515	$57.8193	$57.708	$57.723	$57.7168
Headset Ball Bearings	$57.0255	$57.0255	$57.0255	$57.0255	$57.0255
Paint - Silver	$53.2035	$51.7293	$51.975	$51.9418	$51.9555
Touring-Panniers, Large				$51.56	$51.56
Paint - Black	$50.2635	$50.5533	$50.505	$50.5115	$50.5088
Crown Race	$50.2635	$50.2635	$50.2635	$50.2635	$50.2635
Adjustable Race	$50.26	$50.2635	$50.2635	$50.2635	$50.2634
Thin-Jam Lock Nut 1	$49.6965	$49.8015	$49.811	$49.8043	$49.803
Thin-Jam Hex Nut 1		$49.7753	$49.791	$49.7753	$49.7818
Lock Washer 10		$49.6965	$49.6965	$49.6965	$49.6965
Paint - Yellow	$48.909	$49.2818	$49.2995	$49.2918	$49.2877
Cup-Shaped Race	$48.762	$48.762	$48.762	$48.762	$48.762
ML Mountain Pedal	$48.2895	$48.2895	$48.2895	$48.2895	$48.2895
ML Road Pedal		$48.2895	$48.2895	$48.2895	$48.2895
Thin-Jam Lock Nut 16		$47.7015	$47.7015	$47.7015	$47.7015
Thin-Jam Hex Nut 16		$47.6753	$47.691	$47.6753	$47.6818
Lower Head Race		$47.6805	$47.6805	$47.6805	$47.6805
Thin-Jam Lock Nut 9	$47.523	$47.628	$47.6411	$47.628	$47.6295
Lock Washer 8		$47.5965	$47.5965	$47.5965	$47.5965
Lock Washer 4		$47.523	$47.523	$47.523	$47.523
Flat Washer 2	$43.344	$47.579	$47.579	$47.579	$47.506
Thin-Jam Hex Nut 9	$45.12	$47.523	$47.6175	$47.6018	$47.5016
Chainring Bolts	$46.4205	$47.7606	$47.4473	$47.495	$47.4784
Keyed Washer		$47.4705	$47.4705	$47.4705	$47.4705

Figure 6.21. Table Average price of Product by time and by vendor

ProductName	2012	2013	2014	Total
HL Mountain Tire				
Sport Fan Co.	$43.0395	$43.0395	$43.0395	**$43.0395**
Trikes, Inc.	$42.5145	$42.5145	$42.5145	**$42.5145**

Figure 6.22. Filtered table with "ProductName" = "HL Mountain Tire"

Looking at the table we can see the average price of this product over the years from 2012 ($43.0395), 2013 ($43.0395), 2014 ($43.0395) and the average for each year is ($43.0395) from which we can know the product. This product has a current average price of ($43,0395) to assist with price negotiations and purchasing decisions.

In addition, the table also shows the average supply price from suppliers providing the product "HL Mountain Tire" which is "Sport Fan Co." with an average price of $43.0395 and supplier “Trikes, Inc” with a price of $42.5145 for this product. Based on this, it can be decided that choosing the supplier "Trikes, Inc" with a price of $42,5145 is more suitable and reduces unnecessary costs.

6.2.3. Detail Vendor Information

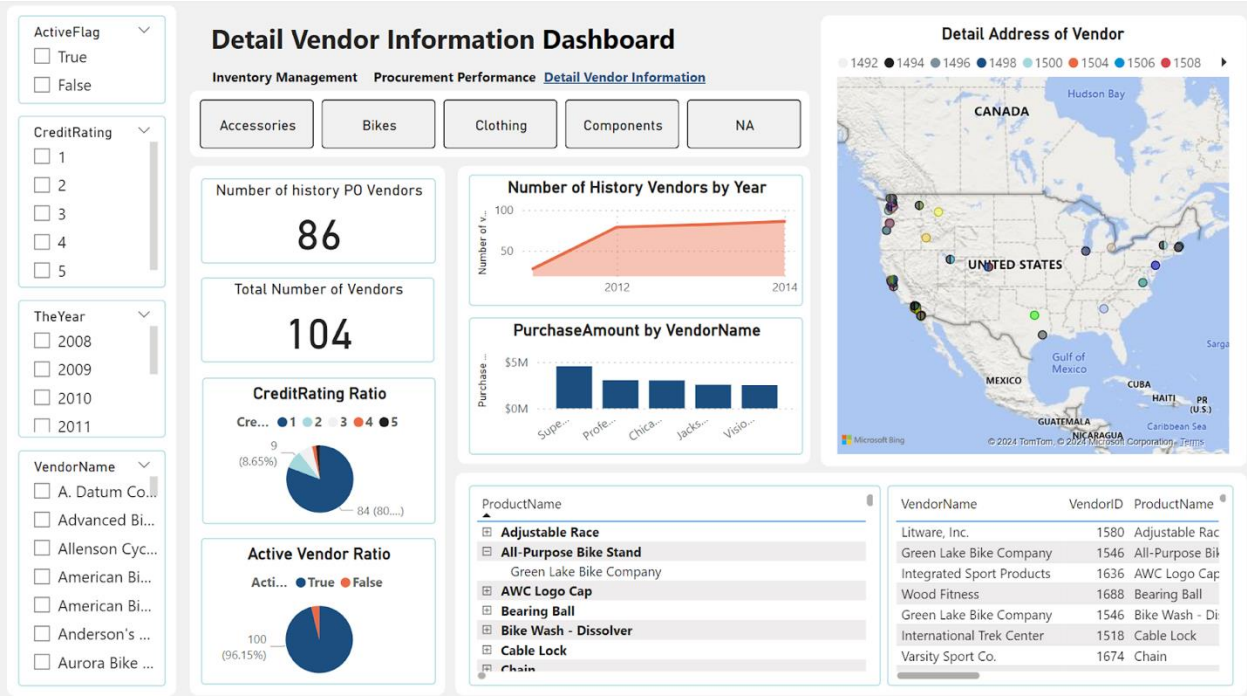


Figure 6.23. Detail Vendor Information Dashboard

This dashboard focuses on detailed information about vendors and purchase orders. It provides a visual summary of data related to vendors, such as the total number of vendors, their purchase history, credit ratings, and average product prices.

Title and category switcher: This section displays the name of the dashboard and buttons to switch the category of product which the vendor provides, including Accessories, Bikes, Clothing, Components and NA. Besides, there are also 3 buttons to switch between dashboards for convenience.

Filter and slicers: This section includes filters which enable users to filter data based on various criteria, and slicers which allow users to interactively slice and dice their data to

focus on specific subsets. Users can filter by ActiveFlag, CreditRating, Year and Vendor Name.

Basic Information of Vendor:

- Detail Address of Vendor: This table shows the vendor's detailed address through an interactive map. This map displays the distribution of vendors over geographic locations.
- Detail Vendor Information: This table displays the Vendor Name, their ID and their products' name.

Vendor Management Focus: This section shows visualizations that demonstrate detailed patterns and insights related to Vendors. It illustrates metrics such as "Number of history PO Vendors" and "Total Numer of Vendors".

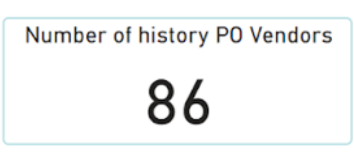


Figure 6.24. Number of History PO Vendors

Number of History PO Vendors: The number of historical purchase order (PO) vendors provide comprehensive information about partners that play the role as Vendor in previous purchasing order. The value is 86, which indicates that there is a variety in the historical PO vendors and vendor rating is important in order to choose the best Vendor from this provided list.

Figure 6.25. Total Number of Vendors

Total Number of Vendors: This metric shows the total number of vendors, which is 104. In compare with the Number of History PO Vendors 86, the ratio is 82 percent, which shows that most available vendors were chosen to make purchases with.

Figure 6.26. Number of History Vendors by Year

Number of History Vendors by Year: This visualization shows the number of historical vendors by year. From this perspective, there is a rapid increase in the number of vendors until 2012. From then, the number of vendors slightly increases by year.

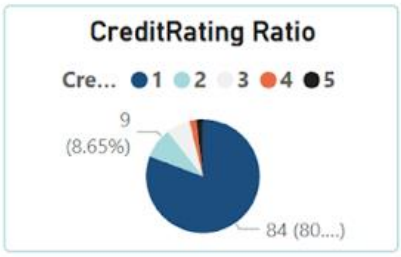


Figure 6.27. CreditRating Ratio

CreditRating Ratio: This chart shows credit rating information which range from 1 to 5, evaluating how valuable are vendors from historical purchases. It can easily be seen that 84% of vendors are rated with the highest score possible. However, the score of 4 only makes up a very small percent of the population. This opens up a possible factor which discriminates between those of rating 4 and those of rating 5.

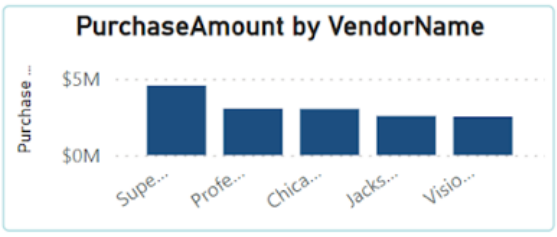


Figure 6.28. Purchase Amount by VendorName

Purchase Amount by VendorName: This bar chart shows the purchase amount by vendor name. The vendors with the most purchases are Superior Bicycles, following are Professional Athletic, Chicago City saddles, Jackson Authority and Vision Cycles inc. This insight helps identify the most familiar and credible to make purchases with.

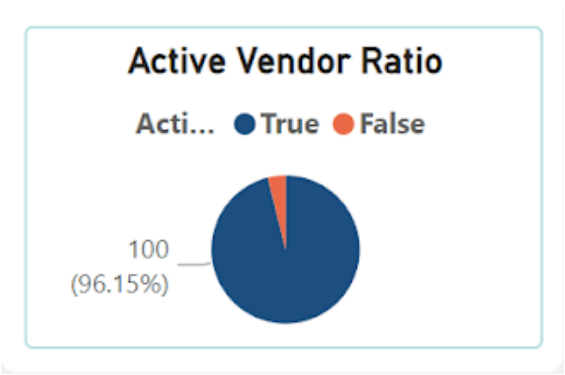


Figure 6.29. Active Vendor Ratio

Active Vendor Ratio: This section shows the ratio of active vendors versus inactive ones. It can be seen that active vendors makes up a absolute percentage of 96.15%, which state that most vendors are actives.

Average Product Price by Vendor and Time: This section shows the average product price by vendor and time.

- Adjustable Race: This product is priced at $50.26 in 2011, $50.2635 in 2012, and $50.2635 in 2013.

- All-Purpose Bike Stand: This product is sold by Green Lake Bike Company (Vendor ID 1546).

- AWC Logo Cap: This product is sold by Wood Fitness (Vendor ID 1688) and Integrated Sport Products (Vendor ID 1636).

- Bearing Ball: This product is priced at $41.916 in both 2011 and 2012. It is sold by Green Lake Bike Company (Vendor ID 1546).

- Bike Wash - Dissolver: This product is sold by International Trek Center (Vendor ID 1518).

- Cable Lock: This product is priced at $96.15. It is sold by Aurora Bike Center.

6.3. Evaluation and Discussion

The Inventory Management Dashboard serves as a comprehensive tool, offering real-time insights into crucial inventory metrics, empowering warehouse managers to monitor levels, turnover rates, and anticipate potential stock outages or excess stock situations. Through dynamic visualizations like bar charts and line graphs, stakeholders

gain visibility into inventory quantities, locations, and specific product details, facilitating proactive decision-making regarding purchase requisitions and stock optimization. Additionally, the dashboard's ability to track available products and safety stock levels enables managers to make informed choices, further enhancing inventory management efficiency. The customizable views, drill-down analysis features, and proactive alerts make the dashboard an indispensable resource for evaluating inventory performance and identifying opportunities for cost savings.

On the other hand, the Procurement Performance Analysis Dashboard provides purchasing team leaders with a comprehensive overview of vendor relationships and procurement effectiveness. Through visualizations of average product prices, historical data, and price fluctuations, leaders can assess vendor performance and negotiate prices strategically. Moreover, insights into vendor ratings, purchase order success rates, and product quantities enable the identification of top-performing vendors, aiding in segmentation for future business activities. The dashboard's customizable segmentation options, intuitive interfaces, and detailed price trend analysis facilitate strategic vendor selection, effective negotiation, and overall improvement in procurement processes.

The analysis of pricing trends and supplier credit ratings underscores the importance of adaptability and stability in procurement strategies. Efficient cost management strategies become imperative in optimizing expenditure and maximizing returns, given the substantial total purchase amount. The increasing number of purchase orders over time indicates growing procurement needs, emphasizing the necessity for proactive planning and resource allocation. Furthermore, the high purchase order success rate reflects streamlined procurement processes and effective supplier management. Insights into product demand and delivery performance reveal opportunities for optimization within the supply chain, while the low purchase order return rate underscores the effectiveness of quality control measures. Leveraging these insights enables informed decision-making, risk mitigation, and ultimately drives business efficiency and profitability.

Chapter 7: Conclusion and Future Works

7.1. Results

The Business Intelligence (BI) project has achieved significant milestones following the research and implementation process.

Firstly, we have conducted detailed identification and description of business requirements/KPIs and data. Using documentation techniques and analysis, we established a clear and specific set of performance metrics for the BI project, along with building appropriate Data Warehouse and Data Marts to provide a foundation for data integration and analysis.

Through the use of data integration techniques such as ETL and ELT on the SSIS platform, we were able to efficiently and reliably integrate data from various sources into the Data Warehouse. This facilitated multidimensional data analysis and deployment of KPIs on SSAS in a robust and accurate manner.

Lastly, by leveraging Power BI, we constructed visually appealing dashboards and reports, aiding users in understanding and uncovering insights from the data.

7.2. Limitation

Despite the significant achievements, the BI project encountered several limitations and challenges that necessitated further attention.

Firstly, data quality issues emerged during the ETL process, particularly concerning null values and data inconsistencies. Despite efforts to mitigate these issues, maintaining data integrity and completeness proved challenging, potentially compromising the reliability of our analyses.

Moreover, the complexity and incompleteness of certain datasets, notably in the Product tables where discrepancies between Product and Productinventory data arose, posed additional hurdles. Resolving these discrepancies was challenging and may have impacted the accuracy of our analyses and reporting.

Additionally, uncertainties surrounding Cube construction hindered our ability to identify and address underlying issues effectively. This lack of clarity impeded

optimization efforts and may have limited the effectiveness of multidimensional data analysis.

7.3. Future works

In looking ahead, there are several areas that require further attention and development. Firstly, it will be crucial to prioritize the enhancement of data quality assurance measures. This entails implementing more robust processes to identify and resolve issues related to data quality, particularly concerning null values and inconsistencies, to ensure the integrity and reliability of our analyses. Secondly, addressing discrepancies within the Product tables and improving data completeness across various datasets is essential. Intensified efforts to reconcile inconsistencies between Product and Product Inventory data are necessary to enhance the accuracy and completeness of our reporting. Thirdly, refining the Cube construction process to optimize performance and effectiveness in multidimensional data analysis is imperative. This involves clarifying uncertainties and streamlining Cube construction methodologies to uncover deeper insights and increase the overall value of our Business Intelligence project. Additionally, exploring advanced analytics techniques and integrating predictive modeling capabilities into our BI framework could further enhance our analytical capabilities and provide valuable foresight for decision-making. Lastly, providing ongoing training and support for users to maximize the utilization and understanding of BI tools and reports is crucial. Continuously soliciting feedback and iterating on our BI solutions based on user needs and preferences will ensure that our efforts remain aligned with organizational objectives and deliver tangible value.

Reference

[1] Statista. (n.d.). Bicycles - Vietnam | Statista Market Forecast. Retrieved May 02, 2024, from https://www.statista.com/outlook/mmo/bicycles/vietnam.

[2] Statista. (n.d.). Bicycles - Worldwide | Statista Market Forecast. Retrieved May 02, 2024, from https://www.statista.com/outlook/mmo/bicycles/worldwide.

[3] Rick Sherman, Business intelligence guidebook : from data integration to analytics, 1st Edition, Elsevier, 2015, ISBN 978-0-12-411461-6.