VIETNAM NATIONAL UNIVERSITY HO CHI MINH CITY

UNIVERSITY OF ECONOMICS AND LAW

PROJECT REPORT

SUBJECT: BUSINESS INTELLIGENCE AND DECISION SUPPORT SYSTEM

DATA ANALYSIS IN BUSINESS
WAREHOUSE MODULE COMBINE WITH PURCHASING

Lecturers:

PhD. Le Hoanh Su

MSc. Le Ba Thien

Group Members:

Tran Thi Thuy Loi	K224060840
To Nguyen Tuong Vy	K224060853
Nguyen Que Anh	K224060828
Nguyen Huynh Kim Suong	K224060846
Duong Nguyen Phuong Thao	K224060847

Ho Chi Minh City, August 2024

COMMITMENT

We hereby certify that the research findings presented in this essay are the sole product of our own group and are not the result of any cheating or plagiarism. The entire content of the report is presented based on personal viewpoints, accumulated knowledge, or selectively chosen from various legitimate and well-documented sources.

We take full responsibility and accountability as per regulations if any misconduct or cheating is detected.

ACKNOWLEDGEMENTS

We sincerely thank everyone who contributed to the success of this research. First, we appreciate Mr. Le Hoanh Su and Mr. Le Ba Thien for sharing their expertise in implementing this plan in practical work. Their invaluable guidance was highly valued throughout the project.

We also thank the academic advisor for their continuous support and encouragement. Their insightful suggestions significantly enhanced the quality of our research.

Finally, we express our gratitude to our friends for their various forms of contributions to this study. Their assistance was instrumental in our successful completion of the research.

Thank you all for your invaluable contributions.

DOCUMENT MATERIALS

FULL	BI_FINAL_GROUP 6 - Google Drive
SSIS	SSIS - Google Drive
SSAS	SSAS - Google Drive
SQL QUERY	SQL QUERY - Google Drive
DASHBOARD	DASHBOARD - Google Drive
CODE FORECAST	CODE - Google Drive
WORK ROSTER	Work Roster
SLIDE	SLIDE - Google Drive
DASHBOARD PUBLISHED	Power BI [expired license] Power BI [expired license]

TABLE OF CONTENT

LIST OF FIGURE

LIST OF TABLE

LIST OF ACRONYMS

BI	Business Intelligence
SSAS	SQL Server Analysis Services
KPIs	Key Performance Indicators
SARIMA	Seasonal Auto-Regressive Integrated Moving Average
WWI	Wide World Importers
ETL	Extract, Transform and Load
DW	Data Warehouse
IDE	Integrated Development Environment
MDX	Multidimensional Expressions
DAX	Data Analysis Expressions
SSMS	SQL Server Management Studio
XMLA	XML for Analysis
SCD	Slowly Change Dimension
DSS	Decision Support Systems
SAP	System Application Programming
HR	Human Resources

ABSTRACT

This project aimed to enhance the operational efficiency of Wide World Importers through the implementation of robust Business Intelligence (BI) solutions. The initiative commenced with an extensive data collection process, ensuring the aggregation of relevant and high-quality data from various sources within the organization. SQL Server Analysis Services (SSAS) were employed to conduct a thorough analysis of the processed data housed in the data warehouse. This analysis was pivotal in developing a suite of Key Performance Indicators (KPIs) designed to measure and evaluate various aspects of operational efficiency such as inventory levels, order fulfillment rates, and supplier performance. To facilitate real-time monitoring and strategic decision-making, interactive dashboards were created using Power BI. These dashboards offered a comprehensive and visual representation of the performance metrics, allowing stakeholders to quickly grasp the operational health of both the warehouse and the purchasing departments. A significant component of the project was the implementation of an SARIMA (Seasonal Auto-Regressive Integrated Moving Average) forecasting model. This statistical model was meticulously developed to predict future product demand trends with a high degree of accuracy. By anticipating demand fluctuations, the company could optimize stock levels, reduce holding costs, and minimize stock outs or overstock situations.

CHAPTER 1: OVERVIEW OF THESIS

1.1. The reason for choosing the topic

Business case: Wide World Importers (WWI) is a wholesale novelty goods importer and distributor operating from the San Francisco bay area.

As a wholesaler, WWI's customers are mostly companies who resell to individuals. WWI sells to retail customers across the United States including specialty stores, supermarkets, computing stores, tourist attraction shops, and some individuals. In 2023, the total retail sales volume of goods and consumer service revenue at current prices is estimated to reach 6,231.8 billion VND, an increase of 9.6% over the previous year. Total retail sales of goods and consumer service revenue in the fourth quarter of 2023 increased by 9.3% over the same period last year. Overall in 2023, total retail sales of goods and consumer service revenue increased by 9.6% compared to the previous year.

WWI also sells to other wholesalers via a network of agents who promote the products on WWI's behalf. While all of WWI's customers are currently based in the United States, the company is intending to push for expansion into other countries/regions.

WWI buys goods from suppliers including novelty and toy manufacturers, and other novelty wholesalers. They stock the goods in their WWI warehouse and reorder from suppliers as needed to fulfill customer orders. They also purchase large volumes of packaging materials, and sell these in smaller quantities as a convenience for the customers.

To ensure sustainable development and effective growth, the company needs to use a solution to analyze data and make accurate and timely business decision.

1.2. Topic goal

1.2.1. General goal

The objective of the thesis is to study the overview of business intelligence solutions, implementation processes and decision support systems. From that, we will apply this researched and learned knowledge to analyze data and build an intelligent reporting system to support management decision-making based on WWI's available database.

Enhance the efficiency and optimize the business processes of Wide World Importers (WWI) through the implementation of Business Intelligence (BI) solutions to provide timely, accurate, and valuable information for strategic and operational decision-making.

1.2.2. Specific goal

Optimize warehouse management by detecting relationships between products to reduce search time and improve operational efficiency.

Forecast inventory demand to adjust stock levels appropriately, avoiding shortages or excess inventory.

Optimize the replenishment process by arranging frequently purchased items close to each other to reduce movement time.

Enhance inventory accuracy by predicting changes in stock levels to minimize inventory losses.

Optimize purchasing strategy by analyzing purchase data to improve the ordering process and reduce costs.

Identify opportunities for supplier collaboration to improve purchasing conditions.

Improve supplier relationship management by analyzing purchasing patterns to make strategic decisions about maintaining or changing suppliers.

Analyze purchase order performance to improve future processes.

1.3. Subject and research scope of the project

1.3.1. Subject

The subject of our project is the Wide World Importers database. Specifically, we are focusing on the company's business data related to the performance of their purchasing and warehouse.

1.3.2. Research scope

- Time scope: 02/01/2013 - 31/05/2016
- Space scope: The data is collected from retail customers of Wide World Importers across the United States including specialty stores, supermarkets, computing stores, tourist attraction shops, and some individuals.
- The number of transactions: 10,796.

1.4. Tools used

- ***Visual Studio***

A powerful development tool used to complete the entire development cycle in one place. This comprehensive Integrated Development Environment (IDE) allows you to write, edit, debug, build code, and deploy applications. Visual Studio includes a compiler, code completion tools, source control, extensions, and many other features to enhance every stage of the software development process.

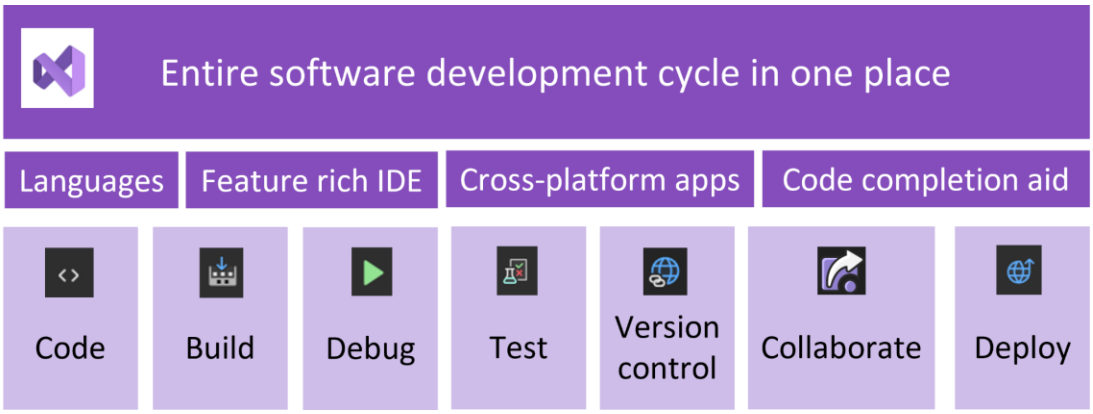

Figure 1.1. Software development cycle in Visual Studio

(Source: Microsoft Learn)

- ***SQL Server Management Studio (SSMS)***

An integrated environment for managing any SQL infrastructure. Use SSMS to access, configure, manage, administer, and develop all components of SQL Server, Azure SQL Database, Azure SQL Managed Instance, SQL Server on Azure VM, and Azure Synapse Analytics. SSMS offers a comprehensive utility that combines a large set of graphical tools with numerous rich script editors to provide access to SQL Server for developers and database administrators at all skill levels.

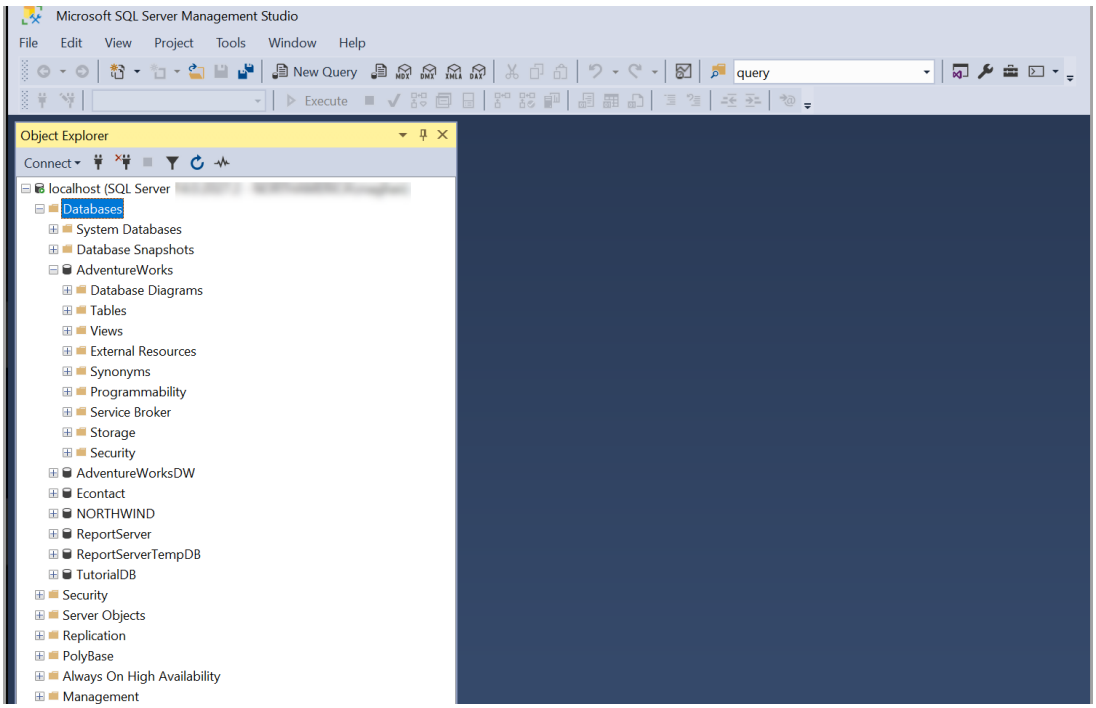

Figure 1.2. Screenshot of SQL Server Management Studio

(Source: Microsoft Learn)

- ***SQL Server Analysis Services (SSAS)***

An analytical data tool (VertiPaq) used in decision support and business analysis. It provides enterprise-level semantic data models for business reports and client applications like Power BI, Excel, Reporting Services reports, and other data visualization tools. Installed as an on-premises server or virtual machine, SQL Server Analysis Services supports tabular models at all compatibility levels (depending on the version), multidimensional models, data mining, and Power Pivot for SharePoint.

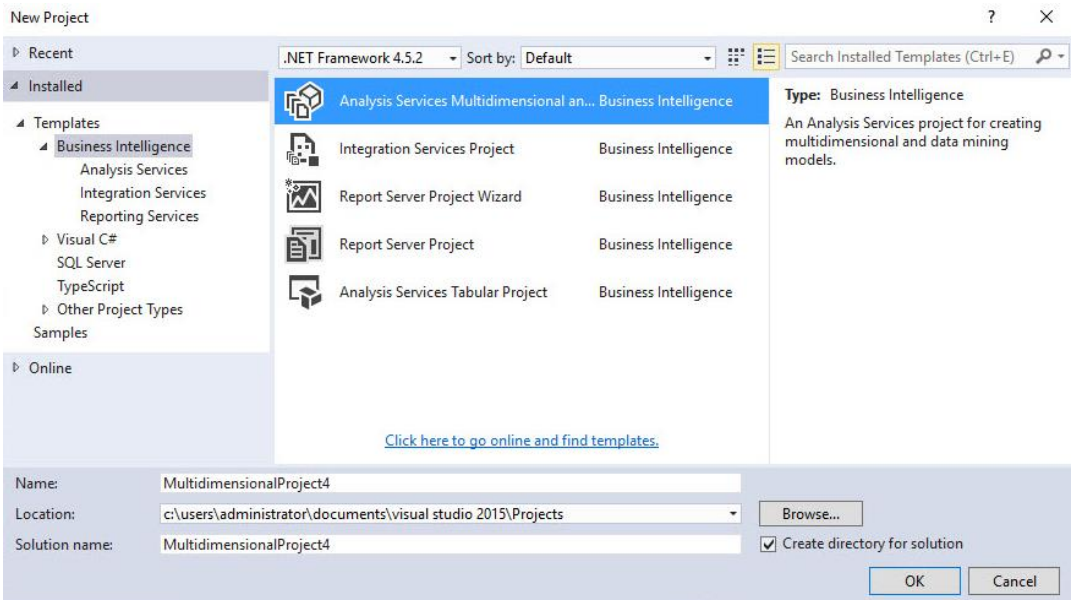

Figure 1.3. Screenshot of SQL Server Analysis Services

(Source: SQLShack)

- ***Power BI***

A collection of software services, apps, and connectors that transforms disparate data sources into interactive, visual insights. It connects to various data sources, including Excel and cloud or on-premises data warehouses, to visualize and share crucial information.

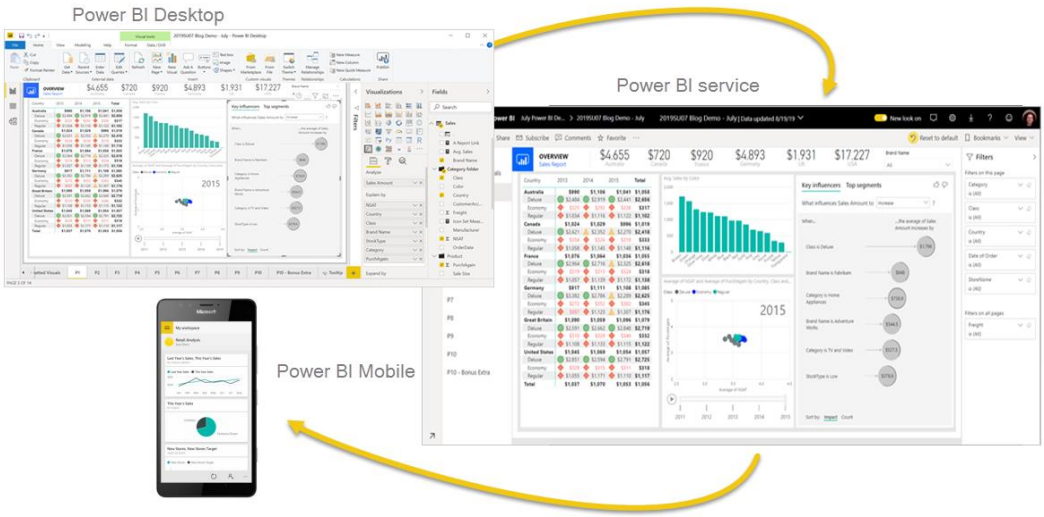

Figure 1.4. Integration of Power BI Desktop, Service, and Mobile

(Source: Microsoft Learn)

1.5. Research implications

Data-Driven Decision Making: Researching and implementing BI solutions in the Warehouse and Purchasing modules allows for more informed and data-driven decision-making processes. This leads to better strategic planning and operational efficiencies.

Operational Efficiency: BI tools help in identifying inefficiencies and areas for improvement within warehouse operations and purchasing processes. By addressing these inefficiencies, businesses can reduce costs and improve service levels.

Strategic Insights: The analysis of historical and real-time data provides strategic insights that can be used to enhance supply chain management, vendor relationships, and overall business operations.

Performance Monitoring: BI solutions enable continuous monitoring of key performance indicators (KPIs), allowing businesses to track performance against goals and make timely adjustments.

1.6. Structure of report

This project consists of 6 chapters:

- **Chapter 1: Overview of thesis**

In Chapter 1, we describe the business case of Wide World Importers and the need for business intelligence solutions to define the goals of the thesis. Besides that, we specify the subject, scope and list the tools and software used as well as the value brought for the project.

- **Chapter 2: Theoretical basis**

In Chapter 2, we present knowledge about Business Intelligence (BI), KPIs, ETL processes and techniques, Data Warehouses, and Data Marts. We also cover tools to build business intelligence solutions and decision support systems (BI and DSS) such as Pivot Table in Excel, Power BI, and Tableau. This foundational knowledge is necessary and important for implementing the project.

- **Chapter 3: Analysis of user requirement and data description**

In Chapter 3, we focus on analyzing user requirements and data description for the Wide World Importers of the Warehouse and Purchasing Departments. The project leverages the WWI dataset for BI analysis, addressing challenges related to data quality, complexity, integration and size for effective Business Intelligence.

- **Chapter 4: Data warehouse implementation**

Chapter 4 focuses on implementing a data warehouse using the Galaxy model, which integrates Master Data and Transaction Data into Fact and Dimension tables. This approach supports comprehensive business reporting and strategic decision-making by categorizing data and managing relationships between various data elements effectively.

- **Chapter 5: Results - Data analytics and visualization**

Chapter 5 utilizes SSAS for multidimensional data analysis, implementing MDX and KPIs for effective data evaluation. It also includes the creation of worksheets and dashboards using Power BI, providing a detailed view to support decision-making.

- **Chapter 6: Conclusion and future work**

In Chapter 6, we will summarize the results achieved and present limitations and directions for further development of the project.

CHAPTER 2: THEORETICAL BASIS

2.1. Overview of Business Intelligence

2.1.1. Introduce Business Intelligence model and solution

2.1.1.1. Business Intelligence definition

Business intelligence (BI) refers to computer-based techniques used in spotting, digging-out, and analyzing business data, such as sales revenue by products and/or departments, or by associated costs and incomes. [3]

BI technologies provide historical, current, and predictive views of business operations. Common functions of business intelligence technologies are reporting, online analytical processing, analytics, data mining, business performance management, benchmarking, text mining, and predictive analytics.

Currently, BI systems are increasingly applied in data-based business environments and have become a powerful assistant for companies possessing the available data assets to achieve competitive advantage.

2.1.1.2. How BI works

BI platforms traditionally rely on data warehouses for their baseline information. Then, BI presents the results to the user in the form of reports, charts and maps, which might be displayed through a dashboard. [4]

The steps taken in BI usually flow in this order:

- *Data sources:* Identify the data to be reviewed and analyzed, such as from a data warehouse or data lake, cloud, Hadoop, industry statistics, supply chain, CRM, inventory, pricing, sales, marketing or social media.
- *Data collection:* Gather and clean data from various sources. This data preparation might be manually gathering information in a spreadsheet or an automatic extract, transform and load (ETL) program.

- *Analysis:* Look for trends or unexpected results in the data. This might use data mining, data discovery or data modeling tools.
- *Visualization:* Create data visualizations, graphs and dashboards that use business intelligence tools such as Tableau, Cognos Analytics, Microsoft Excel or SAP. Ideally this visualization includes drill-down, drill-through, drill-up features to enable users to investigate various data levels.
- *Action plan:* Develop actionable insights based on analysis of historical data versus key performance indicators (KPIs). Actions might include more efficient processes, changes in marketing, fixing supply chain issues or adapting customer experience issues.

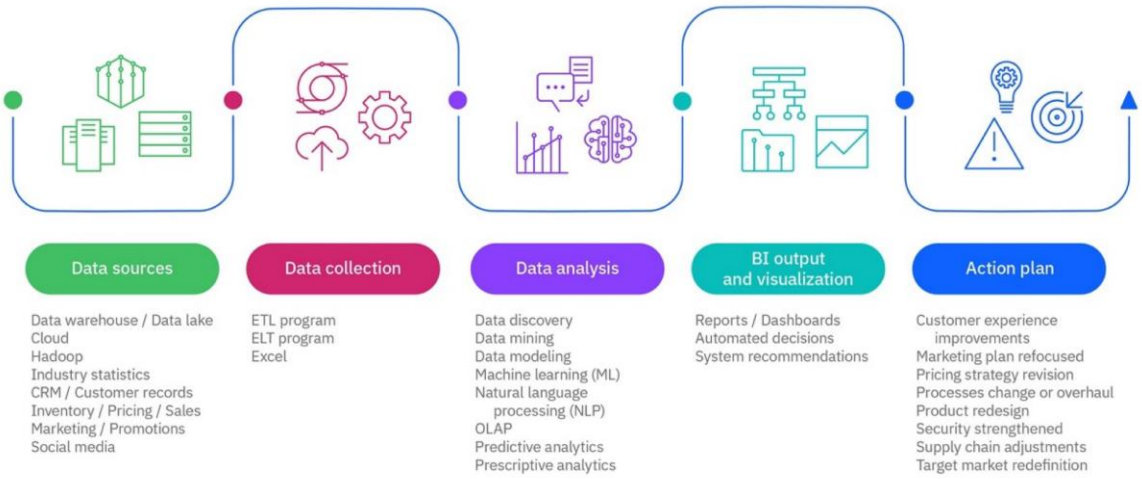

Figure 2.1. The steps taken in BI

2.1.2. The benefits of BI in the business

There are many reasons why enterprises decide to adopt Business intelligence (BI) solutions. Some of them could originate from its diverse support functions such as hiring, production, and marketing, etc.

[5] Below will be some primary benefits that companies can achieve after applying BI solutions:

- *Clearer reporting:* BI gives organizations the ability to ask questions in plain language and get answers they can understand. Dashboards can prioritize the most important insights, saving time for both data experts and nontechnical team members.
- *Consolidated data:* By providing an accurate picture of the business and market, BI provides an organization with the means to design a business strategy.
- *Create new efficiencies:* Organizations can monitor business operations against benchmarks and fix or make improvements on an ongoing basis—all fueled by data insights.
- *Deeper data insights:* BI helps organizations become more data-driven, to continually improve business performance, gain competitive advantage, and locate new customers and new opportunities.
- *Faster decision making:* As progress is monitored and analyzed digitally, better informed decisions can be made more quickly for faster adjustments in the marketplace.
- *Increase customer satisfaction:* When customer service staff have access to customer data and insights, they can provide requested information and resolve issues more quickly.
- *Increase employee satisfaction:* Self-help access to important business data can optimize workflows so that staff can do their jobs faster, with fewer added or repetitive steps.

In general, BI solutions help businesses obtain deeper data insights, make faster and more informed decisions, improve customer experience as well as employee satisfaction by providing an accurate, prompt and complete picture of the business and market.

2.1.3. The process of building a BI solution for businesses

Implementing a BI solution for enterprises will include the following steps [5]:

- *Defined business goals and objectives:* This involves identifying the key business problems business wants to solve, the data that need to analyze, and the desired insights from this analysis.
- Identifying the key performance indicators (KPIs) will be a useful method that helps measure success. By clearly defining your goals and objectives, BI projects can be ensured to be aligned with the overall business strategy and goals.
- *Gather requirements:* Determine the data sources that will need to analyze, the types of reports and visualizations needed to create, and the key stakeholders who will be involved in the project. This can help ensure that the business intelligence project is designed to meet the specific needs of the organization.
- *Select a BI platform:* When selecting a BI platform, it's important to consider factors such as cost, ease of use, scalability, and the specific features and capabilities that are most important to your organization. Right BI platform selection will help ensure that the business intelligence project is built on a solid foundation and we have the tools we need to analyze the data effectively.
- *Plan the project:* This involves identifying the specific tasks and activities that will be required to implement the project, as well as the timeline and budget for the project, which will ensure that you have the necessary resources and support to complete the project on time and within proper financial condition.
- *Develop the solution:* Building the solution involves designing and building the reports, dashboards, and other tools that will be used to analyze your data. It also involves integrating the BI platform with your existing systems and data sources, which helps ensure that the business intelligence project is effective, efficient, and meets the specific needs of the organization.
- *Launch the project and train users:* This involves rolling out the solution to the organization and providing training and support to users. This will help ensure that users are able to use the solution effectively and we are able to gain the insights we need from the data.
- *Set up support and continuous improvement processes:* This final step is to provide ongoing support and continuous improvement of the solution, which involves monitoring the performance of the solution, identifying areas for

improvement, and making changes and enhancements as needed. It also involves providing ongoing training and support to users to ensure that they are able to use the solution effectively and we are able to continue to gain valuable insights from the data.

2.2. ETL Process

2.2.1. Introduction of ETL

Extraction-Transformation-Loading (ETL) tools are pieces of software responsible for the extraction of data from several sources, their cleansing, customization and insertion into a data warehouse. Usually, these processes must be completed in a certain time window; thus, it is necessary to optimize their execution time [6].

2.2.2. Benefits of ETL

ETL is essential in data analytics and data warehousing systems because it simplifies and automates the process of integrating data from multiple sources, reducing effort and cost for data processing.

ETL allows businesses to consolidate data from multiple databases and other sources into a single repository with data that has been properly formatted and qualified in preparation for analysis. This unified data repository allows for simplified access for analysis and additional processing. It also provides a single source of truth, ensuring that all enterprise data is consistent and up-to-date.

ETL's main benefits are:

- *Quality:* ETL improves data quality by transforming data from different databases, applications, and systems to meet internal and external compliance requirements. This consolidation provides historical context since all relevant data is cataloged for discovery, thus removing blind spots in decision making.

- *Consistency:* ETL simplifies analysis by transforming data to follow a universal standard. ETL also improves the accuracy of calculations and predictions when all data is archived and searchable.
- *Speed:* ETL improves the speed of decision making by removing the need to query multiple data sources which all may have varying response times to build a complete picture.

2.2.3. ETL Process

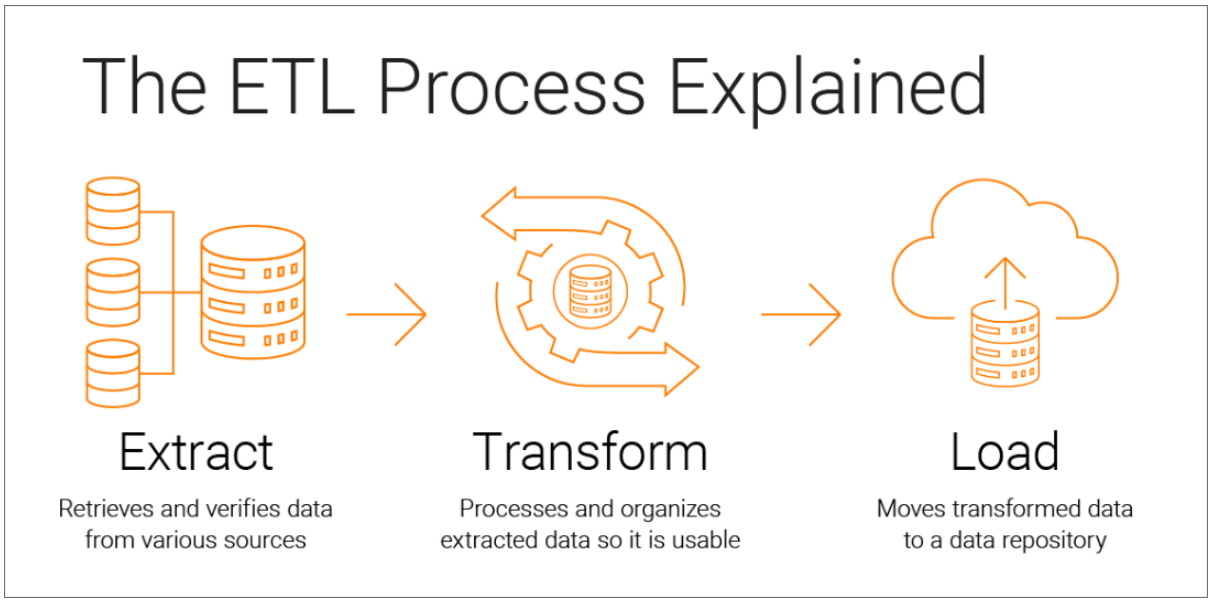

Figure 2.2. ETL Process

Source: Informatica

Step 1 - Extract: In this process, the first step is to extract the data from the target sources, which are usually heterogeneous, such as business systems, APIs, sensor data, marketing tools, and transaction databases, among others.

Step 2 - Transform: The next step is to transform the raw data, which has been extracted from the sources, into a format that can be accessed by different applications. The goal of this stage is to clean, map and transform data so that it is operationally useful. It involves several types of transformations that guarantee the quality and integrity of data.

Step 3 - Load: The Load process involves loading the transformed and refined data into the data warehouse or data analytics system. This process includes verifying the integrity and accuracy of the data and confirming that the data has been successfully loaded into the system.

2.3. Data warehouse and Data mart

2.3.1. Definitions of Data warehouse and Data mart

2.3.1.1. Data warehouse

Data warehouse is a type of data management system designed to support business intelligence (BI) activities, with a focus on data analysis to facilitate better business decision-making. Unlike other data systems, data warehouses are primarily used for running queries and performing analyses, often encompassing vast amounts of historical data. These warehouses are built by aggregating data from various sources, which is then transformed, processed, and stored in a centralized repository to support analysis and reporting needs.

By centralizing and consolidating extensive datasets from multiple origins, a data warehouse enhances an organization's ability to generate valuable insights from its data, thereby aiding in decision-making processes.

2.3.1.2. Data mart

Data Mart is a simple version of Data Warehouse, dedicated to a particular business line or department. Its purpose is to give users access to a targeted portion of the organization’s data for specific analysis or reporting needs. Unlike the comprehensive nature of a data warehouse that caters to the entire organization, data mart is customized to address the requirements of a specific group, such as sales, finance, or marketing, making it more efficient for targeted data analysis and reporting.

The main function of a Data Mart is to deliver pertinent information necessary for making critical decisions within a specific segment of the company. There are three main types of data marts:

- *Dependent Data mart:* A dependent data mart is constructed atop a central data warehouse. When analysis on a specific topic is needed, the data mart retrieves the relevant data from the data warehouse. An advantage of this approach is that most data management tasks are handled in the central repository, reducing the need for extensive technical expertise at the data mart level. Additionally, crucial aspects such as technology, data management, and storage (including backups) are managed centrally.
- *Independent Data mart:* This type can operate without supporting information from Data Warehouse. Companies often establish independent data marts for specific purposes, with data stored both internally and externally. An advantage of this model is that individual business units can manage the data mart that best fits their needs.
- *Hybrid Data mart:* These types will integrate data in a Data Warehouse existing from external sources. Combined data marts are highly flexible and can handle large volumes of data. A key consideration for this model is the human resources needed to operate it. Technical administrative expertise is required both at the central data warehouse and at the data mart level.

2.3.2. Galaxy Schema

In a data warehouse, a galaxy schema is a type of star schema where dimensions are organized into multiple levels. The lowest level, which is the center of the galaxy, contains the fact table. The next level comprises dimensions directly linked to the fact table, while the outermost level includes dimensions not directly connected to the fact table.

Characteristics of the Galaxy Schema:

- Galaxy Schema is multidimensional acting as a strong design consideration for complex database systems.
- Galaxy Schema reduces redundancy to near zero redundancy as a result of normalization.

- Galaxy Schema is known for high data quality and accuracy and lends to effective reporting and analytics.

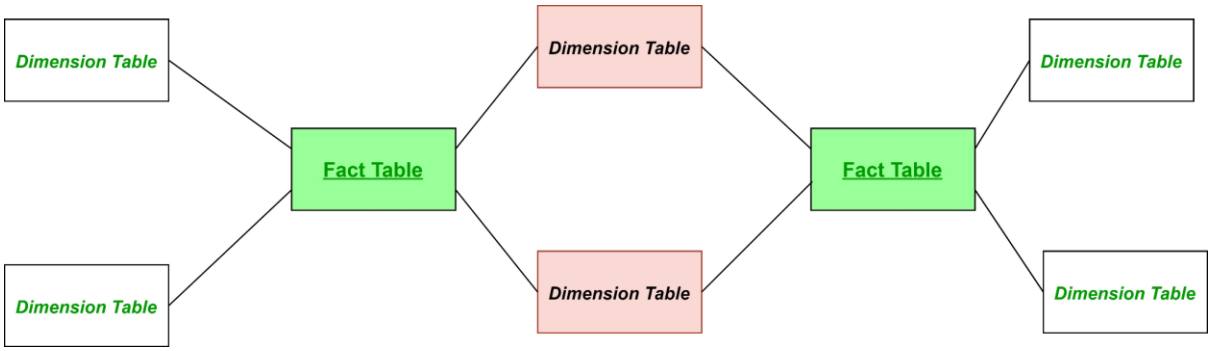

Figure 2.3. Galaxy Schema

Source: GeeksforGeeks

2.3.3. The cases apply DW and DM

Data Warehouse

- Personnel responsible for decision-making based on data volume.
- Users involved in complex and customized processes to gather information from various sources.
- Individuals needing simple technology to access data.
- Those seeking a systematic approach for decision-making.
- People aiming to enhance performance with large datasets for reports, networks, or charts.
- Individuals looking to uncover "hidden patterns" in data and analyze group trends.

Sometimes, having all your data in one place is more beneficial to the bottom line. These use cases illustrate when a data warehouse should be used instead of a data mart.

- Systems Integration: Companies aiming to enhance their systems and business processes can use security devices, smartwatches, and other data-driven technologies to predict future trends and patterns based on historical data. This aids in generating metrics and reports that enable teams to swiftly adapt to changes.

- Centralized Data for Impact or Profit: A health insurance company reporting on profitability benefits from a centralized data warehouse to collect information from sales, marketing, finance, and operations. Data warehouses enable the creation of dashboards for visualizing this data.
- Company-Wide Performance Evaluations: A retail company can utilize data warehouses to assess team performance across the entire organization. Business intelligence analysts can develop dashboards and reports based on customer value and usage patterns to evaluate marketing, sales, and customer service teams.

Data Mart

- Employees in business units or departments require data analysis and reporting to support their tasks.

Some common use cases:

- The marketing team seeks demographic data on customers who bought a beauty product in the summer of this year to enhance brand positioning for the next year. Since financial and operational data are not required for this analysis, a data mart is more appropriate.
- Sales representatives at a retail company can utilize a data mart to consolidate month-over-month and year-over-year data into a single dashboard for tracking their performance.
- In a shipping department, a data mart can monitor the complete duration and expenses from order placement to customer delivery. This shipping data mart can collaborate with the sales department data mart to assess overall shipping efficiency and costs.

2.3.4. Advantages and disadvantages of Data warehouse

Advantages

- Enhance accessibility for end users to diverse business data types.
- Enhance data consistency.

- Additional data on data.
- Potential for cost savings and improved productivity.
- Centralize integration of relevant data from disparate sources.
- Establish computing infrastructure capable of accommodating system and business structure changes.
- Enable end users to execute complex queries or generate reports without impacting system performance.

Disadvantages

- Less suitable for handling unstructured data.
- Prone to becoming outdated over time.
- Challenges in modifying data types, data range, data source scheme, index and query.
- It is difficult to access to users.
- Users may face difficulty in developing diverse business rules within the Data Warehouse.
- Significant investment required for user training and Data Warehouse deployment.

2.4.KPIs

2.4.1. Definition of KPIs

KPI stands for Key Performance Indicator, is a quantitative measure of performance over time for an organization or individual in a specific business or project. KPI provides clear goals and defines the success of business activities, helping businesses set targets and milestones to evaluate progress. Additionally, KPI provides insights that help everyone in the organization make better decisions. From finance and human resources to marketing and sales, these key performance indicators assist all areas of a business in advancing at a strategic level. KPIs are typically established to measure the effectiveness of activities such as revenue, profit, product or service quality, costs, and completion time.

2.4.2. Types of KPIs

2.4.2.1. Financial Metrics and KPIs

Financial performance indicators often focus on revenue and profit margins. Net profit, the most accurate measure of profitability, shows the amount of revenue remaining as profit after accounting for all expenses, taxes, and interest payments during a given period. Examples of financial KPIs include:

- *Liquidity ratio*: Measures a company's ability to manage short-term debt obligations based on its available short-term assets.
- *Profitability ratio*: Measures how efficiently a company generates revenue while keeping costs low.
- *Solvency ratio*: Measures a company's long-term financial health by assessing its ability to meet long-term debt obligations.
- *Turnover ratio*: Measures the speed at which a company can complete a specific task.

2.4.2.2. Human Resource and Staffing KPIs

Companies benefit from analyzing specific KPIs for their employees. From revenue to retention to satisfaction, companies gather valuable insights about their workforce. Examples of HR-related KPIs include:

- *Overtime Hours*: Monitoring overtime to assess fatigue risks and workforce adequacy.
- *Employee Satisfaction*: Typically involves company-wide surveys to evaluate employee perceptions across various aspects of the company.
- *Turnover Rate*: Measures the frequency and speed at which employees leave their positions.
- *Employee Salary*: Used to evaluate and manage employee compensation and its impact on organizational performance.

2.4.2.3. Process Performance Metrics and KPI

Process metrics are essential for measuring and monitoring operational performance across an organization. These KPIs analyze how tasks are executed and identify issues related to process or efficiency. Examples of process performance metrics include:

- *Production Efficiency*: Typically measured by production time for each stage divided by total processing time.
- *Productivity*: The number of units produced divided by production time per unit, measuring the speed of the production process.
- *Error Rate*: The total number of errors divided by the total number of units produced.
- *Quality Rate*: Focuses on the proportion of positive outcomes produced rather than negative ones.

2.4.2.4. Marketing KPIs

Marketing KPIs aim to better understand the effectiveness of marketing campaigns and advertisements. These metrics typically measure conversion rates, which indicate how often potential customers take specific actions in response to a particular marketing medium. Examples of marketing KPIs include:

- *Social media traffic*: Monitors views, follows, likes, shares, interactions, and other measurable engagements on company social media profiles.
- *Conversion rate on call-to-action content*: Focuses on promotional programs that urge customers to take specific actions.
- *Click-through rate*: Measures how often specific email distributions receive clicks from recipients.
- *Public relations*: measure the effectiveness and impact of public relations activities and campaigns.

2.4.2.5. Sales KPIs

The ultimate goal of a company is to generate revenue through sales. While revenue is typically measured using financial KPIs, sales KPIs take a more detailed approach by leveraging non-financial data to better understand the sales process. Examples of sales KPIs include:

- *Customer Lifetime Value (CLV)*: This represents the total amount a customer is expected to spend on your products throughout the business relationship.
- *Customer Acquisition Cost (CAC)*: This shows the total sales and marketing expenses needed to acquire a new customer.
- *Average Conversion Time*: Measures the time from initial contact with a potential customer to signing a business contract.
- *KPIs for salespeople*: measure salespeople performance, productivity, and effectiveness in achieving sales targets and objectives.

2.4.3. Advantages and Disadvantages of KPIs in Business

2.4.3.1. Advantages of KPIs

Key Performance Indicators (KPIs) offer significant benefits, including identifying and closing learning gaps. If goals aren't met, KPIs can reveal areas where employees need more training. For example, aiming to increase sales conversions by 20% but seeing no improvement after three months indicates a need for additional sales training.

KPIs also empower employees to take action by providing clear, measurable goals like “send 30 sales emails daily” or “follow up with leads within 1 hour.” These specific targets guide employees and drive performance towards broader objectives.

Additionally, KPIs allow for the measurement of outcomes and results. They help assess whether training and strategies are effective. If a goal is achieved, KPIs help determine what contributed to the success. If not, they help identify the reasons, such as overly ambitious targets or gaps in employee knowledge. Overall, KPIs are essential for evaluating and improving organizational performance.

2.4.3.2. Disadvantages of KPIs

While KPIs are essential tools for measuring and improving organizational performance, they come with certain drawbacks. One significant disadvantage is that KPIs require time to provide actionable information. When an organization invests in a new employee training program, the impact of this training may not be immediately visible. Depending on the specific goals and the frequency of tracking, it may take several months before the KPIs reflect any noticeable changes in team performance. This delay can be frustrating, but it's important to remember the adage, "A watched pot doesn't boil." Patience is crucial; organizations must allow sufficient time for KPIs to yield meaningful insights.

Another challenge with KPIs is their steep learning curve. For those unfamiliar with identifying and implementing KPIs, starting with too many at once can be overwhelming. Just as an overambitious workout after a long period of inactivity can lead to exhaustion and discouragement, trying to manage numerous KPIs simultaneously can result in data overload and confusion. This can cause frustration and a sense of failure for managers. Therefore, it's advisable to begin with a few key KPIs, gradually increasing the number as familiarity and confidence grow. By taking a measured approach, organizations can effectively leverage KPIs without falling into the trap of overcomplication.

2.5. Data analysis and visualization

2.5.1. Theory and methods in data analysis

Data analysis is simply the process of transforming collected data into meaningful information. Various techniques, such as modeling, are used to identify trends, relationships, and thereby draw conclusions to support decision-making processes. However, data needs to be prepared before it can be utilized in the data analysis process.

The data analysis can be categorized into the following six main methods (Taherdoost, 2021):

- **Descriptive:** Recognized as the first type of data analysis, it is known as the method with the least amount of effort. Thus, it can be used for large volumes of data. Here the data is used to perform a data set (Start, 2006).
- **Exploratory:** This method is used to explore the unknown relationships and discover new connections, and define future studies or questions (Start, 2006).
- **Inferential:** Inferential analyzing method uses a small sample to conclude a bigger population. It means, data from a subject sample of the world is used to test a general theory about its nature. The types of data sets that can be used in this method are observational, retrospective data set, and cross sectional time study (Bhattacherjee, 2012).
- **Predictive:** Predictive analysis utilizes historical and current facts to reach future predictions. It can also use data from a subject to predict the values of another subject. There are different predictive models; however, a simple model with more data can work better in general. Therefore, the prediction data set and also the determination of the measuring variables are important aspects to consider (MacGregor, 2013).
- **Explanatory:** This analyzing method is used to determine the consequences happening to one variable when changing another one using randomized trial data sets (Bhattacherjee, 2012).
- **Mechanistic:** This method needs the most effort to determine the exact changes in the variables which can lead to changes in other ones using randomized trial data sets. It can be also concluded that mechanistic analysis is hardly inferable. Thus, when you need high precision in your result and you should minimize your errors, for example in the engineering and physical sciences, it can be a choice.

2.5.2. Visualization

2.5.2.1. Definition

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data. Additionally,

it provides an excellent way for employees or business owners to present data to non-technical audiences without confusion.

2.5.2.2. Types of data visualization

Numerous taxonomies exist for classifying data visualization methods. Focusing on the six-category: capture a trend, visualize relationships, part-to-whole charts, visualize a single value, capture distributions and visualize a flow.

Capture a Trend: These data visualizations allow people to display a trend over time.

- Line chart: The most straightforward way to capture how a numeric variable is changing over time.
- Multi-line chart: Captures multiple numeric variables over time. It can include multiple axes allowing comparison of different units and scale ranges.
- Area chart: Shows how a numeric value progresses by shading the area between line and the x-axis.
- Stacked area chart: Most commonly used variation of area charts, the best use is to track the breakdown of a numeric value by subgroups.
- Spline chart: Smoothened version of a line chart. It differs in that data points are connected with smooth curves to account for missing values, as opposed to straight lines.

Visualize Relationships: Display relationships between data points.

- Bar chart: One of the easiest charts to read which helps in quick comparison of categorical data. One axis contains categories and the other axis represents values.
- Column chart: Also known as a vertical bar chart, where the categories are placed on the x-axis. These are preferred over bar charts for short labels, date ranges, or negatives in values.

- Scatter plot: Most commonly used chart when observing the relationship between two variables. It is especially useful for quickly surfacing potential correlations between data points.
- Connected scatterplot: A hybrid between a scatter plot and a line plot, the scatter dots are connected with a line
- Bubble charts: Often used to visualize data points with 3 dimensions, namely visualized on the x-axis, y-axis, and with the size of the bubble. It tries to show relations between data points using location and size
- World cloud chart: A convenient visualization for visualizing the most prevalent words that appear in a text. This can be used to visualize the relationship between different words that appear together or capture a trend on the most commonly prevalent words.

Part-to-whole Charts: These data visualizations allow people to show sub-categories within a large category.

- Pie chart: One of the most common ways to show part to whole data. It is also commonly used with percentages.
- Donut pie chart: The donut pie chart is a variant of the pie chart, the difference being it has a hole in the center for readability.
- Heat maps: Heatmaps are two-dimensional charts that use color shading to represent data trends.
- Stacked column chart: Best to compare subcategories within categorical data. Can also be used to compare percentages.
- Treemap charts: 2D rectangles whose size is proportional to the value being measured and can be used to display hierarchically structured data
- part of the whole chart.

Visualize a Single Value: These data visualizations allow people to visualize a single data point

- Card: Cards are great for showing and tracking KPIs in dashboards or presentations.
- Table chart: Best to be used on small datasets, it displays tabular data in a table.
- Gauge chart: This chart is often used in executive dashboard reports to show relevant KPIs.

Capture Distributions: These data visualizations allow people to visualize the distribution of a variable.

- Histograms: Shows the distribution of a variable. It converts numerical data into bins as columns. The x-axis shows the range, and the y-axis represents the frequency.
- Box plot: Shows the distribution of a variable using 5 key summary statistics—minimum, first quartile, median, third quartile, and maximum.
- Violin plot: A variation of the box plot. It also shows the full distribution of the data alongside summary statistics.
- Density plot: Visualizes a distribution by using smoothing to allow smoother distributions and better capture the distribution shape of the data.

Visualize a flow: These data visualizations allow you to visualize how data points flow into each other.

- Sankey chart: Useful for representing flows in systems. This flow can be any measurable quantity.
- Chord chart: Useful for presenting weighted relationships or flows between nodes. Especially useful for highlighting the dominant or important flows.

- Network chart: Similar to a graph, it consists of nodes and interconnected edges. It illustrates how different items have relationships with each other.

2.5.3. MDX language for analyzing multi-dimension data and OLAP

Multidimensional Expressions (MDX) is the query language used to work with and retrieve multidimensional data in Microsoft SQL Server Analysis Services. MDX is based on the XML for Analysis (XMLA) specification, with specific extensions for SQL Server SQL Server Analysis Services. MDX utilizes expressions composed of identifiers, values, statements, functions, and operators that SQL Server Analysis Services can evaluate to retrieve an object (for example a set or a member), or a scalar value (for example, a string or a number).

MDX queries and expressions in SQL Server SQL Server Analysis Services are used to do the following:

- Return data to a client application from a SQL Server SQL Server Analysis Services cube.
- Format query results.
- Perform cube design tasks, including the definition of calculated members, named sets, scoped assignments, and key performance indicators (KPIs).
- Perform administrative tasks, including dimension and cell security.

2.5.4. DAX language

Data Analysis Expressions (DAX) is a formula expression language used in Analysis Services, Power BI, and Power Pivot in Excel. DAX formulas include functions, operators, and values to perform advanced calculations and queries on data in related tables and columns in tabular data models.

2.5.5. Slowly Changing Dimensions (SCD)

A Slowly Changing Dimension (SCD) is a dimension that stores and manages both current and historical data over time in a data warehouse. It is considered and

implemented as one of the most critical ETL tasks in tracking the history of dimension records.

The three types of SCDs are:

- **Type 1 SCDs - Overwriting**

In a Type 1 SCD the new data overwrites the existing data. Thus the existing data is lost as it is not stored anywhere else. This is the default type of dimension you create. You do not need to specify any additional information to create a Type 1 SCD.

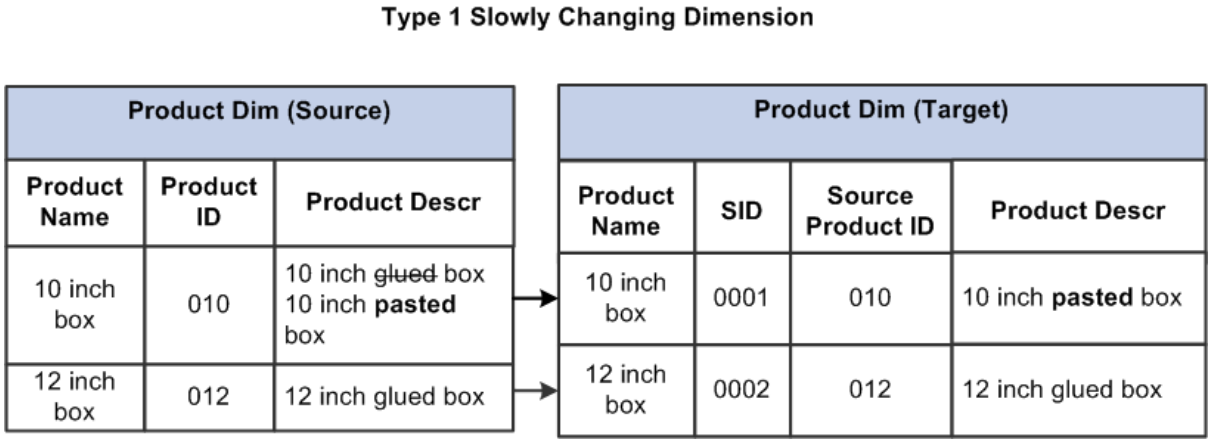

Figure 2.4. Type 1 Slowly Changing Dimension

(Source: Oracle)

- **Type 2 SCDs - Creating another dimension record**

A Type 2 SCD retains the full history of values. When the value of a chosen attribute changes, the current record is closed. A new record is created with the changed data values and this new record becomes the current record. Each record contains the effective time and expiration time to identify the time period between which the record was active.

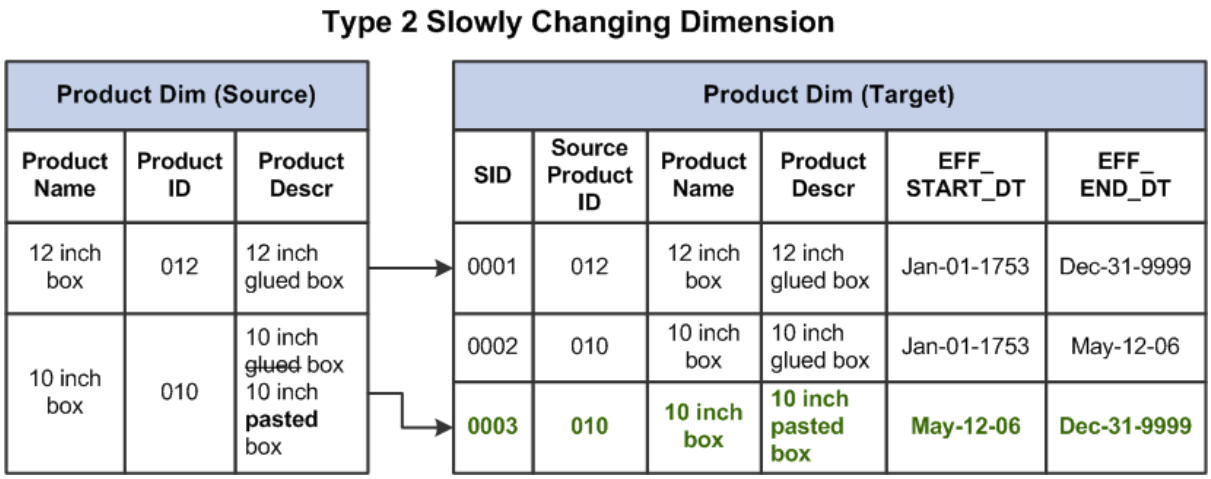

Figure 2.5. Type 2 Slowly Changing Dimension

(Source: Oracle)

- **Type 3 SCDs - Creating a current value field**

A Type 3 SCD stores two versions of values for certain selected level attributes. Each record stores the previous value and the current value of the selected attribute. When the value of any of the selected attributes changes, the current value is stored as the old value and the new value becomes the current value.

CHAPTER 3. ANALYSIS OF USER REQUIREMENTS AND DATA DESCRIPTION

3.1. Business Process

3.1.1. Related department

- Warehouse Department

The Warehouse Department handles the storage and management of inventory. It is responsible for receiving goods from suppliers, storing products in the warehouse and preparing items for shipment. They ensure that stock levels are maintained, manage the distribution of goods and keep accurate records of inventory.

- Purchasing Department

This department is responsible for procuring the goods, raw materials needed by Wide World Importers. They monitor inventory levels, assess supplier performance and handle payments for received goods. The Purchasing Department manages the process of creating purchase orders with suppliers, negotiating terms and ensuring timely delivery of products.

3.1.2. The purpose of Supply Chain Management

The primary purpose of supply chain management (SCM) is to provide customers with the right products at the right time, place and form while maximizing efficiency and profitability for the company. This is achieved through the coordination and optimization of all processes involved in transforming raw materials into final products and delivering them to consumers.

Firstly, the key goal of SCM is aiming to effectively match supply with customer demand, avoiding shortages or excess inventory. Secondly, by streamlining processes, eliminating waste, and leveraging technologies like big data analytics, SCM helps companies cut excess costs and deliver products more efficiently. Thirdly, SCM provides customers with utility in the form of time, place, form, and possession.

Ultimately, Wide World Importers can gain a significant competitive edge by offering superior customer service, lower prices, and faster delivery times.

3.1.3. Supply Chain Management Process

3.1.3.1. Purchasing

Step 1: After identifying the business needs, a purchase requisition is sent to the purchasing department or manager. These requisitions are forwarded to potential suppliers for quotes.

Step 2: The supplier delivers the goods according to the purchase order within the agreed time frame.

Step 3: The verified invoice is approved for payment. Payment is made to the supplier upon receiving the goods, using the payment details on the invoice.

3.1.3.2. Warehouse

Step 1: Receiving and storing goods is vital for warehouse efficiency and accurate supply chain management. This includes receiving goods from suppliers, organizing items for easy access and inventory tracking, and improving processes like receiving and picking.

Step 2: The picking and packing process is crucial for fulfilling customer orders. If the item is available, Wide World Importers packs and ships it to the customer, ensuring product safety and preventing delivery errors. When stock is insufficient, additional orders are placed with suppliers. Available items are sent immediately, while out-of-stock items are delivered later.

Step 3: Shipping is a complex process involving order processing, sorting, picking, packaging, and labeling, all critical for timely and accurate delivery. Wide World Importers uses its own trucks or other shipping services to deliver goods.

3.2. Business Requirements Analysis

3.2.1. Method

Our team chose an interview and survey method to analyze business requirements with the number of stakeholders.

Interview: Start by interviewing some key stakeholders to understand their requirements and desires. The interviews aimed to gather insights into the current business situation, existing data sources, and the specific analytical needs of each department.

Survey: Based on the information collected from interviews, design a detailed survey and send it to a larger group of stakeholders. Surveys help collect data from more people and confirm information gathered from interviews.

After collecting comments, analyzing the current situation and existing data sources, we have identified the analytical problems as below.

3.2.1. Results

3.2.1.1. Functional Requirements

Warehouse Department

- **Inventory Management:** Optimize stock levels, identify slow-moving items, and reduce stockouts.
- **Warehouse Operations:** Assess warehouse productivity and space utilization.
- **Order Fulfillment:** Analyze order processing times and fulfillment rates.

Purchasing Department

- **Supplier Performance:** Evaluate supplier reliability, delivery times, and cost-effectiveness.
- **Cost Analysis:** Compare supplier costs and negotiate better terms.

- **Inventory Reorder Points:** Calculate optimal reorder points for different products.

3.2.1.2. Non-functional Requirements

- **Security:** The system must ensure the protection of sensitive information by preventing unauthorized access and misuse.
- **Usability:** The system must have a friendly user interface, allowing users to easily manipulate, access and manage system functions.
- **Performance:** The system must be able to handle large volumes of data and provide real-time analytics to the team.
- **Scalability:** The system needs to be scalable, accommodating increases in data volume and user traffic as it grows over time.

3.3. Data source and challenges

Data Source: For this project, we are utilizing the Wide World Importers (WWI) dataset, it is a sample database provided by Microsoft, designed to demonstrate the features of SQL Server. It's a comprehensive dataset that includes various business scenarios for wholesale companies, offering a realistic environment to explore business intelligence and data analytics.

Challenges:

- *Data Quality:* The WWI dataset, like many sample datasets, have issues such as missing values, duplicate entries, and inconsistencies between tables. These issues can arise due to the synthetic nature of the data or during data generation processes.
- *Data Complexity:* The dataset represents a simplified business model, which may not fully capture the complexities of real-world operations. It provides a snapshot of data at a specific point in time, limiting analysis of trends and patterns over time.

- *Data Integration:* The dataset typically includes a limited number of data sources, which may not reflect the diversity of data integration scenarios in real-world environments.
- *Data Size:* The large volume of data, with over 10,000 entries, poses significant challenges in terms of storage, processing, and analysis. Efficient data management strategies, such as scalable storage solutions and powerful processing capabilities, are necessary to handle the data size.

3.4. IT requirements analysis

In Business Intelligence (BI) analysis for the Sales, Warehouse, and Purchasing modules, several key components are essential to build a robust infrastructure. These components ensure that data is collected, processed, analyzed, and visualized effectively to support informed decision-making. Here are the important components of infrastructure in BI for the Sales, Warehouse, and Purchasing modules:

Table 1. IT requirements analysis

Requirement	Description
Database Management System (DBMS)	This is a software system used to manage and store business's data. In BI for Sales, Warehouse, and Purchasing, the DBMS ensures that related data is stored and managed safely and accurately.
Data Warehouse (DWH)	Centralized Data Storage: A data warehouse or data lake where data from various sources is consolidated.
ETL Process (Extract - Transform - Load)	Extract, Transform, Load (ETL) tools and processes to integrate data from multiple sources into the data warehouse. This process

	ensures that data is processed and normalized before being stored in the DWH.
Business Intelligence Tools	Dashboards: Interactive dashboards for real-time data visualization (e.g., Power BI, etc.). BI tools help to quickly and accurately analyze and make decisions based on information and data related to Sales, Warehouse, and Purchasing.
Infrastructure	Infrastructure includes physical components such as servers, storage systems, network systems, and other essential elements that facilitate the smooth and efficient functioning of the BI system.

Implementing a comprehensive BI infrastructure for the Sales, Warehouse, and Purchasing modules involves integrating these components to ensure seamless data flow, accurate analysis, and effective decision-making.

CHAPTER 4: DATA WAREHOUSE IMPLEMENTATION

4.1. Designing Data Warehouse

94.1.1. Bus Matrix

Table 2. Bus Matrix

Business processes	Dim StockItem	Dim Supplier	Dim TransactionType	Dim PackageType	Dim PaymentMethod	DimTime
Warehouse						
Receive goods from suppliers	x	x	x		x	x
Organize items for easy access and inventory tracking	x			x		
Pick items according to customer orders	x			x		

Pack and prepare items for shipment	x			x		x
Ensure product safety and prevent delivery errors	x					
If items are available, pack and ship to the customer	x		x		x	x
Process orders, sort, pick, pack, and label items	x			x		
Business processes	**Dim StockItem**	**Dim Supplier**	**Dim TransactionType**	**Dim PackageType**	**Dim PaymentMethod**	**DimTime**
Purchasing						
Send purchase requisition to		x	x			x

purchasing department or manager						
Forward requisitions to potential suppliers for quotation		x				x
Review quotation from suppliers		x				x
Select appropriate supplier and send purchase order	x	x	x		x	x
Supplier delivers goods according to the purchase order	x	x	x	x		x
Verify and confirm the received goods	x	x	x	x		x
Verify the invoice		x	x		x	x

from the supplier						
Approve the invoice for payment		x	x		x	x
Make payment to the supplier based on the payment details on the invoice		x	x		x	x

4.1.2.Master Data

Master Data refers to a type of data that is highly static, meaning it rarely changes and typically holds long-term value. Because Master Data is utilized by many departments within a company, it is closely tied to all business processes of the organization. Master Data is one of the criteria used in the process of data analysis and evaluation. It serves as the key and core of the organization.

Table 3. Master Data

Objective	Description
Purchasing.Suppliers	This table contains information about the suppliers from whom the company purchases goods.
Purchasing.SupplierCategories	This table categorizes suppliers into different categories.
Warehouse.StockItems	Contains information about the items stored in the warehouse.
Warehouse.PackageTypes	Ways that stock items can be packaged (e.g., box, carton, pallet, kg, etc.)
Application.TransactionTypes	Types of customer, supplier, or stock transactions (e.g., invoice, credit note, etc.)
Application.PaymentMethods	Ways that payments can be made (ie: cash, check, EFT, etc.)

4.1.3.Transaction Data

Transaction Data refers to dynamic data used by specific departments and related to the transactions of a business. Its relevance is confined within a certain timeframe. For the mentioned statistics, Transaction Data provides the factual basis behind the criteria of Master Data. Additionally, Transaction Data is dependent on Master Data;

without Master Data, there can be no Transaction Data. Therefore, Transaction Data is often renewed and replaced.

Table 4. Transaction Data

Objective	Description
Purchasing.PurchaseOrders	This table contains details about purchase orders issued by the company to suppliers. This table details the items included in each purchase order.
Purchasing.PurchaseOrderLines	This table contains details about purchase orders issued by the company to suppliers. This table details the items included in each purchase order.
Purchasing.SupplierTransactions	All financial transactions that are supplier-related (invoices, payments)
Warehouse.StockItemTransactions	Transactions covering all movements of all stock items (receipt, sale, write-off).

4.1.4. Fact and Dimension Table

4.1.4.1 Dimension tables

Dim Stock Item: Information about Item

Table 5. DimStockItem

Data Warehouse		Data Source		Data type	Allow Null	Rules
Table	**Column name**	**Table**	**Column name**			

DimStockItem	StockItemKey			int		Auto increment
	StockItemID	Warehouse.StockItems	StockItemID	int		From source
	StockItemName	Warehouse.StockItems	StockItemName	nvarchar (100)		From source
	StockGroupName	Warehouse.StockItems	StockGroup Name	nvarchar (50)		From source
	Barcode	Warehouse.StockItems	Barcode	nvarchar (50)	x	From source
	Brand	Warehouse.StockItems	Brand	nvarchar (50)	x	From source
	Size	Warehouse.StockItems	Size	nvarchar (20)	x	From source
	TypicalWeight PerUnit	Warehouse.StockItems	TypicalWeightPerUnit	decimal(18, 3)		From source
	ColorName	Warehouse.Colors	ColorName	nvarchar (20)	x	From source
	PackageTypeName	Warehouse.PackageTypes	PackageType Name	nvarchar (50)		From source
	FromDate			datetime		FromDate
	ToDate			datetime	x	ToDate

DimSupplier: Information of Supplier

Table 6. DimSupplier

Data Warehouse		Data Source		Data type	Allow Null	Rules
Table	**Column name**	**Table**	**Column name**			
DimSupplier	SupplierKey			int		Auto increment
	SupplierID	Purchasing.Suppliers	SupplierID	int		From source
	SupplierName	Purchasing.Suppliers	SupplierName	nvarchar(100)		From source
	DeliveryPostalCode	Purchasing.Suppliers	DeliveryPostalCode	nvarchar(10)	x	From source
	SupplierReference	Purchasing.Suppliers	SupplierReference	nvarchar(20)		From source
	SupplierCategoryName	Purchasing.SupplierCategories	SupplierCategoryName	nvarchar(50)		From source
	FromDate			datetime		FromDate
	ToDate			datetime	x	ToDate

Dim Transaction Type: Information of the type of transaction

Table 7. DimTransactionType

Data Warehouse		Data Source		Data type	Allow Null	Rules
Table	**Column name**	**Table**	**Column name**			
DimTran	TransactionTyp			int		Auto

sactionType	eKey					increment
	TransactionTypeID	Application.TransactionTypes	TransactionTypeID	int		From source
	TransactionTypeName	Application.TransactionTypes	TransactionTypeName	nvarchar(50)		From source
	FromDate			datetime		FromDate
	ToDate			datetime	x	ToDate

DimPackageType: Information of the types of Package

Table 8. DimPackageType

Data Warehouse		**Data Source**		**Data type**	**Allow Null**	**Rules**
Table	**Column name**	**Table**	**Column name**			
DimTransactionType	PackageTypeKey			int		Auto increment
	PackageTypeID	Warehouse.PackageTypes	TransactionTypeID	int		From source
	PackageTypeName	Warehouse.PackageTypes	TransactionTypeName	nvarchar(100)		From source
	FromDate			datetime		FromDate
	ToDate			datetime	x	ToDate

DimPaymentMethod: Information of Payment Method

Table 9. DimPaymentMethod

Data Warehouse		Data Source		Data type	Allow Null	Rules
Table	**Column name**	**Table**	**Column name**			
	PaymentMethodKey			int		Auto increment
	PaymentMethodID	Application.PaymentMethods	PaymentMethodID	int		From source
	PaymentMethodName	Application.PaymentMethods	PaymentMethodName	nvarchar(50)		From source
	FromDate			datetime		FromDate
	ToDate			datetime	x	ToDate

DimTime: Information of Datetime

Table 10. DimTime

Data Warehouse		Data Source		Data type	Allow Null	Rules
Table	**Column name**	**Table**	**Column name**			
DimTime	DateKey			datetime		Auto increment
	TheDay			int		

	TheDayName			nvarchar (30)		
	TheWeek			int		
	TheISOWeek			int		
	TheDayOfWeek			int		
	TheMonth			int		
	TheMonthName			nvarchar (30)		
	TheQuarter			int		
	TheYear			int		
	TheFirstOfMonth			date		
	TheLastOfYear			date		
	TheDayOfYear			int		

4.1.4.2 Fact tables

FactPurchase: Detail information about Purchasing Order

Table 11. FactPurchase

Data Warehouse		**Data Source**		**Data type**	**Allow Null**	**Rules**
Table	**Column name**	**Table**	**Column name**			
FactPurchase	PurchaseOrderID	Purchasing.PurchaseOrders	PurchaseOrderID	int		From source

	PurchaseOrderLineID	Purchasing.PurchaseOrderLines	PurchaseOrderLineID	int		From source
	OrderDate	Purchasing.PurchaseOrders	OrderDate	datetime		From source
	IsOrderFinalized	Purchasing.PurchaseOrders	IsOrderFinalized	bit		From source
	SupplierKey	DimSupplier	SupplierKey	int		From source
	SupplierName	Purchasing.Suppliers	SupplierName	nvarchar(100)		From source
	OrderedOuters	Purchasing.PurchaseOrderLines	OrderedOuters	int		From source
	ReceivedOuters	Purchasing.PurchaseOrderLines	ReceivedOuters	int		From source
	StockItemKey	DimStockItem	StockItemKey	int		From source
	ExpectedUnitPricePerOuter	Purchasing.PurchaseOrderLines	ExpectedUnitPricePerOuter	decimal(18, 2)		From source
	PackageTypeKey	DimPackageType	PackageTypeKey	int		From source
	PackageTypeNa	Warehouse.Pa	PackageTy	nvarchar		From

	me	ckageTypes	peName	(50)		source
	TransactionTypeKey	DimTransactionType	TransactionTypeKey	int		From source
	TransactionAmount	Purchasing.SupplierTransactions	TransactionAmount	decimal(18, 2)		From source
	TaxAmount	Purchasing.SupplierTransactions	TaxAmount	decimal(18, 2)	x	From source
	AmountExcludingTax	Purchasing.SupplierTransactions	AmountExcludingTax	decimal(18, 2)		From source
	TransactionDate	Purchasing.SupplierTransactions	TransactionDate	date		From source
	FinalizationDate	Purchasing.SupplierTransactions	FinalizationDate	date	x	From source
	OutstandingBalance	Purchasing.SupplierTransactions	OutstandingBalance	decimal(18, 2)		From source
	PaymentMethodKey	DimPaymentMethod	PaymentMethodKey	int		From source
	PaymentMethodName	Application.PaymentMethods	PaymentMethodName	nvarchar(50)		From source

	RunAs			nvarchar (64)	x	
	ModifiedDate			datetime	x	

Fact Stock Item Holding: Information of holding stock items

Table 12. Fact Stock Item Holding

Data Warehouse		**Data Source**		**Data type**	**Allow Null**	**Rules**
Table	**Column name**	**Table**	**Column name**			
FactStockHodling	StockItemKey	DimStockItem	StockItemKey	int		From source
	QuantityOnHand	Warehouse.StockItemHoldings	QuantityOnHand	int		From source
	BinLocation	Warehouse.StockItemHoldings	BinLocation	nvarchar (20)		From source
	LastStocktakeQuantity	Warehouse.StockItemHoldings	LastStocktakeQuantity	int		From source
	LastCostPrice	Warehouse.StockItemHoldings	LastCostPrice	decimal(18, 2)		From source
	ReorderLevel	Warehouse.StockItemHoldings	ReorderLevel	int		From source
	TargetStockLevel	Warehouse.StockItemHoldings	TargetStockLevel	int		From source
	Quantity	Warehouse.Stoc	Quantity	decimal(		From

		kItemTransactions		18, 3)		source
	TransactionOccurredWhen	Warehouse.StockItemTransactions	TransactionOccurredWhen	datetime		From source
	TransactionTypeKey	DimTranSactionType	TransactionTypeKey	int		From source
	TransactionTypeName	Application.TransactionTypes	TransactionTypeName	nvarchar (50)		From source
	ModifiedDate			nvarchar (64)	x	
	RunAs			datetime	x	

4.1.5.Data warehouse model

The Galaxy model is the Data warehouse model used here, with DimStockItem and DimTransactionType tables applied for both FactPurchase and FactStockHolding tables.

The Galaxy model, a popular data model in the Business Intelligence (BI) field, is used to manage a company's business information. This model has two primary functions: Purchasing (corresponding to FactPurchase) and warehouse (corresponding to FactStockHolding).

The FactPurchase table is used to manage the purchasing related information of the Purchasing segment. This table contains metrics such as number of products bought, purchasing price and purchasing on demand. These metrics are connected to dim tables like DimTime, DimSupplier and DimStockItem, DimPackageType, DimTransactionType and DimPaymentMethod.

The FactStockHolding table is used to manage information relating to stock item holding of the warehouse segment. This table contains metrics like stock quantity, location,etc. Dim tables are used to categorize stock items holding information and stored in dim tables like DimStockItem and DimTransactionType.

The Galaxy model enables the creation of reports and analyses tailored to specific business functions, which can then be integrated to inform strategic decision-making across the company.

4.1.5.1. Data Warehouse Model

The proposed DW model is described in the figure below (Figure 4.1)

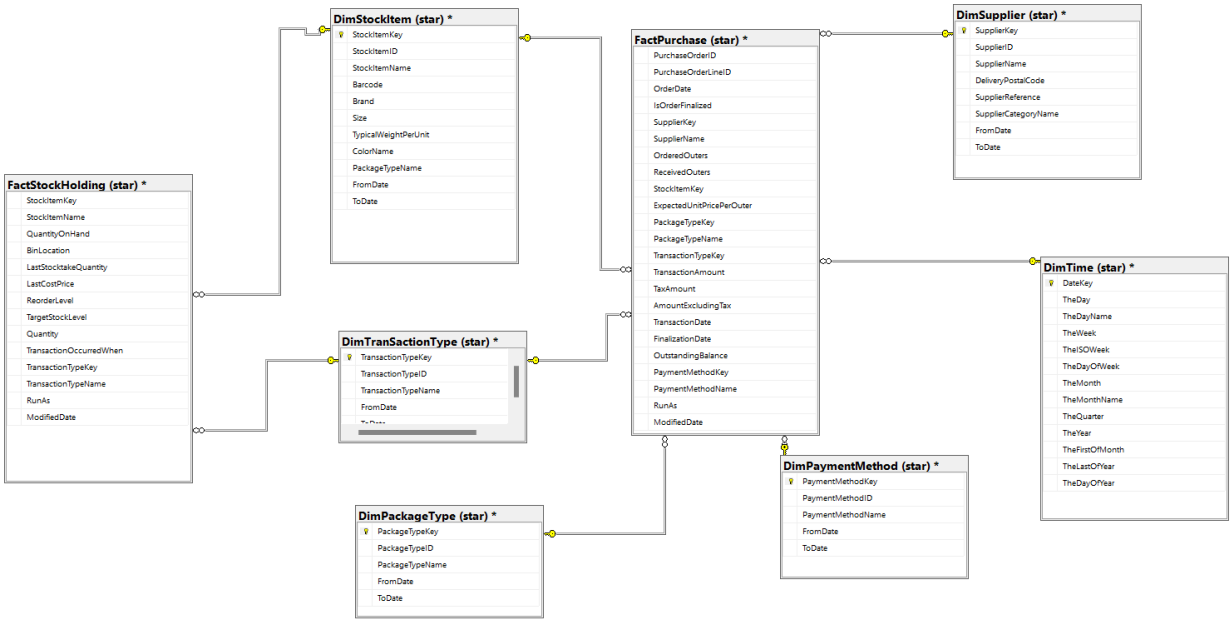

Figure 4.1. The proposed Data Warehouse model

Source: The author synthesizes suggestions

Table 13. The description of the tables in the Data Warehouse model

Objective	Description
DimStockItem	Information about Stock Item
DimSupplier	Information about Supplier

DimTime	Detail information about Time
DimPaymentMethod	Information about Payment Method
Dim TransactionType	Information about Transaction Type
DimPackageType	Information about Package Type
FactPurchase	Detail information about Purchase Order
FactStockHolding	Detail information about Stock Holding

4.1.5.2. The main relationships in the data warehouse model schema

Table 14: The table describes the relationship in the Data Warehouse model

No.	Relationship	Relationship Type	Description
1	DimSupplier - FactPurchase	1 - n	One supplier may have one or many lines on FactPurchase, each line on FactPurchase belongs to only one supplier.
2	DimStockItem - FactPurchase	1 - n	One stock item may have one or many lines on FactPurchase, each line on FactPurchase belongs to only one stock item.
3	DimTime - FactPurchase	1 - n	One time can have one or many lines on FactPurchase, each order belongs to only one time
4	DimTransactionType-FactPurchase	1 - n	One transaction type may have one or many lines on FactPurchase, each line on FactPurchase belongs to only one

			transaction type .
5	DimPackageType - FactPurchase	1 - n	One package type may have one or many lines on FactPurchase, each line on FactPurchase belongs to only one package type .
6	DimPaymentMethod-FactPurchase	1 - n	One payment method may have one or many lines on FactPurchase, each line on FactPurchase belongs to only one payment method .
7	DimStockItem - FactStockHolding	1 - n	One stock item may have one or many lines on FactStockHolding, each line on FactStockHolding belongs to only one stock item.
8	DimTransactionType-FactStockHolding	1 - n	One transaction type may have one or many lines on FactStockHolding, each line on FactStockHolding belongs to only one transaction type.

4.2. ETL Process

After defining the data warehouse model, the next step is to use the SSIS (SQL Server Integration Services) tool to convert the data from the original source to the Data Warehouse.

4.2.1. Steps for Dimension Table's ETL Process

Utilize Slowly Changing Dimension Type 2 (SCD2) for integrating data into Dimension tables to effectively track and manage historical changes in Dimension properties. By preserving the complete history of values, we can analyze data trends and make informed business decisions based on the historical context of the data.

The following will illustrate the data integration for the Dimension table:

- *Step 1:* Define a data source for the Dimensions table, primarily using the Wide World Importers sample database.
- *Step 2:* Convert the source data attributes into columns for the Dimensions table, ensuring that data types, lengths, and formats are consistent between the source and the target.
- *Step 3:* Set up a Dimension table in the target database or data warehouse with the appropriate columns and data types. Add extra columns to the Dimension table for tracking historical changes, such as "FromDate", "ToDate", and the surrogate key for the previous version of the record.
- *Step 4:* Load the transformed data into the Dimensions table, ensuring to handle SCD2 changes properly.
- *Step 5:* Verify that the data in the Dimensions table is accurate, complete, and consistent with the source data.

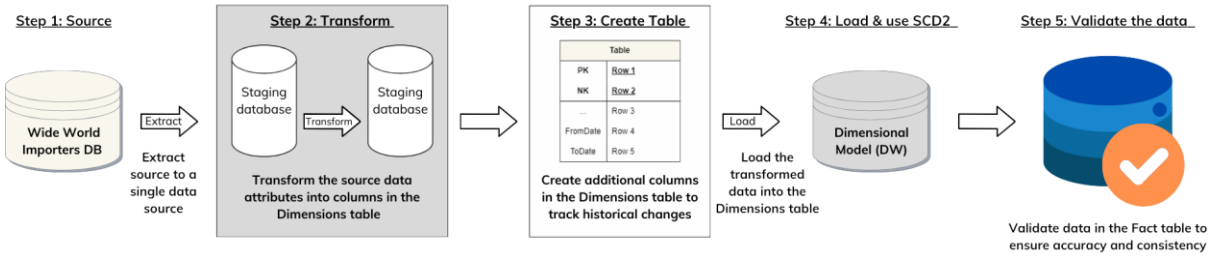

Figure 4.2. Dimension Table's ETL Process

Source: The author synthesizes suggestions

According to the aforementioned steps, we respectively have the results of ETL process of Dimension tables as below:

4.2.1.1. ETL Process of DimStockItem table

The following figure illustrates the ETL process of DimStockItem Table:

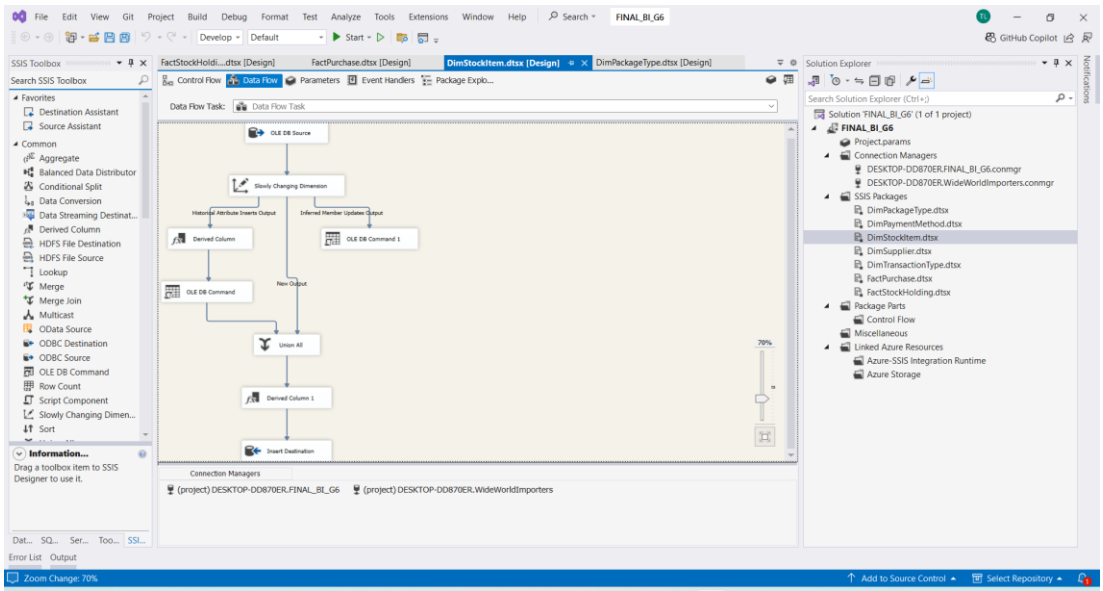

Figure 4.3. DimStockItem Table's ETL Process

Source: The author synthesizes suggestions

After setting the link, we conduct the execution of DimStockItem Package and below is the successful package execution.

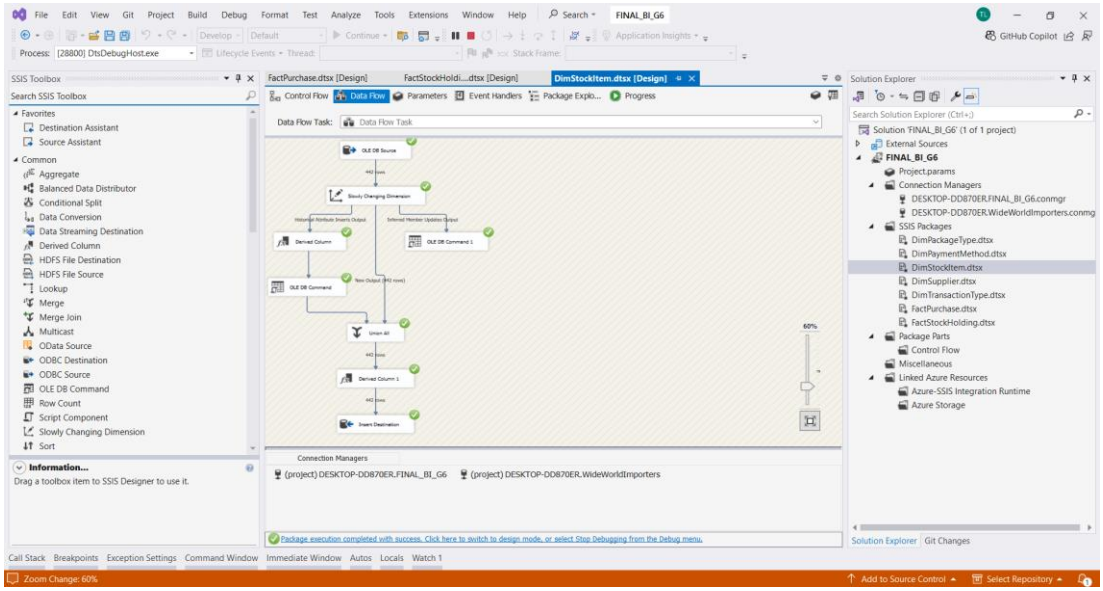

Figure 4.4. DimStockItem Table's successful package execution

Source: The author synthesizes suggestions

Then, we use SQL Server to check whether the data is converted into the DimStockItem table or not. It is obvious that the DimStockItem table has already got the data with 442 rows, which is similar to the result we have got from SSIS.

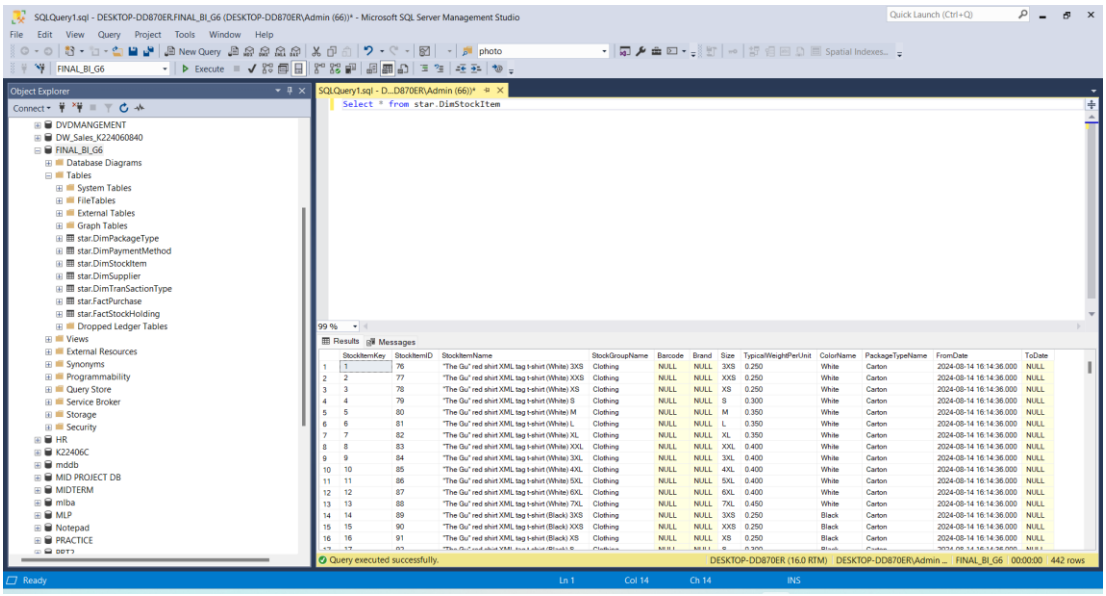

Figure 4.5. DimStockItem Table's check

Source: The author synthesizes suggestions

4.2.1.2. ETL Process of DimSupplier table

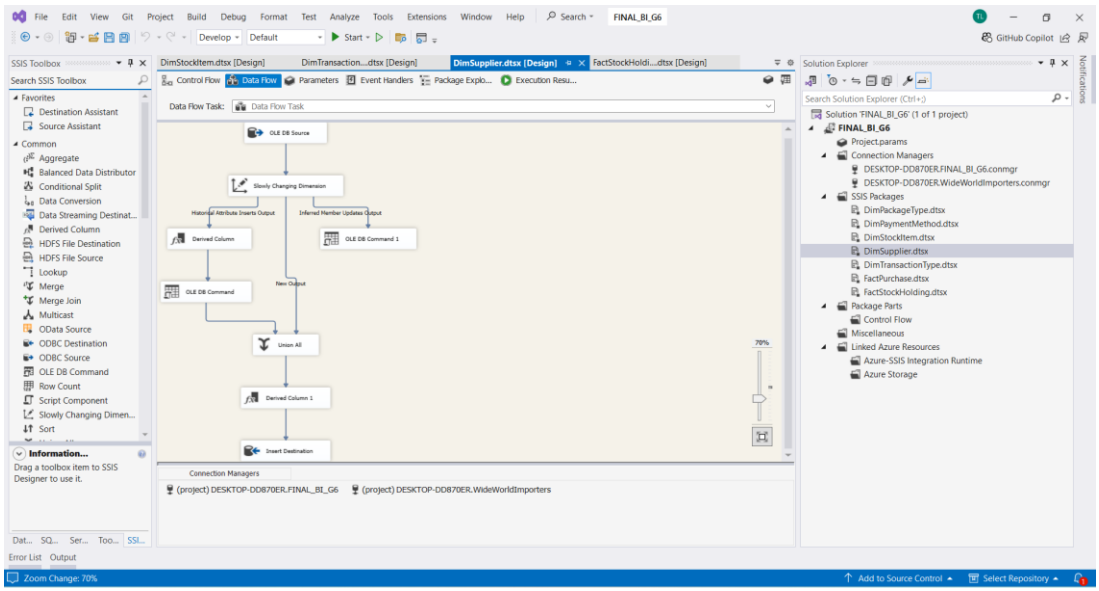

Figure 4.6. DimSupplier Table's ETL Process

Source: The author synthesizes suggestions

The figure above shows the ETL process of the DimSupplier table. Similarly, we execute the DimSupplier package after linking the item from the toolbox and get the successful results as below:

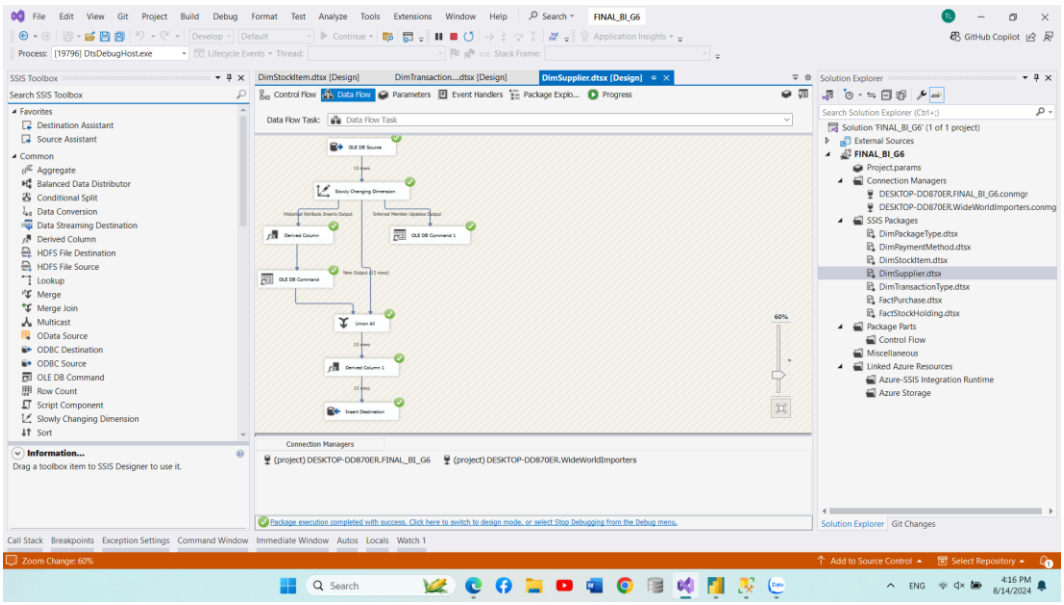

Figure 4.7. DimSupplier Table's successful package execution

Source: The author synthesizes suggestions

Then, we use SQL Server to check the availability of data in the DimSupplier table with 13 rows, which is the same as the results shown in SSIS process.

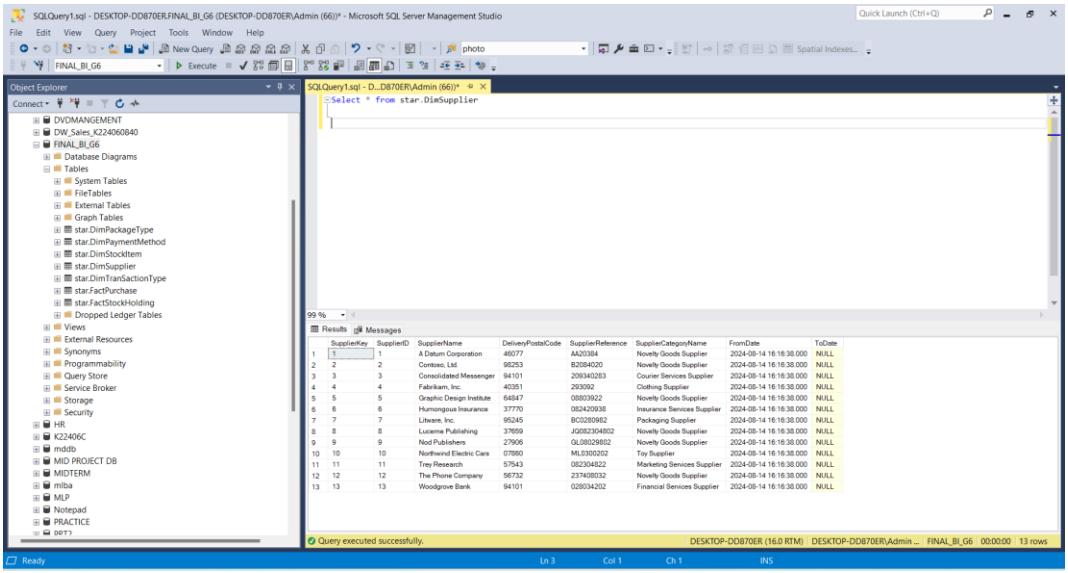

Figure 4.8. DimSupplier Table's check

Source: The author synthesizes suggestions

4.2.1.3. ETL Process of DimTransactionType table

Below is the figure demonstrating the ETL process of the DimTransactionType table:

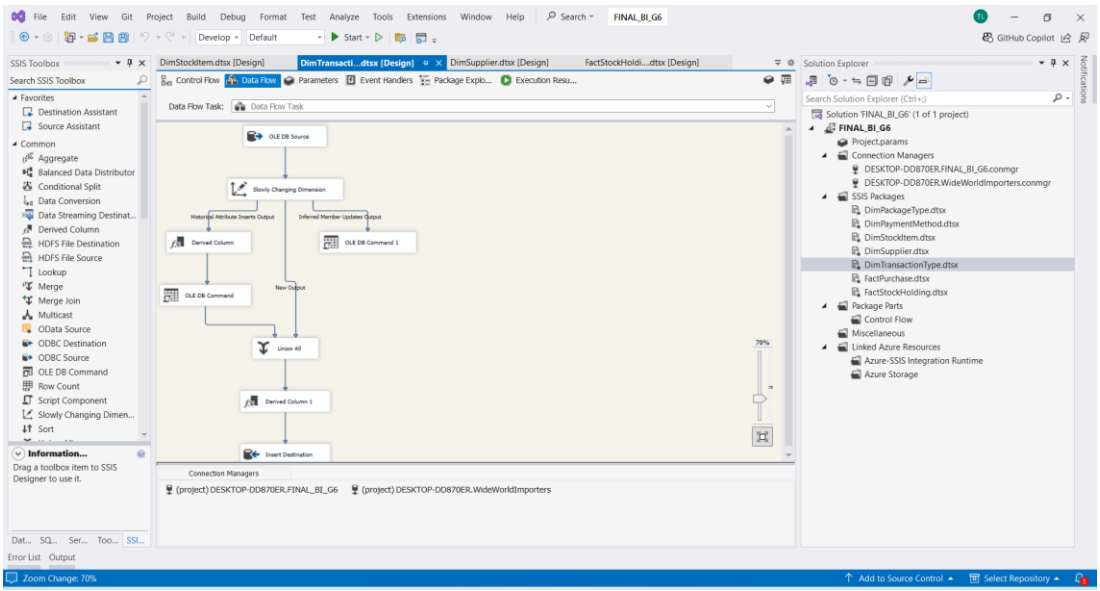

Figure 4.9. DimTransactionType Table's ETL Process

Source: The author synthesizes suggestions

Similarly to the aforementioned table, we execute the DimTransactionType Package after connecting the item from SSIS toolbox and get the following result:

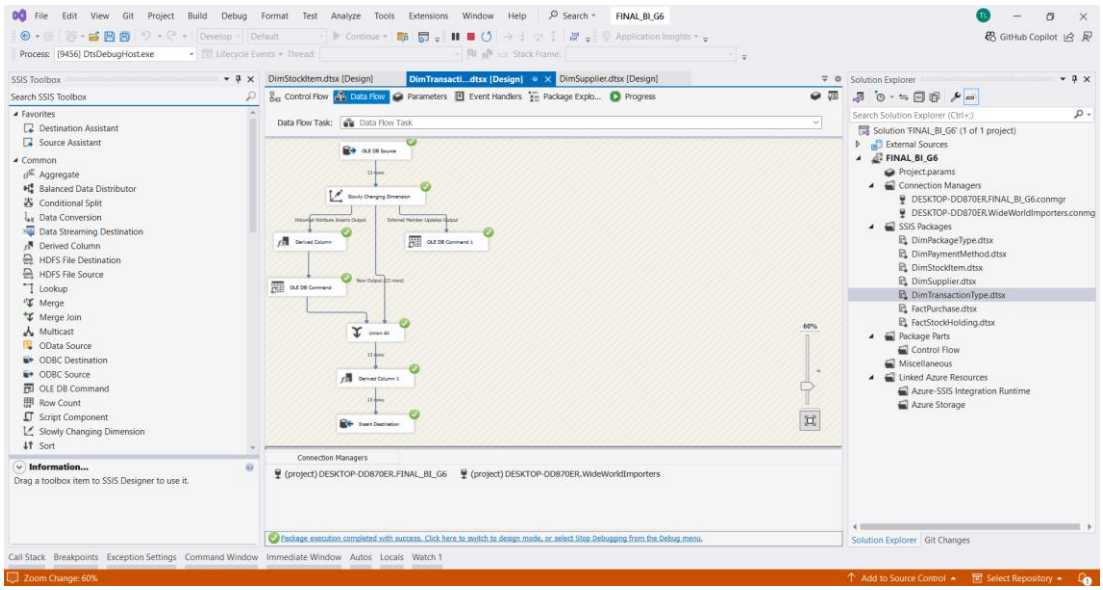

Figure 4.10. DimTransactionType Table's successful package execution

Source: The author synthesizes suggestions

After that, we continually use the SQL Server to check the availability of converted data into the DimTransactionType table and get the results of 13 rows matching the above SSIS process.

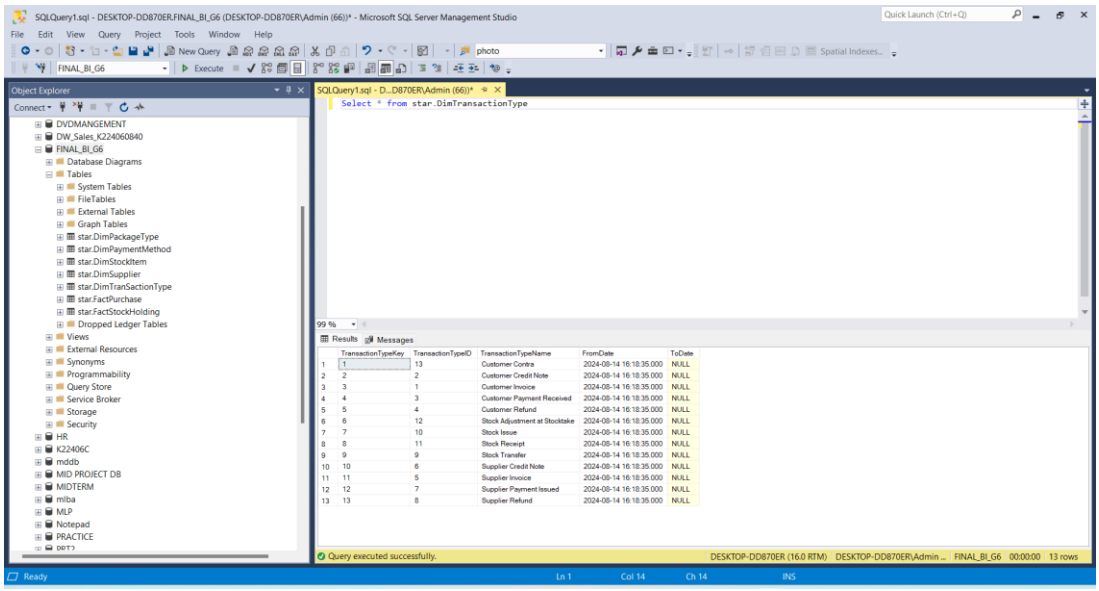

Figure 4.11. DimTransactionType Table's check

Source: The author synthesizes suggestions

4.2.1.4. ETL Process of DimPackageType table

The below figure shows the ETL Process of DimPackageType table:

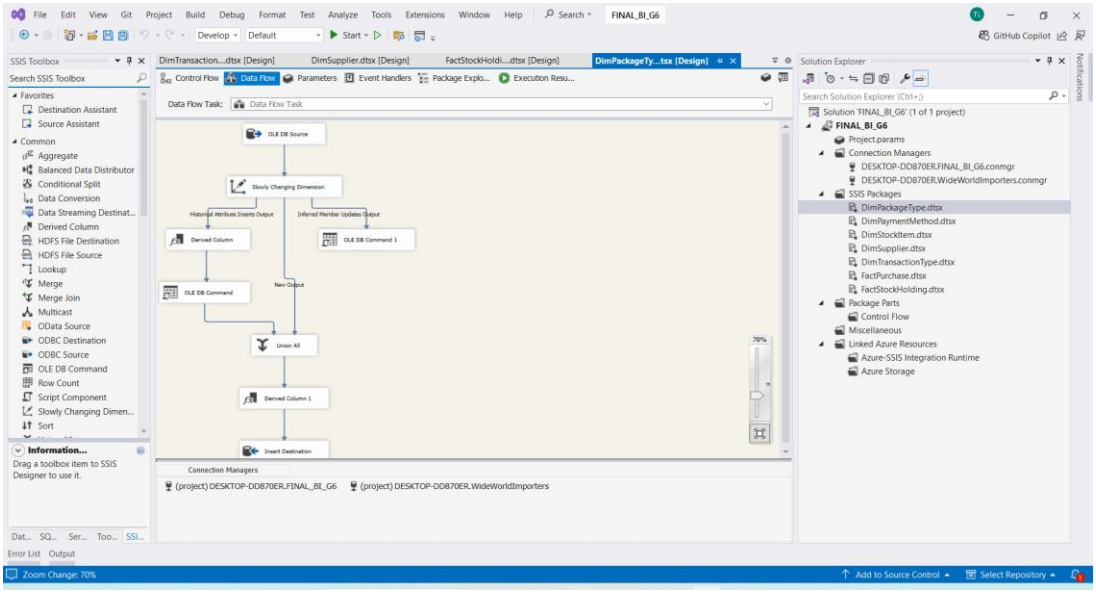

Figure 4.12. DimPackageType Table's ETL Process

Source: The author synthesizes suggestions

After matching the suitable item from the SSIS toolbox, we conduct the package execution of the DimPackageType table and have a completion result as below:

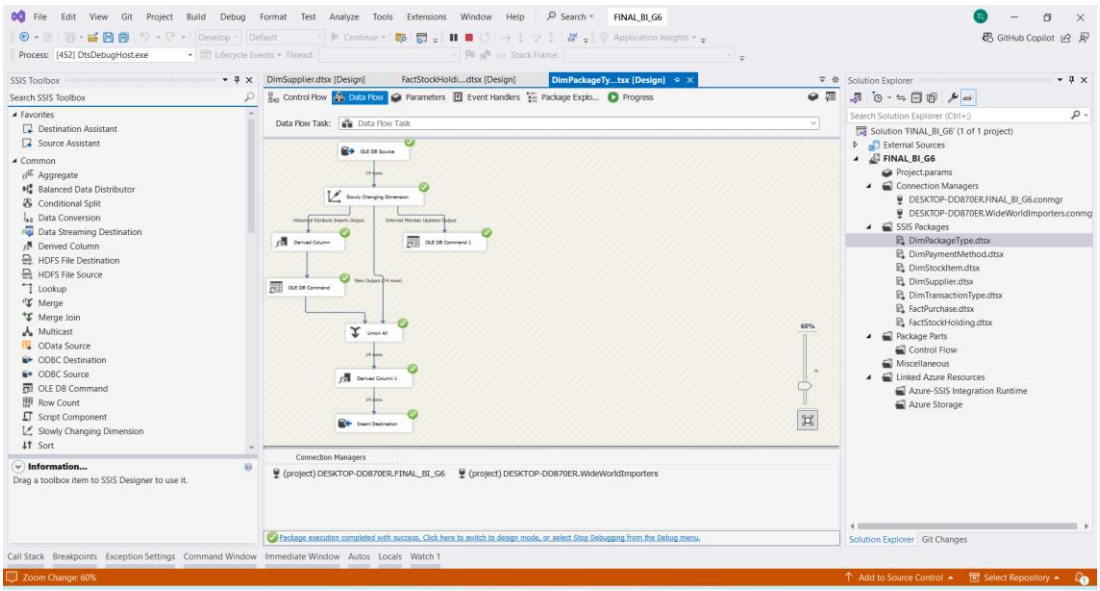

Figure 4.13. DimPackageType Table's successful package execution

Source: The author synthesizes suggestions

Similarly, we still check whether the available data is converted into the DimPackageType table or not in SQL Server, and the result shows that there are 14 rows in this table, which is the same as the successful result from the SSIS process.

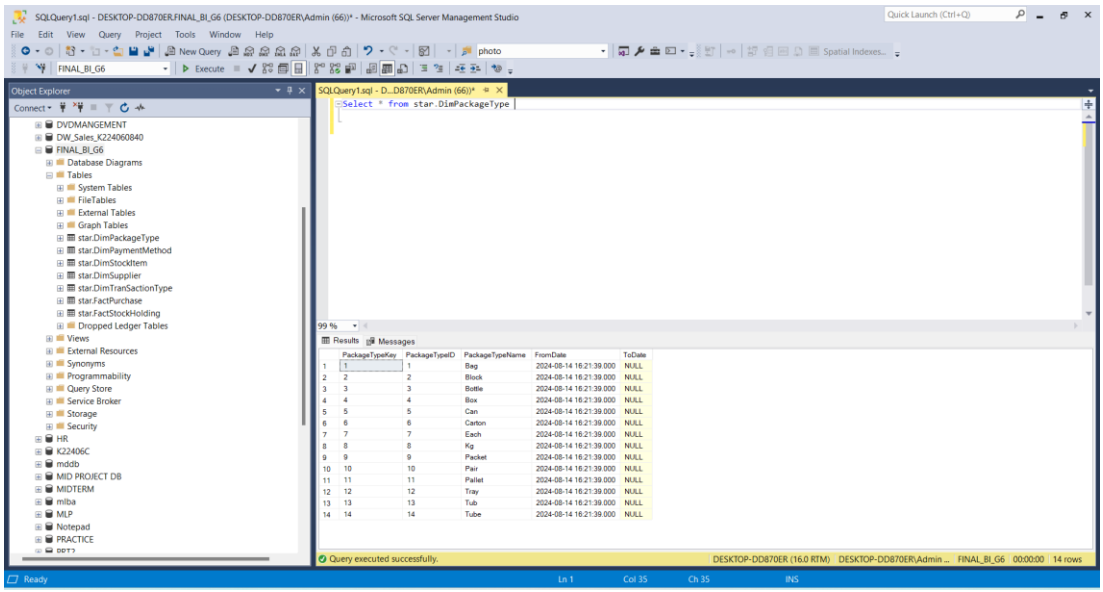

Figure 4.14. DimPackageType Table's check

Source: The author synthesizes suggestions

4.2.1.5. ETL Process of DimPaymentMethod table

The following figure illustrates the ETL Process of the DimPaymentMethod table:

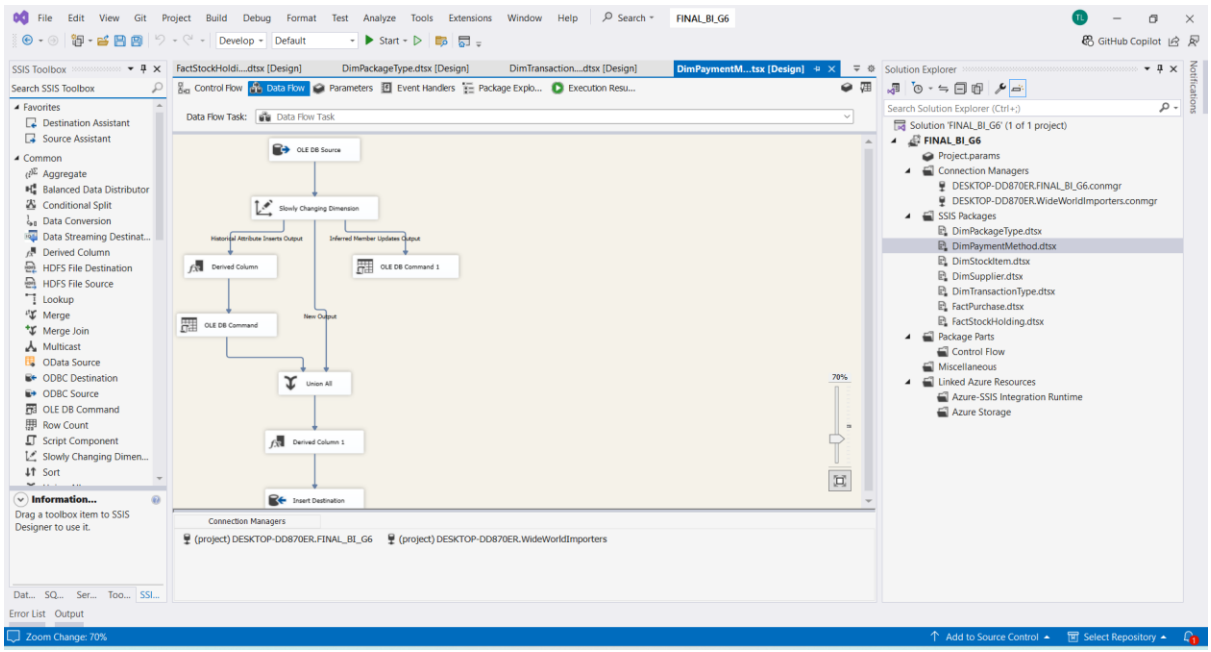

Figure 4.15. DimPaymentMethod Table's ETL Process

Source: The author synthesizes suggestions

After setting the link, we conduct the execution of DimPaymentMethod Package and the below figure demonstrates the successful package execution.

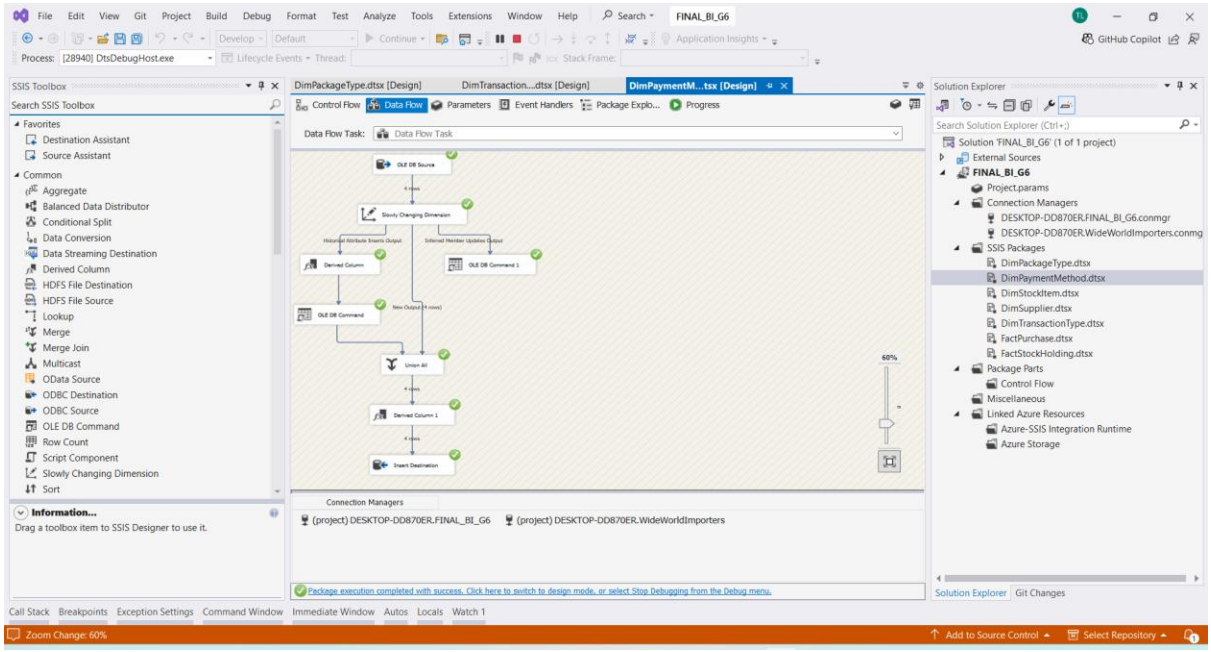

Figure 4.16. DimPaymentMethod Table's successful package execution

Source: The author synthesizes suggestions

After that, we similarly use SQL Server to check if the data is converted into the DimPaymentMethod table and the DimStockItem table has already got the data with 4 rows, which is similar to the result we have got from SSIS.

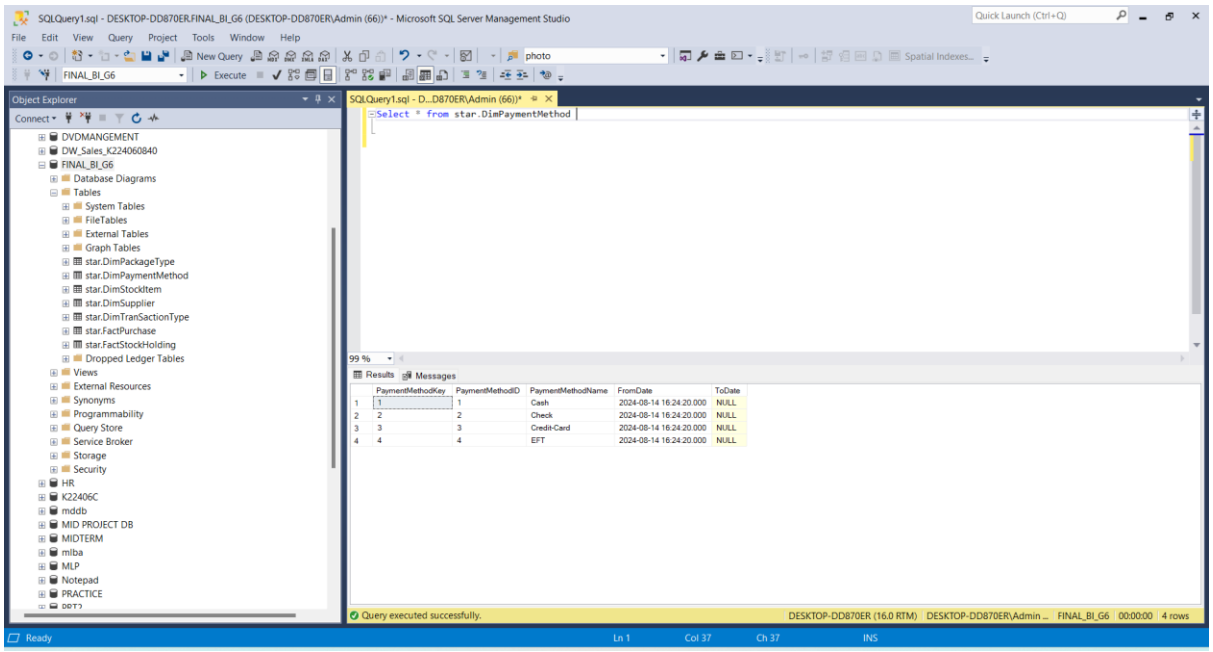

Figure 4.17. DimPaymentMethod Table's check

Source: The author synthesizes suggestions

4.2.2. Fact Table's ETL Process

The process for integrating data into the Fact table involves identifying the source data, transforming it to meet the required format, and updating existing records in the Fact table condition. This ensures the data's accuracy and performance are maintained over time.

The following will illustrate the data integration for the Fact table:

- *Step 1:* Identify the data source for the Fact table, which is drawn from the Wide World Importers sample database, and outline the Dimension tables.
- *Step 2:* Merge join data from source data into a single data source by performing a merge join. Prior to this, ensure that the data is sorted uniformly across the two input sources. Transform the source data attributes to align with the columns of the Fact table, ensuring that data types, lengths, and formats are consistent between the source and the destination.

- *Step 3:* Set up a Fact table in the target database or data warehouse, ensuring it includes the necessary columns and data types. Add an additional column, ModifiedDate, RunAs, to track changes in the data.
- *Step 4:* Load the converted data into the Fact table. This may involve updating existing records in the Dimensions table based on previously defined business rules.
- *Step 5:* Verify the data in the Fact table to confirm its accuracy and consistency with the original source data.

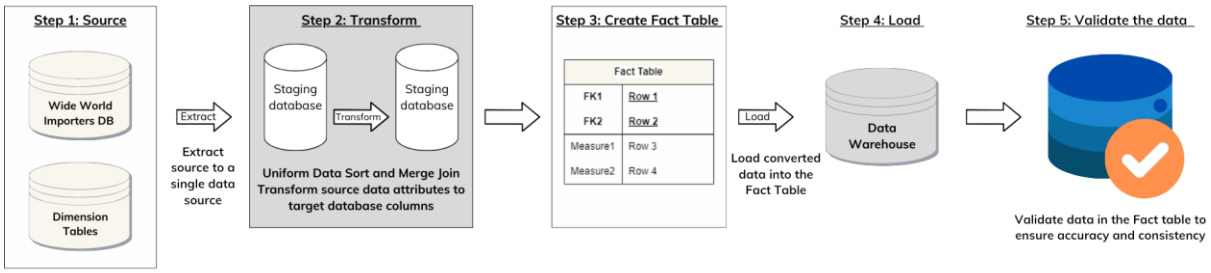

Figure 4.18. Fact Table's ETL Process

Source: The author synthesizes suggestions

According to the aforementioned steps, we have conducted the processes relating to FactPurchase and FactStockHolding and have gotten the results as below:

4.2.2.1. ETL process of FactPurchase Table

The first following figure demonstrates the ETL process of the FactPurchase Table.

And the second table shows the successfully completed package execution of FactPurchase Table after connecting the suitable items from SSIS toolbox.

Then, we continually make use of SQL Server to check the success of the data conversion process from SSIS, which is illustrated in the last figure. Obviously, there are already 8,358 rows obtained from the ETL process, which is matching the rows shown from SSIS.

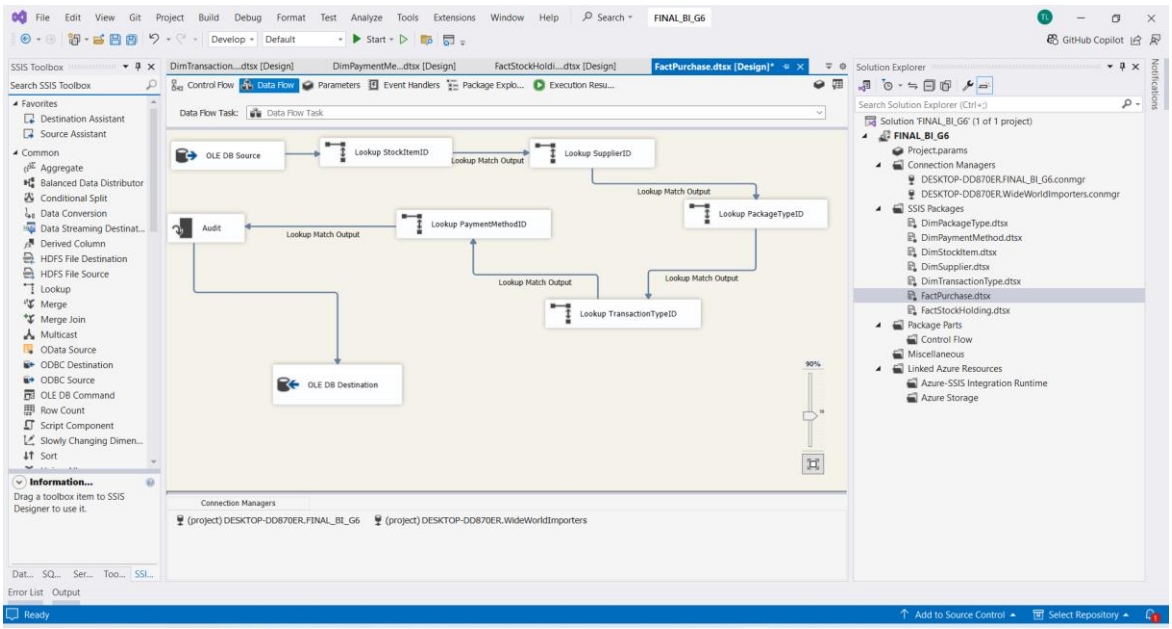

Figure 4.19. FactPurchase Table's ETL Process

Source: The author synthesizes suggestions

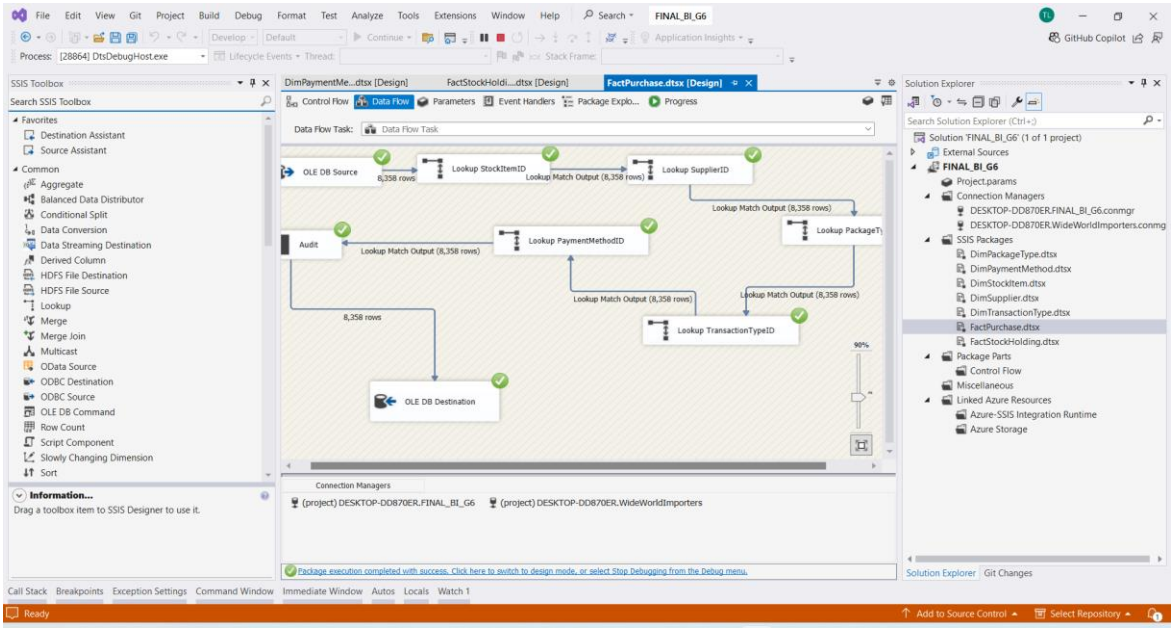

Figure 4.20. FactPurchase Table's successful package execution

Source: The author synthesizes suggestions

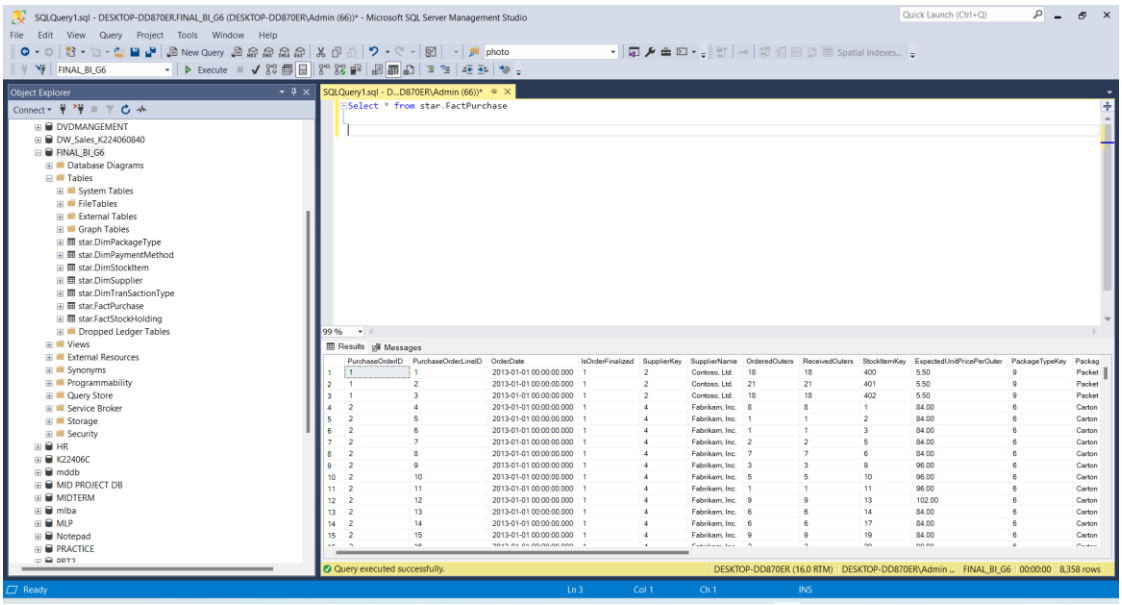

Figure 4.21. FactPurchase Table's check

Source: The author synthesizes suggestions

4.2.2.2. ETL process of FactStockHolding Table

The below figure shows the ETL process of the FactStockHolding Table:

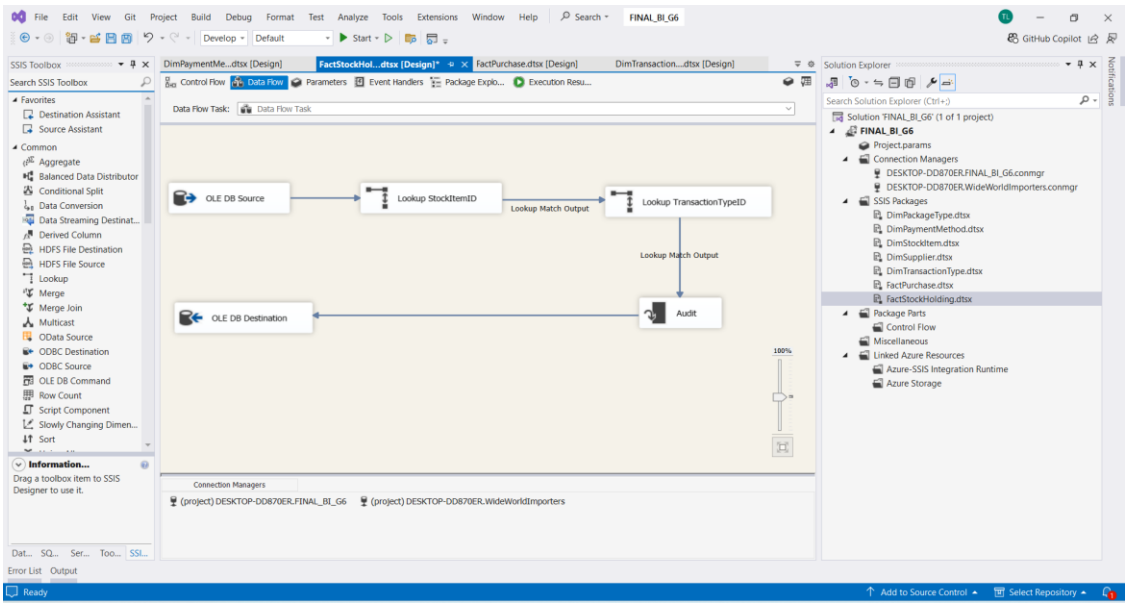

Figure 4.22. FactStockHolding Table's ETL Process

Source: The author synthesizes suggestions

After setting the link, we conduct the execution of FactStockHolding Package and the below figure demonstrates the successful package execution.

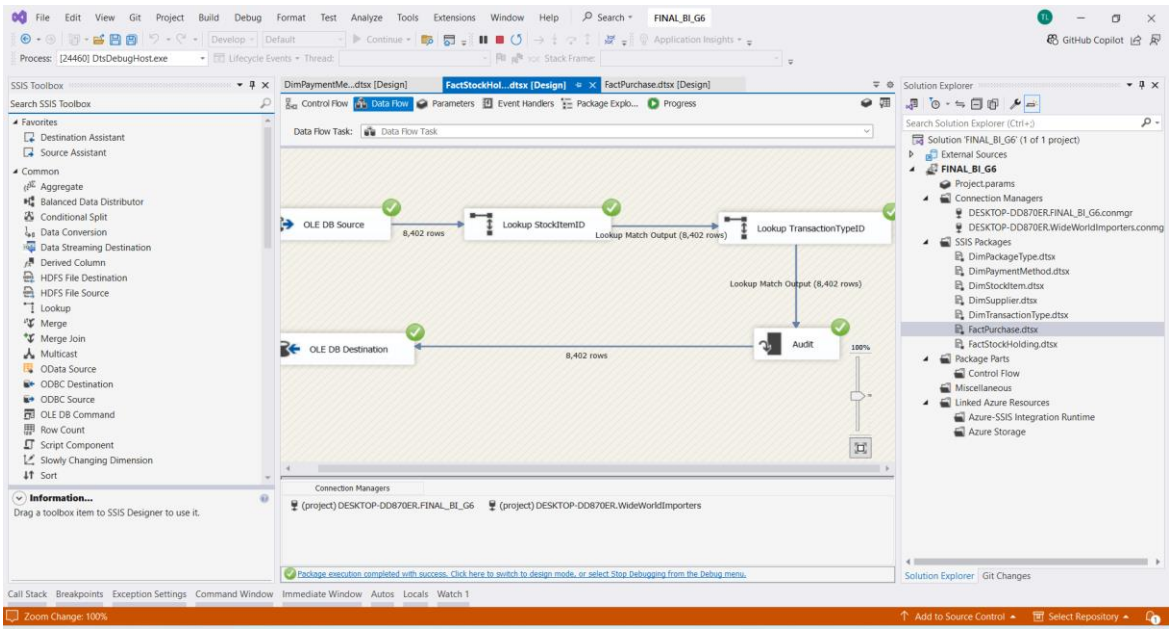

Figure 4.23. FactStockHolding Table's successful package execution

Source: The author synthesizes suggestions

After that, we continually use the SQL Server to check the availability of converted data into the FactStockHolding table and get the results of 8,402 rows matching the above SSIS process.

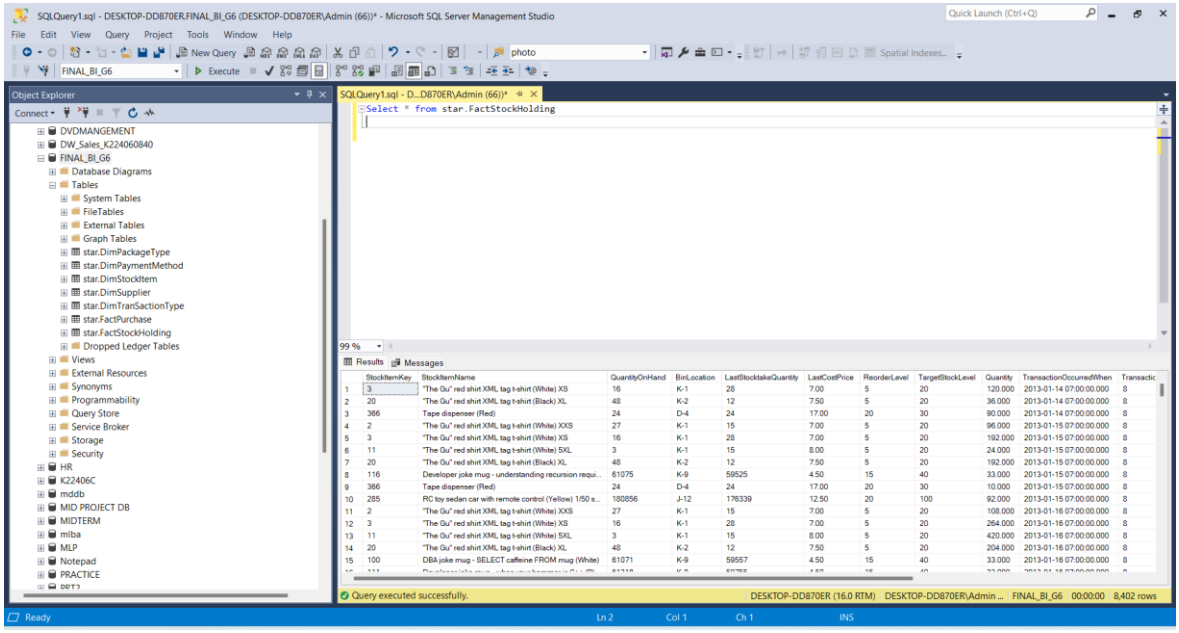

Figure 4.24. FactStockHolding Table's check

Source: The author synthesizes suggestions

CHAPTER 5: RESULTS – DATA ANALYTICS AND VISUALIZATION

5.1 Data analysis

5.1.1. Data analysis with SSAS

To analyze data from Data Warehouse, the project uses SSAS (SQL Server Analysis Service) tool, which is an online analytical processing (OLAP) tool and data mining in Microsoft SQL Server.

The 5 steps to install and analyze data using SSAS technology will be described briefly below:

- *Step 1:* Create a SSAS project (choosing Analysis Service Multidimensional Project).
- *Step 2:* Create Data Source is the data source for building data cube, here is the previously built SQL data warehouse Sale DW.
- *Step 3:* Create a Data Source View, add the tables one by one dimension and facts table contained in the datastore. After successful creation, we will see the same diagram as in Data Warehouse.
- *Step 4:* Create a cube.
- *Step 5:* Linking columns in the Dimension Table.

Next, the project's data dump is depicted in the diagram below (Figure 5.1):

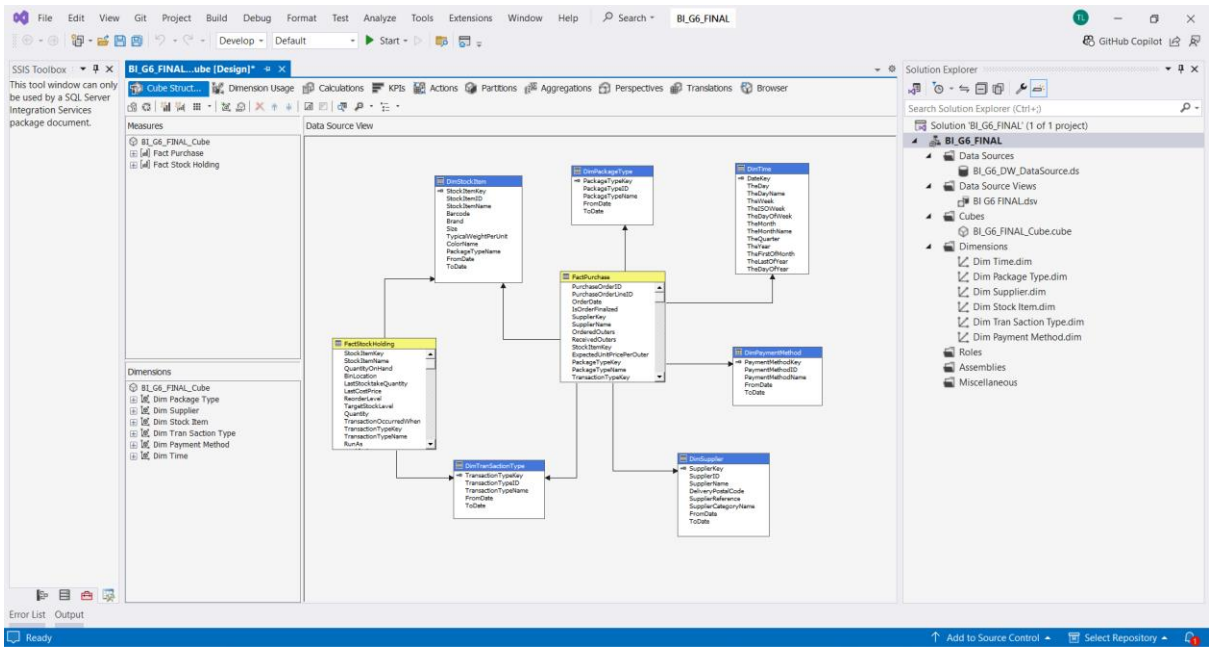

Figure 5.1. Data analysis with SSAS

Source: The author synthesizes suggestions

5.1.2. Build KPIs

In addition to data analysis from the DW data warehouse, the topic offers administrators key performance indicators (KPIs) to assess the efficiency of business operations.

Relying on the dataset analyzed, we realize that there were 7 suppliers having selling transactions with the company instead of a total of 13 suppliers as the dataset provided.

	SupplierName	total paid	the number of transaction
1	Graphic Design Institute	56236.50	42
2	A Datum Corporation	54891.00	10
3	Fabrikam, Inc.	4068204765.50	5796
4	Northwind Electric Cars	337614.00	18
5	The Phone Company	141863.00	15
6	Contoso, Ltd.	940.50	3
7	Litware, Inc.	776144551.70	2474

Figure 5.2. List of Suppliers having transaction with the company

Source: The author synthesizes suggestions

The following figure demonstrates the KPI Supplier Report of WideWorldImporter, showing the expense KPI that the company had to pay for each supplier over years. Here, we would like to propose some discount levels in each year with suppliers, which are strategic, potential, new, etc. in order to decrease the cost of the company. As a result, with Value Expression is total paid amount of year, the Goal Expression will be Value Expression * 0.95 if the supplier is Fabrikam, Inc. (with the highest number of transactions and total paid amount), it will be Value Expression * 0.97 if the supplier is Litware, Inc. which had the second highest number and amount of total transactions, and it will be Value Expression * 0.98 if the supplier is Northwind Electric Cars having a high value of total paid amount. Besides, with the rest of supplier , it will be Value Expression * 1.00 to maintain the relationship with the new ones, from which we will monitor the status of whether the suppliers' costs meeting KPI or not (Status Expression) and that expenses' growth trend is up or down (Trend Expression).

Row Labels	KPI_TotalPaidAmount	KPI_TotalPaidAmount Goal	KPI_TotalPaidAmount Status	KPI_TotalPaidAmount Trend
Clothing				
Fabrikam, Inc.				
2011			◆	→
2012			◆	→
2013	318443199.5	302521039.5	●	↓
2014	1005618732	955337795.4	●	↓
2015	1771776312	1683187496	●	↓
2016	972366522	923748195.9	●	↑
2017			◆	↑
2018			◆	→

Row Labels	KPI_TotalPaidAmount	KPI_TotalPaidAmount Goal	KPI_TotalPaidAmount Status	KPI_TotalPaidAmount Trend
Computing Novelties				
Graphic Design Institute				
2011			◆	→
2012			◆	→
2013	56236.5	56236.5	●	↓
2014			◆	↑
2015			◆	→
2016			◆	→
2017			◆	→
2018			◆	→
The Phone Company				
2011			◆	→
2012			◆	→
2013	141863	141863	●	↓
2014			◆	↑
2015			◆	→
2016			◆	→
2017			◆	→
2018			◆	→

Row Labels	KPI_TotalPaidAmount	KPI_TotalPaidAmount Goal	KPI_TotalPaidAmount Status	KPI_TotalPaidAmount Trend
Novelty Items				
A Datum Corporation				
2011				
2012				
2013				
2014				
2015				
2016	54891	54891		
2017				
2018				
Northwind Electric Cars				
2011				
2012				
2013	337614	330861.72		
2014				
2015				
2016				
2017				
2018				

Row Labels	KPI_TotalPaidAmount	KPI_TotalPaidAmount Goal	KPI_TotalPaidAmount Status	KPI_TotalPaidAmount Trend
Packaging Materials				
Litware, Inc.				
2011				
2012				
2013	18735297.2	18173238.28		
2014	159102853.4	154329767.8		
2015	381259366.7	369821585.7		
2016	217047034.4	210535623.4		
2017				
2018				
Toys				
Contoso, Ltd.				
2011				
2012				
2013	940.5	940.5		
2014				
2015				
2016				
2017				
2018				
Grand Total	4844940862	4844940862		

Figure 5.3. KPI Supplier Report

Source: The author synthesizes suggestions

5.2. Visualization with Power BI

5.2.1. Warehouse

5.2.1.1. Warehouse Overview

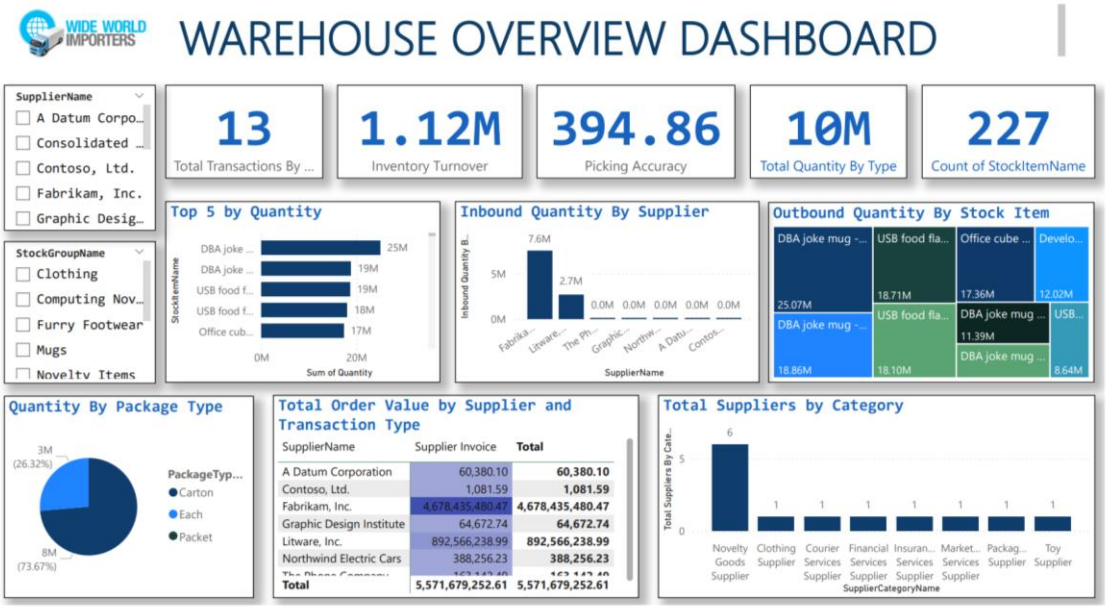

Figure 5.4. Warehouse Overview Dashboard

Sources: The author synthesizes suggestions

The dashboard gives overview figures for Wide World Importers dataset, which provides insights into various aspects of warehouse operations. It visualizes key metrics related to ***total transactions by type, inventory turnover, picking accuracy, total quantity by type and count of stock item name.***

Four cards combined with slicer provide an overview of Total Transactions By Type, Inventory Turnover, Picking Accuracy, Total Quantity By Type and Total Suppliers:

- ***Total Transactions By Type:*** This metric represents the total number of transactions recorded in the system for a specific period. It might refer to the number of different types of transactions, such as *Customer Contra, Customer Credit Note, Customer Invoice, Customer Payment Received, Customer Refund,*

Stock Adjustment at Stocktake, Stock Issue, Stock Receipt, Stock Transfer, Supplier Credit Note, Supplier Invoice, Supplier Payment Issued, Supplier Refund.

- ***Inventory Turnover:*** Inventory turnover measures how often inventory is sold and replaced over a specific period. It is calculated as the cost of goods sold divided by average inventory. A turnover of *1.12M* suggests that the inventory is cycling through at a rate of *1.12 million times* during the period. High turnover rates can indicate efficient inventory management and strong sales, while lower rates might suggest overstocking or weak sales.
- ***Picking Accuracy:*** Picking accuracy reflects the accuracy rate of the picking process within the warehouse. It indicates how many picks were correctly performed out of the total picks made. A high picking accuracy, such as 394.86, suggests that the warehouse operations are efficient, with few errors occurring in the picking process. This can lead to higher customer satisfaction and reduced returns or rework.
- ***Total Quantity By Type:*** This metric represents the total quantity of items categorized by type within the warehouse. A total quantity of 10 million indicates the volume of stock the warehouse manages. This figure helps assess the scale of operations and the complexity of inventory management.
- ***Count of Stock Item Name:*** The metric "Count of Stock Item Name" indicates that the warehouse manages 227 distinct types of products. This diversity suggests a complex inventory system, requiring efficient organization and management. It reflects the warehouse's ability to handle a broad range of items, supporting various markets or customer needs.

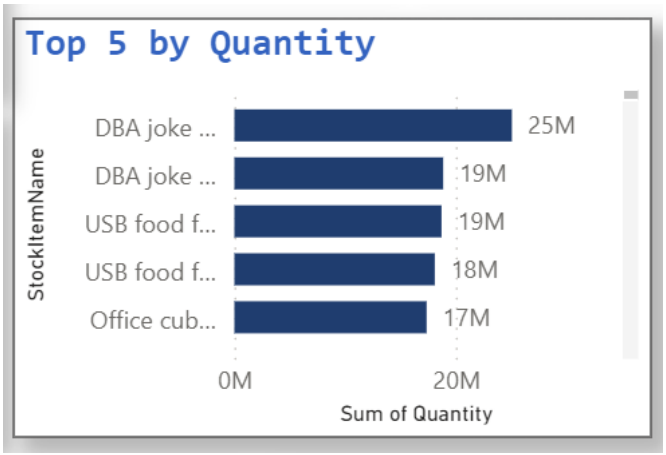

Figure 5.5. Top 5 by Quantity

Sources: The author synthesizes suggestions

Figure 5.5 shows the top five stock items by quantity, highlighting which products have the highest inventory levels or sales volume.

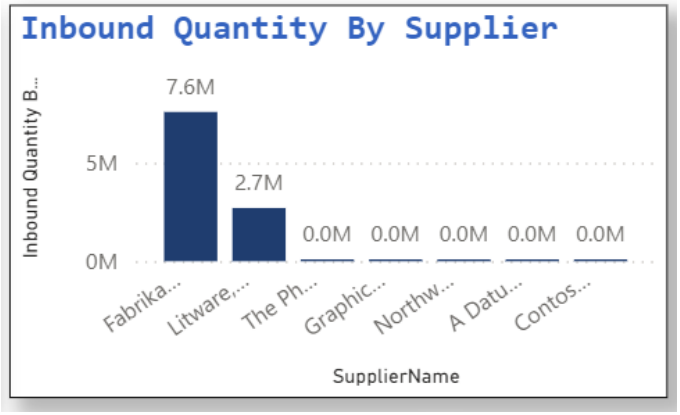

Figure 5.6. Inbound Quantity By Supplier

Sources: The author synthesizes suggestions

Figure 5.6 displays the quantities of goods received from various suppliers, indicating which suppliers contribute the most to the warehouse's inventory. The dominance of Fabrikam, Inc., suggests a reliance on this supplier for a significant portion of inventory, highlighting the importance of maintaining a strong partnership with them.

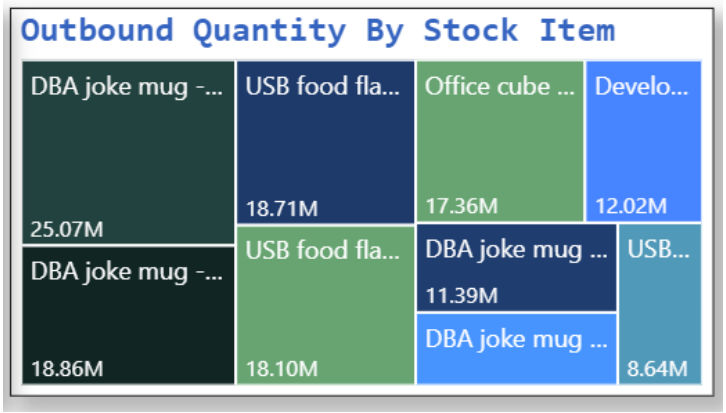

Figure 5.7. Outbound Quantity By Stock Item

Sources: The author synthesizes suggestions

Figure 5.7 provides a treemap visualization of the quantities of stock items leaving the warehouse, helping identify which items are most frequently shipped. The high volume of DBA joke mugs and USB food flasks suggests these items are popular among customers, indicating potential trends in demand.

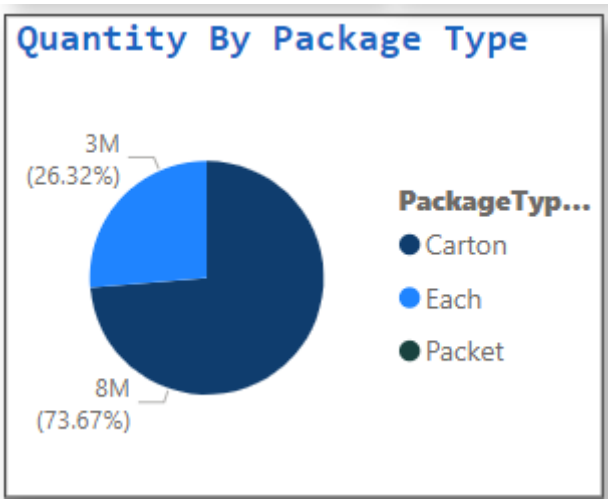

Figure 5.8. Quantity by Package Type

Sources: The author synthesizes suggestions

Figure 5.8 shows the number of products packaged by different packaging types. It provides detailed insight into how products are distributed based on packaging type, helping to analyze and optimize packaging and shipping processes.

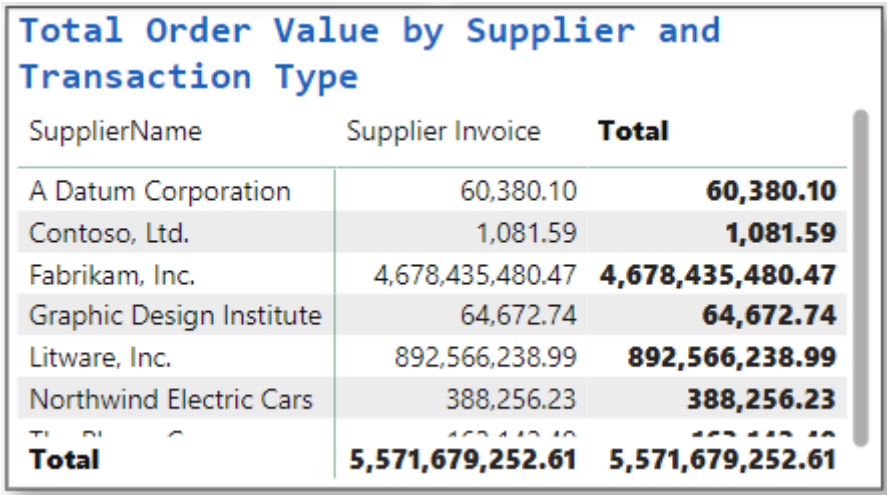

Figure 5.9. Total Order Value by Supplier and TransactionType

Sources: The author synthesizes suggestions

Figure 5.9 displays the financial transactions with suppliers, showing the invoice amounts and total order values for each supplier, providing insight into spending distribution. Key suppliers like Fabrikam, Inc., and Litware, Inc. account for the majority of the expenditure, reflecting their importance in the supply chain. This information is crucial for strategic decision-making regarding supplier relationships and cost management.

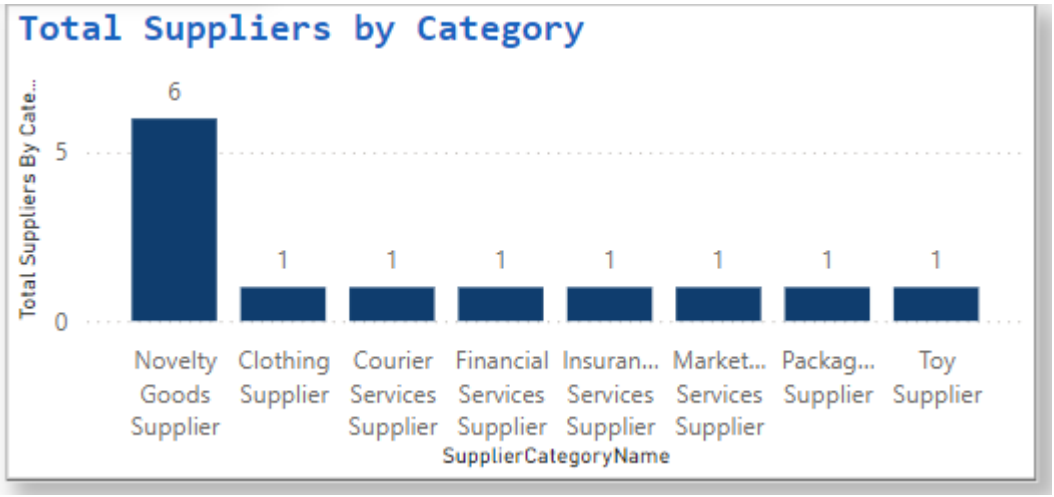

Figure 5.10. Total Suppliers by Category

Sources: The author synthesizes suggestions

Figure 5.10 visualizes total suppliers based on different categories. It helps viewers understand the distribution of suppliers based on the product categories they supply, thereby assisting in analyzing and managing supplier relationships. In the above chart, Novelty Goods Supplier accounted for the largest with 6 units.

5.2.1.2. Warehouse Performance

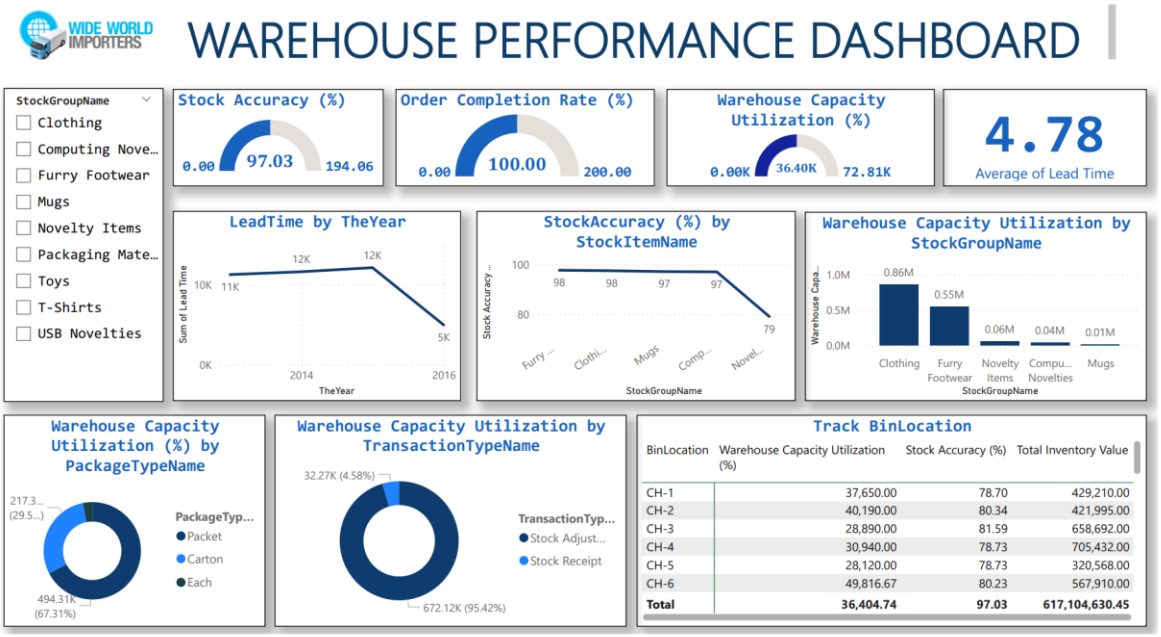

Figure 5.11. Warehouse Performance Dashboard

Sources: The author synthesizes suggestions

The above dashboard gives performance figures for Wide World Importers dataset, containing various key performance indicators and visualizations. It visualizes key metrics related to ***stock accuracy, StockItemKey warehouse capacity, and lead time.***

Four cards combined with slicer provide an overview of Stock Accuracy, Order Fulfillment Rate (%), Warehouse Capacity Utilization (%) and Average of Lead Time:

- ***Stock Accuracy (%):*** This metric indicates the accuracy of the stock levels recorded in the system compared to the physical inventory. A higher percentage signifies better accuracy. In this case, the stock accuracy figure is 97.03%.

- ***Count of StockItemKey:*** This metric counts the number of StockItem.
- ***Warehouse Capacity Utilization (%):*** Warehouse capacity utilization is the measure of the total amount of used capacity in a warehouse. This metric shows how much of the available warehouse space is being used.
- ***Average of Lead Time:*** This metric represents the average time taken from when an order is placed to when it is fulfilled. A card provides a quick summary of the average lead time.

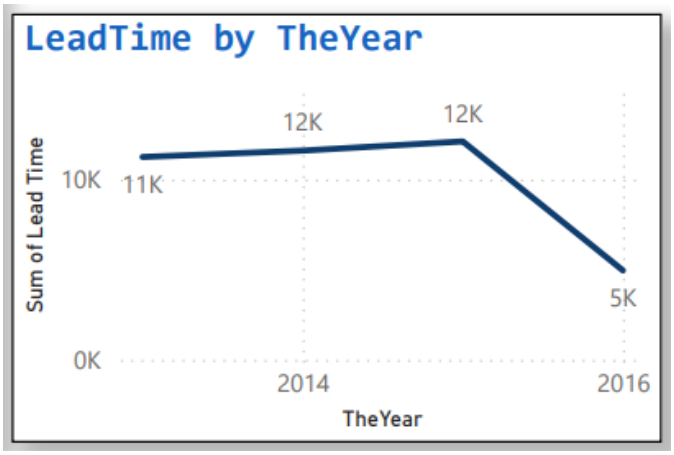

Figure 5.12. Lead Time by TheYear

Sources: The author synthesizes suggestions

Figure 5.12 visualizes the trend of lead time over the years. It helps tracking the order fulfillment speed over years. This is also a criterion for customers to evaluate the effectiveness of the work that the manufacturer performs after the finished product is ready for delivery.

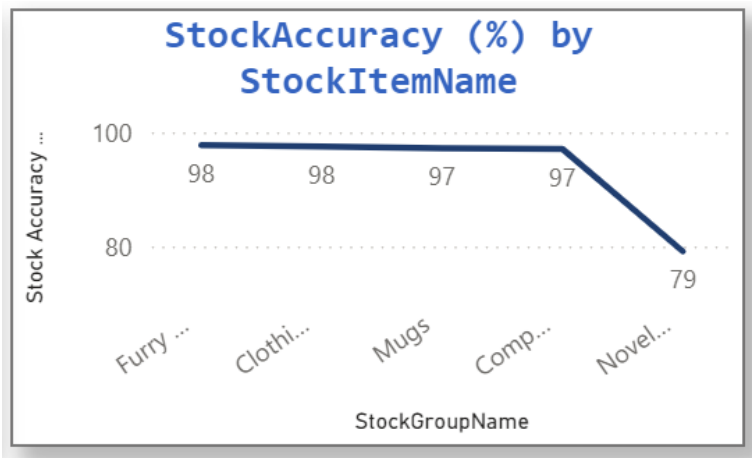

Figure 5.13. Stock Accuracy (%) by StockItemName

Sources: The author synthesizes suggestions

Figure 5.13 provides a detailed breakdown of stock accuracy for each individual stock item. A bar chart allows for easy comparison of accuracy levels across different items.

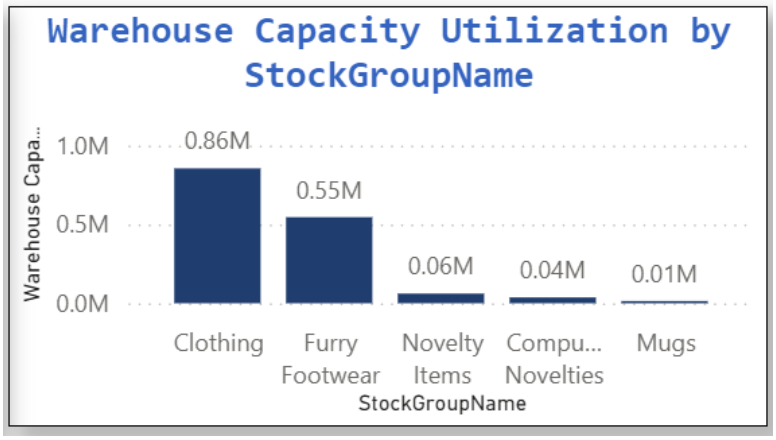

Figure 5.14. Warehouse Capacity Utilization by BinLocation

Sources: The author synthesizes suggestions

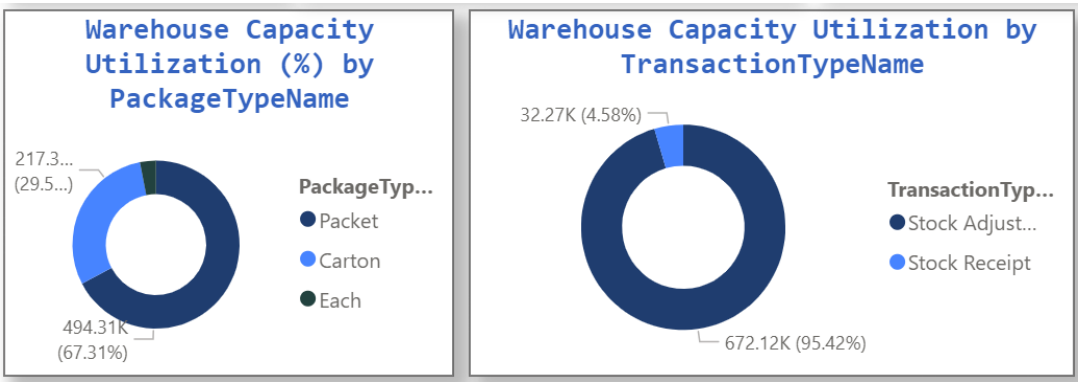

Figure 5.15. Warehouse Capacity Utilization (%) by PackageTypeName and TransactionTypeName

(Sources: The author synthesizes suggestions)

This metric analyzes warehouse capacity utilization based on different package types and transaction types. These charts provide a breakdown of capacity utilization by these categories.

Track BinLocation

BinLocation	Warehouse Capacity Utilization (%)	Stock Accuracy (%)	Total Inventory Value
CH-1	37,650.00	78.70	429,210.00
CH-2	40,190.00	80.34	421,995.00
CH-3	28,890.00	81.59	658,692.00
CH-4	30,940.00	78.73	705,432.00
CH-5	28,120.00	78.73	320,568.00
CH-6	49,816.67	80.23	567,910.00
Total	**36,404.74**	**97.03**	**617,104,630.45**

Figure 5.16. Tracking Bin Location

(Sources: The author synthesizes suggestions)

This matrix provides detailed information about the warehouse capacity utilization, accuracy, and inventory value for each bin location.

5.2.1.3. Inventory Management

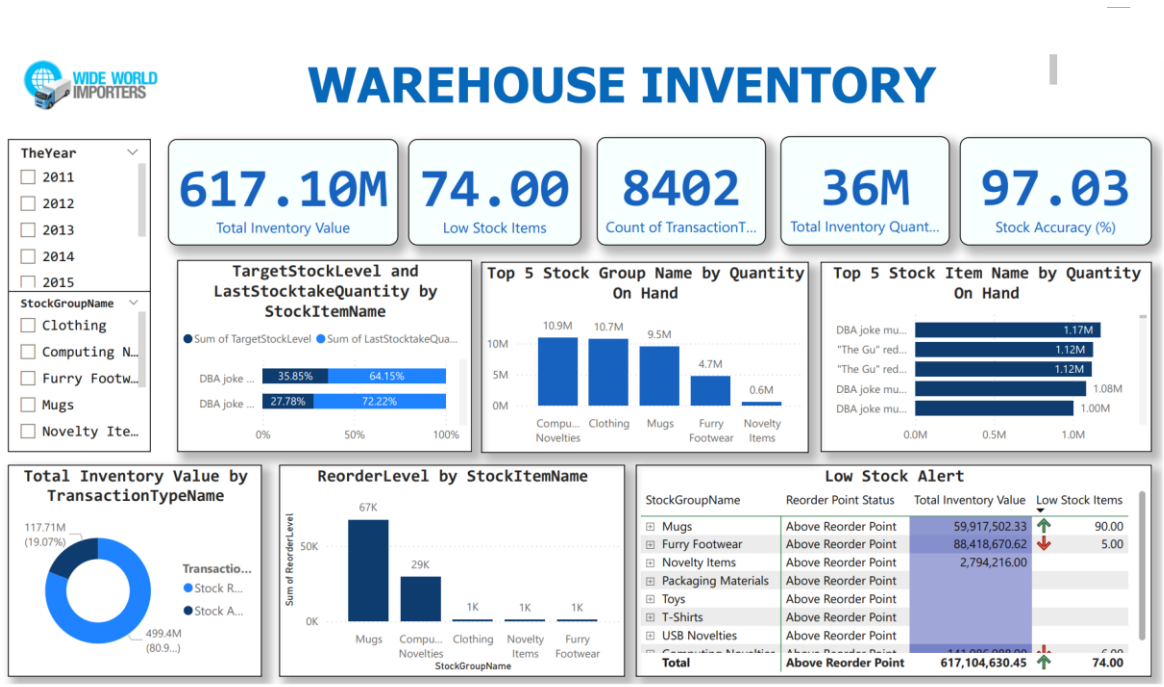

Figure 5.17. Warehouse inventory management

(Sources: The author synthesizes suggestions)

The provided dashboard gives an overview of the warehouse inventory for Wide World Importers, containing various key performance indicators and visualizations.

5 cards combined with slicer provide an overview of Total Inventory Value, Total Inventory Quantity, Low Stock Items, Numbers of Transactions and Stock Accuracy (%):

- ***Total Inventory Value:*** 617.10M. This is the total value of all items in the warehouse, providing an overview of the current assets of the warehouse.
- ***Total Inventory Quantity:*** 36M units. The total number of inventory units, giving an overview of the scale of the warehouse.
- ***Low Stock Items:*** The number of items with stock levels below the threshold, helping warehouse management know which items need to be reordered.
- ***Sum of Total Transactions:*** 8402 transactions. The total number of transactions, indicating the level of activity in the warehouse.
- ***Stock Accuracy (%):*** 97.03%. The percentage accuracy between recorded data and actual counts, indicating the reliability of the warehouse management system.

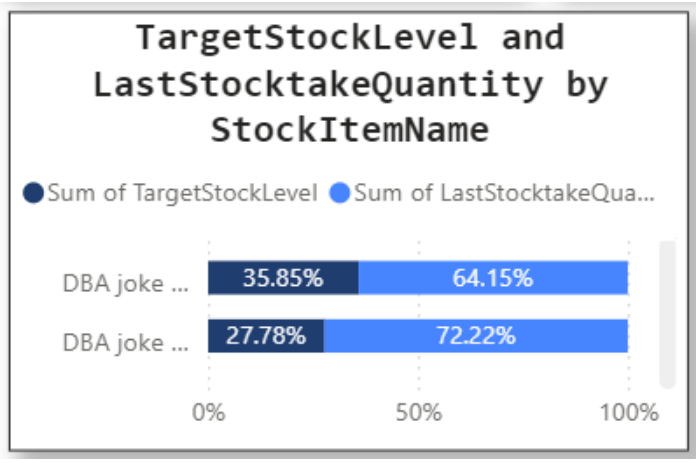

Figure 5.18. Target Stock Level and Last Stocktake Quantity by Stock Item Name

Source: The author synthesizes suggestions

Figure 5.18 illustrates ***Target Stock Level and Last Stocktake Quantity by Stock Item Name.*** A bar chart showing the target stock level and last stocktake quantity for each stock item, helping to compare the target versus the actual quantities.

The above chart helps identify which items are short or over the target inventory level. If the actual inventory is lower than the company's target, they may need to order more. Conversely, if the actual inventory is higher than your target, you may need to adjust your purchasing strategy. Therefore, managers can use information from the chart to make strategic decisions about ordering, inventory optimization, and supply management.

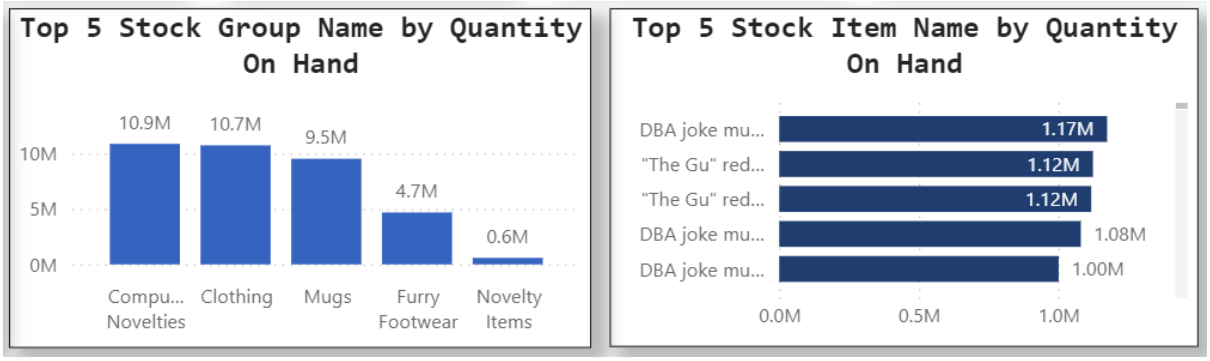

Figure 5.19. Top 5 StockGroupName and StockItemName by QuantityOnHand

Source: The author synthesizes suggestions

Figure 5.19 illustrates the top 5 StockGroupName and StockItemName sorted by QuantityOnHand.

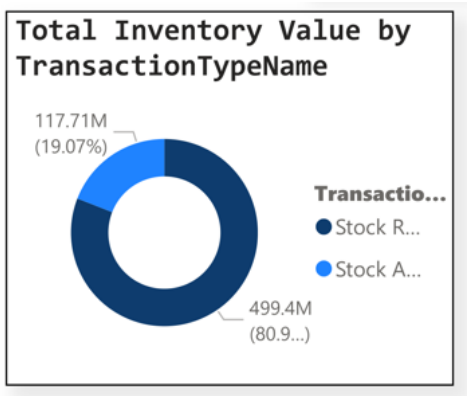

Figure 5.20. Total Inventory Value by TransactionTypeName

Source: The author synthesizes suggestions

Figure 5.20 illustrates the total inventory value divided by transaction types (e.g., Stock Receipt, Stock Adjustment). It provides information on the current asset value of the warehouse based on transaction types.

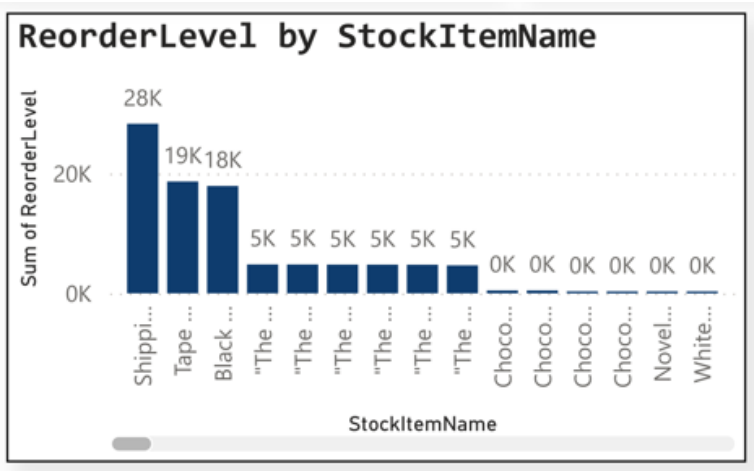

Figure 5.21. ReorderLevel by StockItemName

Source: The author synthesizes suggestions

A bar chart indicating the reorder level for each stock item. It helps identify items that need to be reordered soon to avoid stockouts. This helps ensure that businesses do not face stock outs or excess inventory, both of which can lead to financial losses.

Low Stock Alert

StockGroupName	Reorder Point Status	Total Inventory Value	Low Stock Items
Mugs	Above Reorder Point	59,917,502.33	90.00
Furry Footwear	Above Reorder Point	88,418,670.62	5.00
Novelty Items	Above Reorder Point	2,794,216.00	
Packaging Materials	Above Reorder Point		
Toys	Above Reorder Point		
T-Shirts	Above Reorder Point		
USB Novelties	Above Reorder Point		
Total	**Above Reorder Point**	**617,104,630.45**	**74.00**

Figure 5.22. Low Stock Alert Tracking

Source: The author synthesizes suggestions

A table displaying stock items that are below their reorder points along with their reorder status and total inventory value. It provides crucial information for warehouse management to take timely actions to maintain adequate stock levels.

5.2.2. Purchasing

5.2.2.1. Purchasing Overview

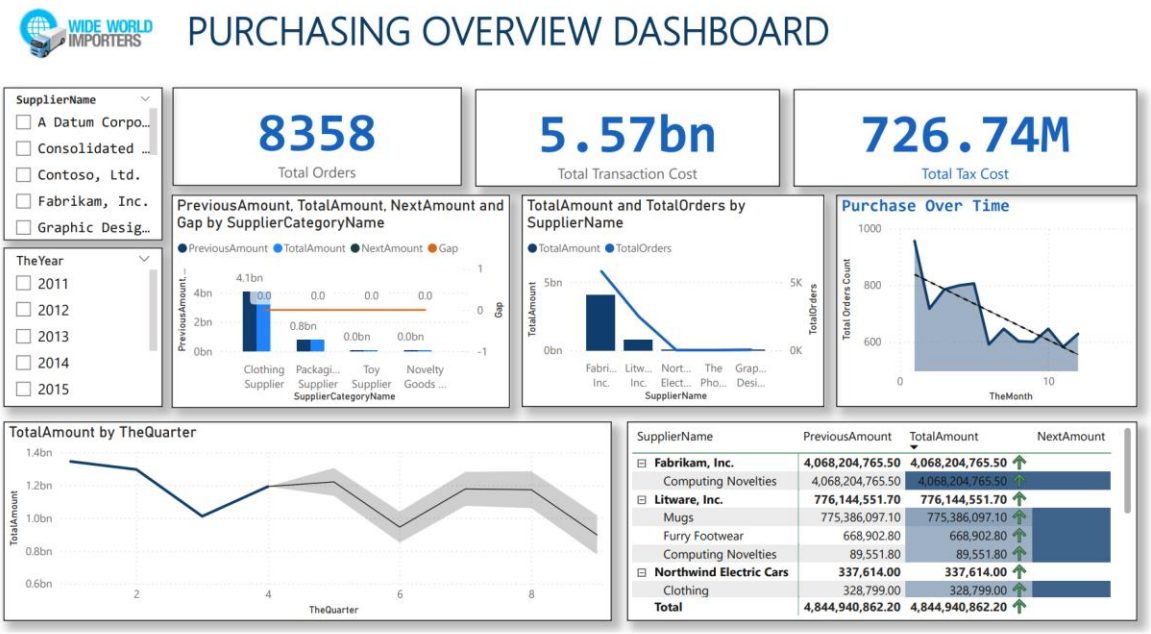

Figure 5.23. Purchasing Overview Dashboard

Source: The author synthesizes suggestions

This is the Purchasing Overview Dashboard of Wide World Importers. The cards in the dashboard represent:

- ***Total Orders****:* Displays the total number of orders placed (8,358 orders). Helps the business assess the volume of purchasing activity and business performance.
- ***Total Transaction Cost****:* The total transaction cost is 5.57 billion. The business can use this information to analyze and control costs, ensuring financial efficiency.
- ***Total Tax Cost****:* The total tax cost is 726.74 million. Helps the business monitor tax costs and manage tax obligations effectively.

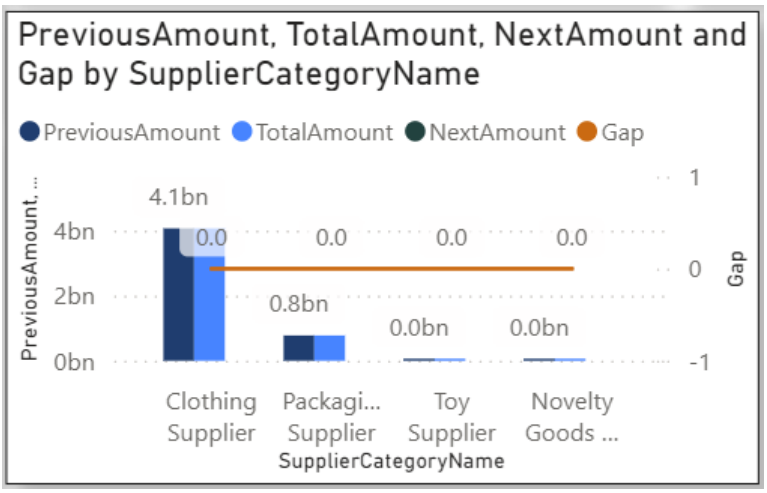

Figure 5.24. PreviousAmount, TotalAmount, NextAmount and Gap by SupplierCategoryName

Source: The author synthesizes suggestions

The chart shows the previous amount, total amount, next amount, and gap by supplier category. Clothing Supplier has a previous amount of 4.1 billion, while other categories have significantly lower amounts or zero. This helps the business compare and evaluate the performance of each supplier category.

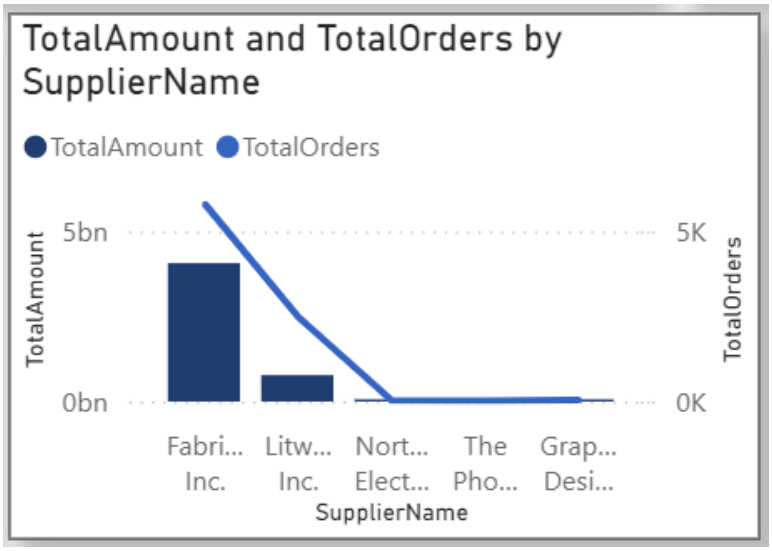

Figure 5.25. TotalAmount and TotalOrders by SupplierName

Source: The author synthesizes suggestions

This chart shows the total amount and total orders by supplier name. Fabrikam, Inc. has the highest total amount of approximately 5 billion and a lower total number of

orders, indicating high-value orders. The business can identify important suppliers and evaluate their performance based on transaction amount and order quantity.

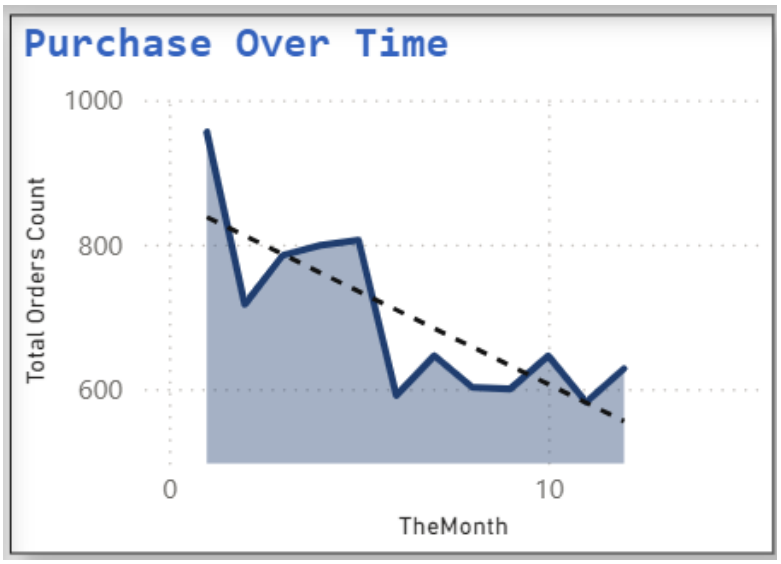

Figure 5.26. Purchase Over Time

Source: The author synthesizes suggestions

The number of orders (Total Orders Count) by month (The Month) is shown in the chart. There is a general downward trend in the number of orders over time, with some months showing slight increases. This helps the business recognize trends and fluctuations in purchasing activity, allowing them to adjust purchasing or promotional strategies.

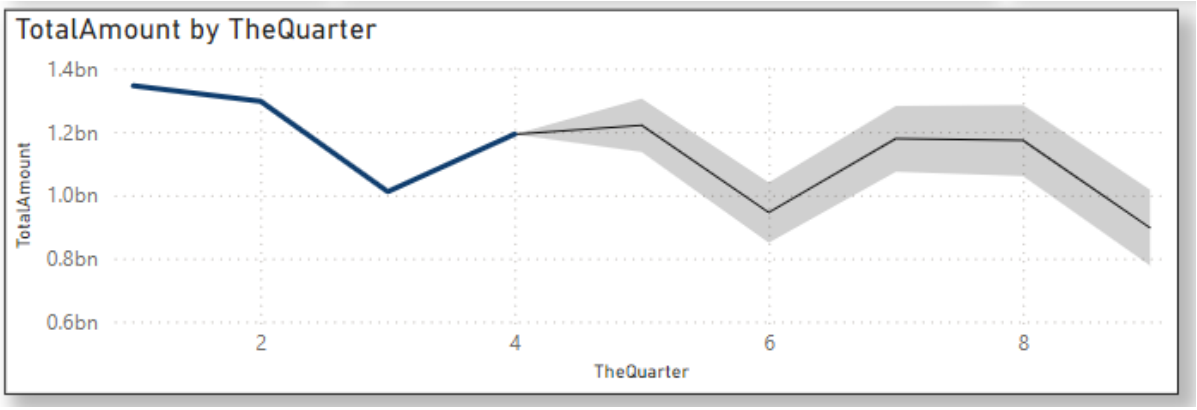

Figure 5.27. TotalAmount by TheQuarter

Source: The author synthesizes suggestions

This figure chart illustrates the TotalAmount across quarters from Q1 to Q8, with values ranging from 0.6 billion to 1.4 billion. The line shows the changes in TotalAmount over the quarters, starting at 1.3 billion in Q1, dropping to 1.0 billion in Q3, and rising again in the following quarters. The shaded gray area around the line represents the confidence interval or variability of TotalAmount across the quarters.

SupplierName	PreviousAmount	TotalAmount	NextAmount
Fabrikam, Inc.	**4,068,204,765.50**	**4,068,204,765.50**	
Computing Novelties	4,068,204,765.50	4,068,204,765.50	
Litware, Inc.	**776,144,551.70**	**776,144,551.70**	
Mugs	775,386,097.10	775,386,097.10	
Furry Footwear	668,902.80	668,902.80	
Computing Novelties	89,551.80	89,551.80	
Northwind Electric Cars	**337,614.00**	**337,614.00**	
Clothing	328,799.00	328,799.00	
Computing Novelties	8,815.00	8,815.00	
Total	**4,844,940,862.20**	**4,844,940,862.20**	

Figure 5.28. Previous Amount, Total Amount and NextAmount by SupplierName and StockGroupName

Source: The author synthesizes suggestions

The table shows supplier names, previous amounts, total amounts, and next amounts. Fabrikam, Inc. has the same previous and current total amount, indicating no change. Other suppliers like Litware, Inc., and Northwind Electric Cars also show minimal changes in amounts. This helps the business monitor changes and make decisions about maintaining or replacing suppliers.

5.2.2.2. Supplier Performance

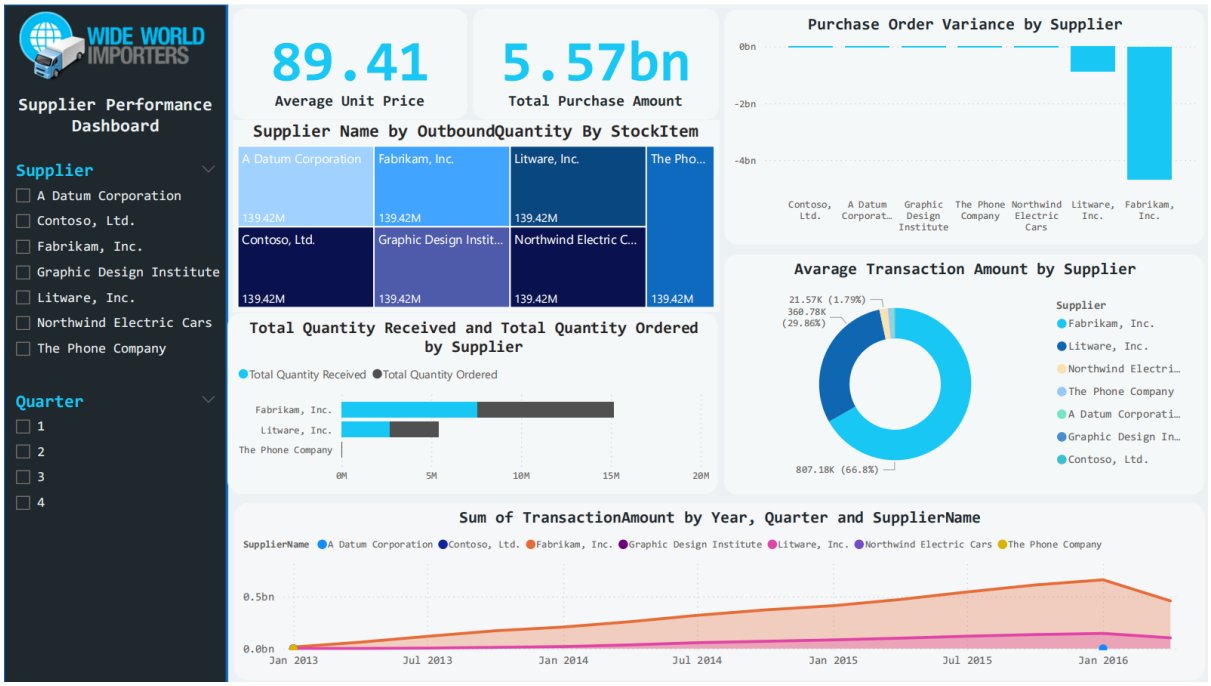

Figure 5.29. Supplier Performance Dashboard

Source: The author synthesizes suggestions

Figure 5.28 provides a comprehensive overview of supplier performance through visualizations. This data helps Wide World Importers evaluate and compare the performance of different suppliers, thereby making strategic decisions and optimizing the supply chain.

Figure 5.30. Average Unit Price and Total Purchase Amount

Source: The author synthesizes suggestions

The average unit price is 89.41 and the total purchase value is 5.57 billion as shown in Figure 5.29 This information provides Wide World Importers with an

overview of their spending size and economic efficiency, helping to evaluate overall spending and plan financial plans appropriately.

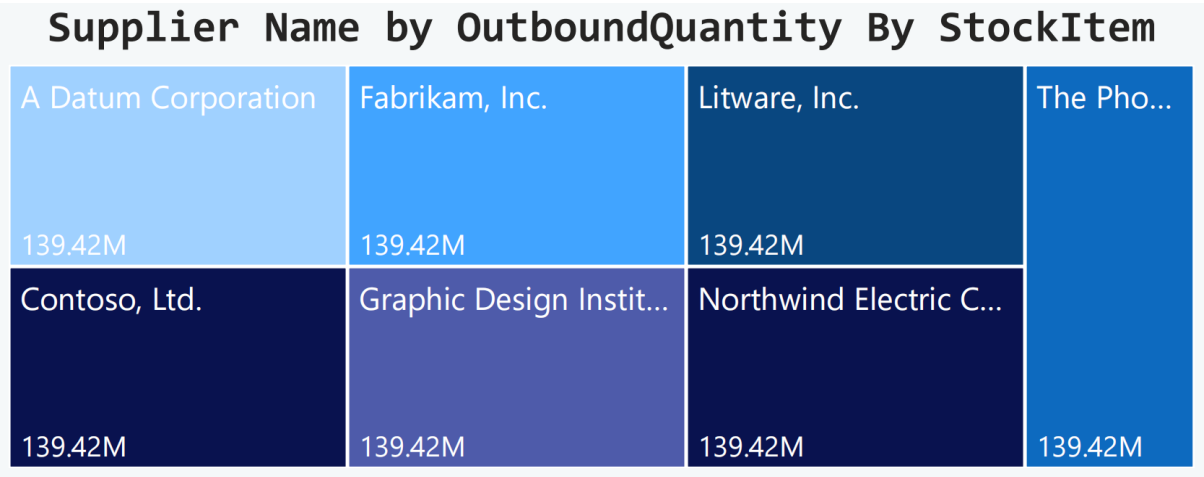

Figure 5.31. Supplier Name by Outbound Quantity

Source: The author synthesizes suggestions

Figure 5.30 shows the outbound quantity of goods for each supplier (Supplier Name) for a specific product (Stock Item). All suppliers have the same outbound quantity of 139.42 million units. This indicates that suppliers provide an equivalent quantity of goods for this product and the market for this product is highly competitive. Although the current suppliers provide an equivalent quantity of goods, the business should consider diversifying its supply sources to reduce risks and increase bargaining power.

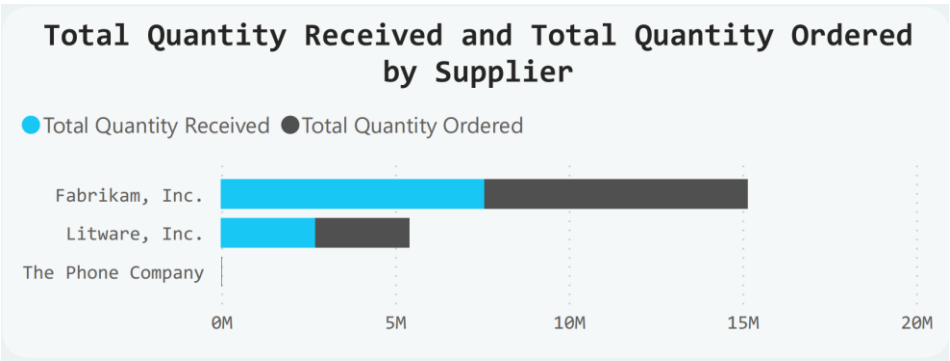

Figure 5.32. Total Quantity Received and Total Quantity Ordered

Source: The author synthesizes suggestions

The total quantity of goods received and the total quantity ordered from each supplier are compared. Figure 5.31 helps Wide World Importers evaluate suppliers' reliability based on their ability to deliver as required. Businesses can review contracts with underperforming suppliers and strengthen partnerships with highly reliable suppliers to ensure a stable supply.

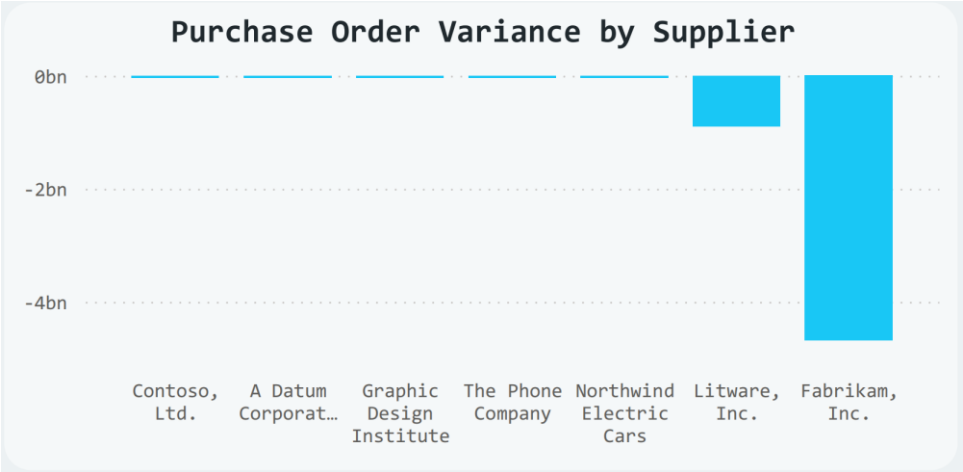

Figure 5.33. Purchase Order Variance

Source: The author synthesizes suggestions

Figure 5.32 shows the difference between ordered quantity and received quantity for each supplier. A large negative value indicates that the supplier often delivers less than ordered. Wide World Importers can reassess and negotiate with suppliers with significant fluctuations to improve delivery accuracy and seek new suppliers if current ones do not meet requirements.

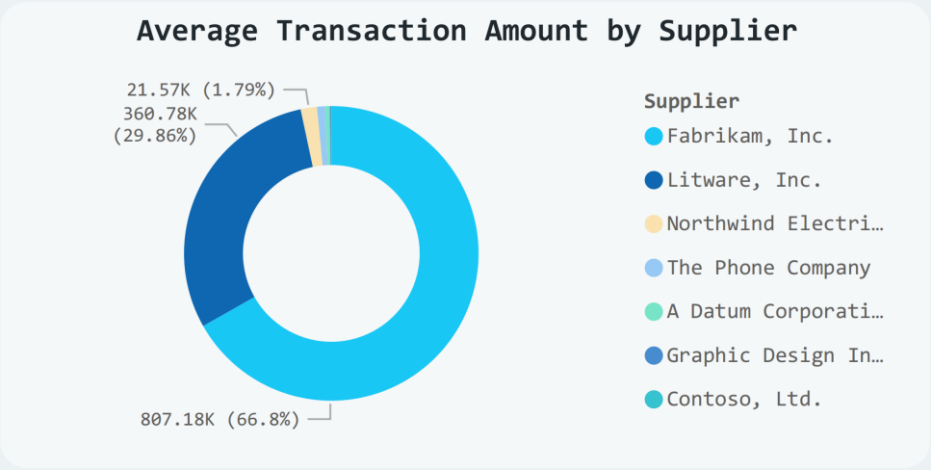

Figure 5.34. Average Transaction Amount

Source: The author synthesizes suggestions

Figure 5.33 illustrates the percentage contribution of each supplier to the total transaction value of Wide World Importers. The chart provides a clear view of how the transaction value is distributed among the suppliers. With a rate of 66.8%, Contoso, Ltd. is the supplier contributing the most to Wide World Importers' transaction value. The company should pay special attention to its partnership with this supplier. Fabrikam, Inc. and Litware, Inc. also account for a significant portion of the total transaction value.

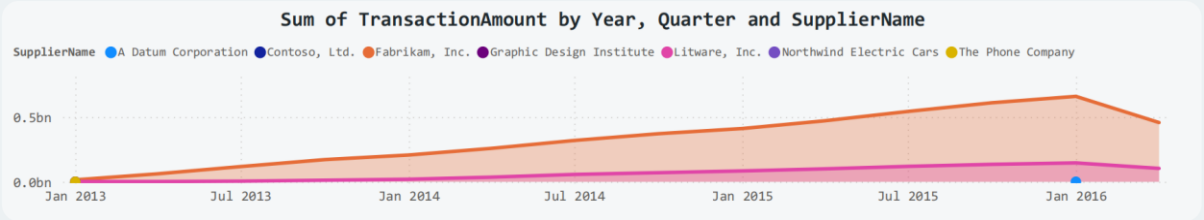

Figure 5.35. Sum of Transaction Amount

Source: The author synthesizes suggestions

The chart shows the change in total transaction value over time (each quarter from early 2013 to early 2016) and is divided by supplier. Figure 5.34 shows the overall growth trend of the total transaction value across quarters. This indicates that Wide World Importers is developing and expanding its operations. The supplier Fabrikam, Inc. has the highest and fastest-growing curve, indicating that it is the largest and most important contributor to the company's revenue. Wide World Importers should focus on strengthening its relationship with Fabrikam, Inc. by increasing transactions, expanding cooperation to new products/services, or prioritizing Fabrikam, Inc. in major projects.

The "Supplier Performance" dashboard provides detailed and clear insight into supplier performance. Wide World Importers can use this information to optimize their purchasing strategies, ensure stable and efficient supply, thereby improving performance and reducing costs in the supply chain.

5.2.2.3. Purchase Order Analysis

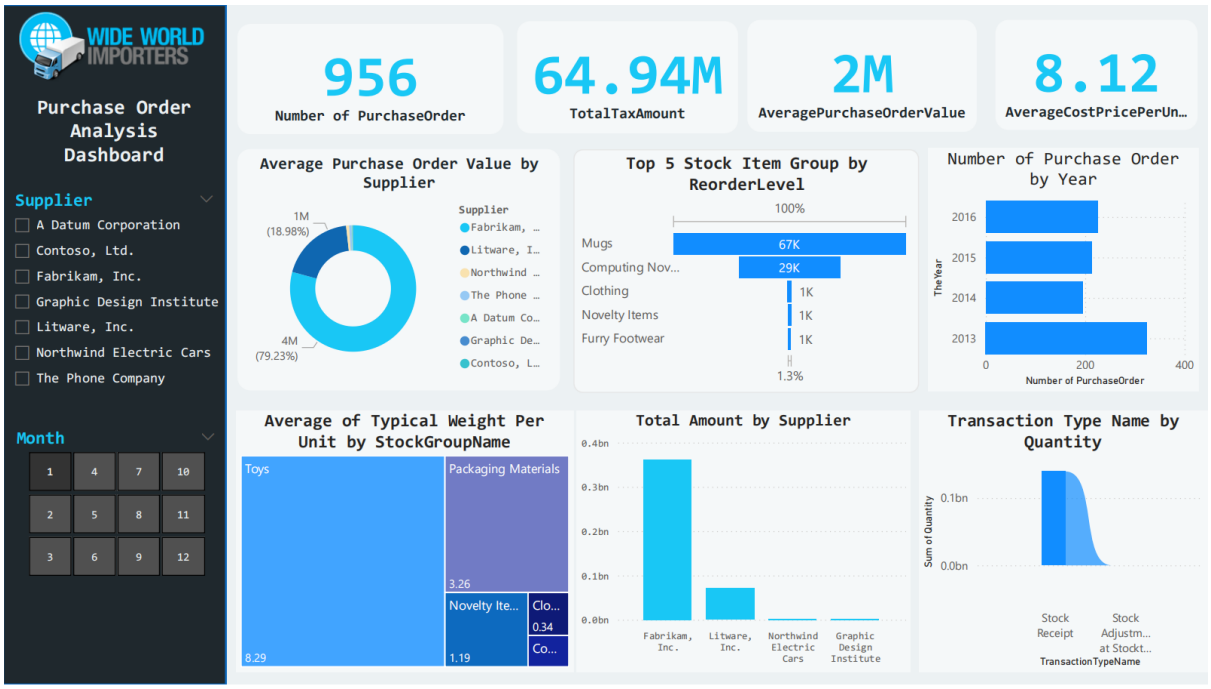

Figure 5.36. Purchase Order Analysis Dashboard

Source: The author synthesizes suggestions

Figure 5.35 analyzes Purchase Orders, providing a comprehensive overview of the financial metrics related to the company's purchase orders. It includes charts and key indicators that help users track and analyze trends, costs, and values related to suppliers. The main metrics presented include average purchase order value by supplier, total tax amount, average cost price per unit…

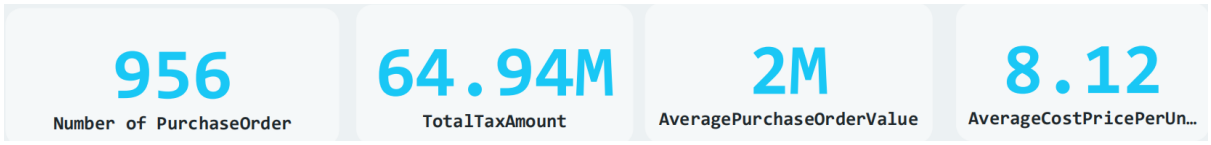

Figure 5.37. Purchase Order Financial Metrics

Source: The author synthesizes suggestions

Figure 5.36 provides summary metrics:

- ***Number of Purchase Orders****:* 956 orders, indicating a fairly active purchasing activity.

- ***Total Tax Amount:*** The total tax amount is 64.94 million. The business can use this information to monitor tax expenses and plan financial strategies accordingly.
- ***Average Purchase Order Value:*** The average purchase order value is 2 million. This helps the business evaluate the average value of each purchase order, improving negotiations with suppliers and managing cash flow more effectively.
- ***Average Cost Price Per Unit:*** The average cost price per unit is 8.12. This information helps the business assess the average cost per unit of goods and identify measures to reduce costs if necessary.

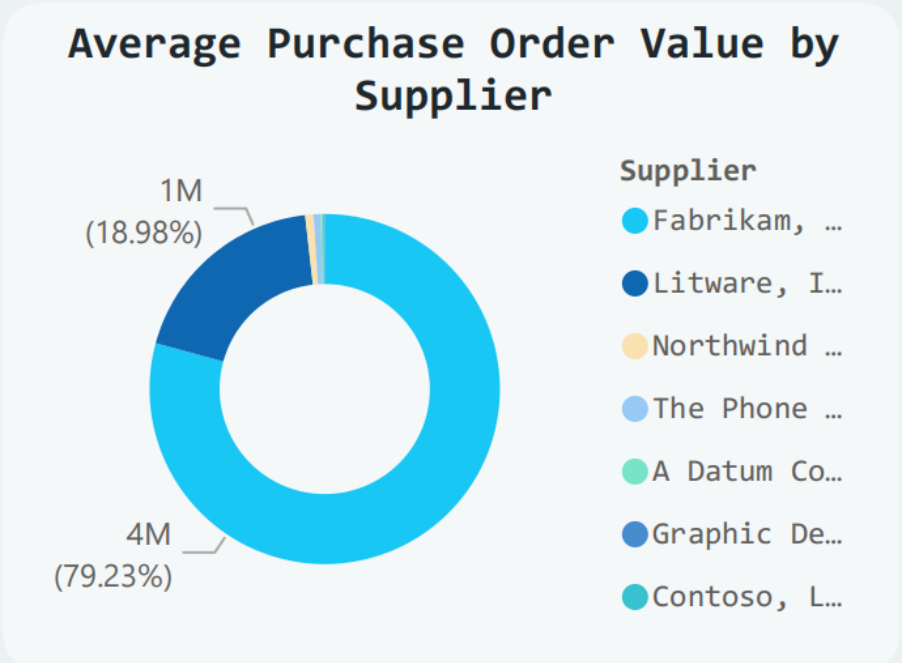

Figure 5.38. Average Purchase Order Value

Source: The author synthesizes suggestions

This chart shows the average order value for each supplier. Figure 5.37 helps the business understand the financial relationship with each supplier and identify key suppliers. Analyzing this chart assists the business in making strategic decisions about maintaining or altering supplier relationships.

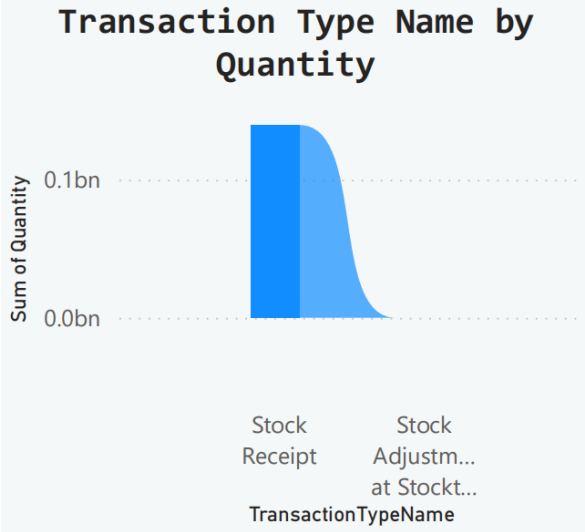

Figure 5.39. Total Quantity by TransactionTypeName

Source: The author synthesizes suggestions

Figure 5.38 illustrates the total quantity by transaction type, including stock receipt and stock adjustment. The business can monitor and manage warehouse activities to ensure efficient stock intake and adjustment processes.

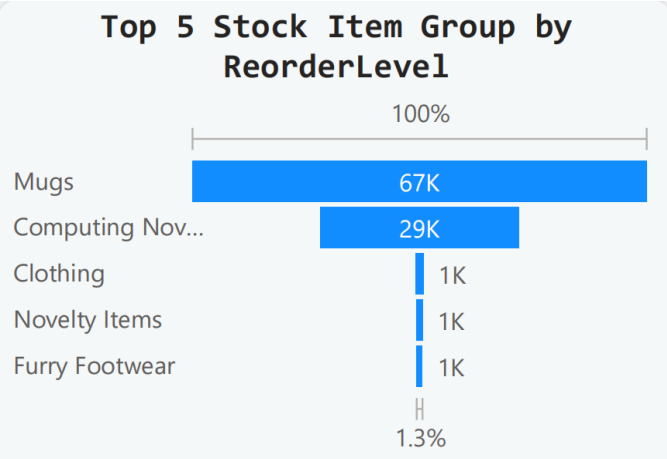

Figure 5.40. Top 5 Stock Groups sorted by ReorderLevel

Source: The author synthesizes suggestions

The bar chart displays reorder levels by stock group. Mugs have the highest reorder level at 67,000, followed by Computing Novelties at 29,000. The business can use this information to maintain appropriate stock levels and avoid shortages.

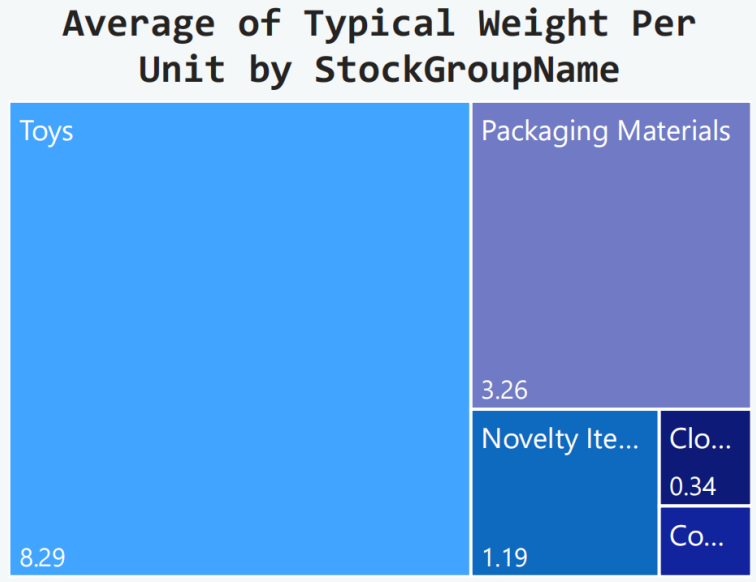

Figure 5.41. Average of Typical Weight

Source: The author synthesizes suggestions

The rectangular chart shows the average weight of each stock group. Toys have the highest average weight at 8.29, while Clothing has the lowest at 0.34. This information helps the business manage and transport goods more efficiently based on weight.

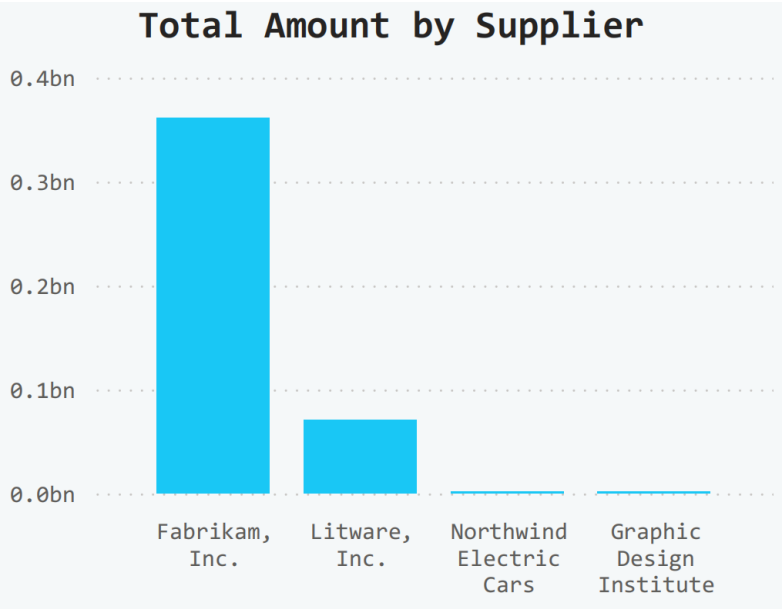

Figure 5.42. Total Amount

Source: The author synthesizes suggestions

Figure 5.42 clearly shows that:

- Fabrikam, Inc. is the leading supplier: It contributes the largest volume of goods, accounting for almost the entire total volume.
- The remaining suppliers contribute very small amounts: Litware, Inc. is second but supplies only one-third of Fabrikam, Inc.'s volume. Northwind Electric Cars and Graphic Design Institute supply almost negligible amounts.

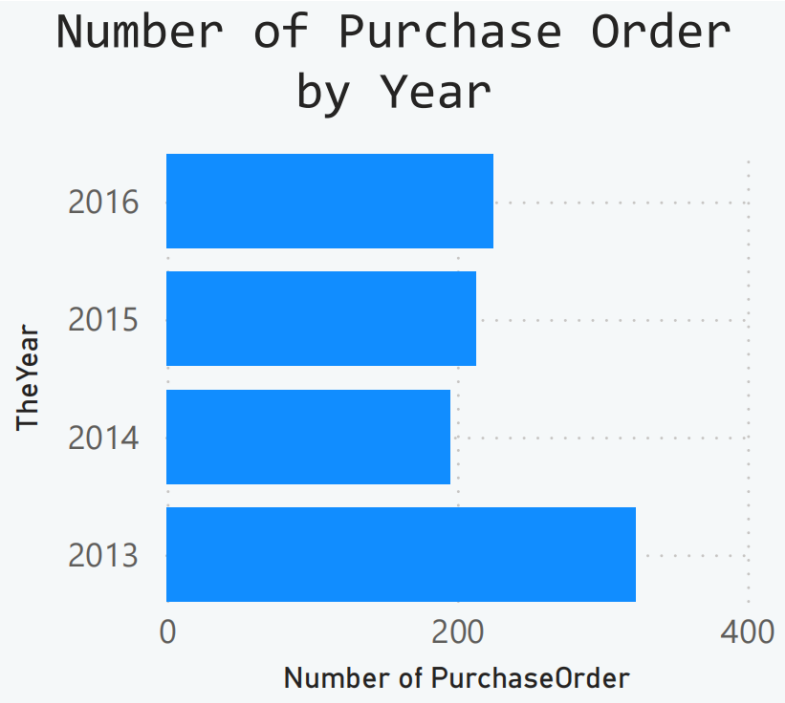

Figure 5.43. Number of Purchase Order by Year

Source: The author synthesizes suggestions

Number of orders by year: There is a gentle growth in the number of orders over the years. It is necessary to further analyze order value and product structure each year to evaluate development trends.

The details and charts in this dashboard offer crucial insights for the business to better understand supplier relationships, manage costs, and plan finances effectively. This data helps the business make informed decisions to optimize costs, manage cash flow, and improve operational efficiency.

5.3. Time series and Forecasting model

Our team utilizes the SARIMA forecasting model to predict the company's purchasing demand.

5.3.1. Overview of forecasting model SARIMA

SARIMA - Seasonal Auto Regressive Integrated Moving Average is a general class of statistical models for time series analysis forecasting. It’s an extension of the non-seasonal ARIMA model, designed to handle data with seasonal patterns.

SARIMA captures both short-term and long-term dependencies within the data, making it a robust tool for forecasting. It combines the concepts of autoregressive (AR), integrated (I), and moving average (MA) models with seasonal components.

Four components: **S + AR + I + MA (p, d, q).**

S (Seasonal Component): repeating patterns in the data. This could be daily, monthly, yearly, or any other regular interval.

AR (Auto Regressive): the time series is linearly regressed on its own past values.

-> p: the number of past values included in the AR model.

$$Yt = c + \varphi 1 Yt\text{-}1 + \varphi 2 Yt\text{-}2 + ... + \varphi p Yt\text{-}p + \varepsilon t$$

I (Integrated): if not stationary, the time series can be different to become stationary, compute the differences between consecutive observations.

-> d: the number of times the time series is different.

$$\nabla yt = yt - yt-1$$

MA (Moving Average): the time series is ‘regressed’ on the past forecast errors.

-> q: the number of past forecast errors included in the MA model.

$$Yt = c + \theta 1 \varepsilon t\text{-}1 + \theta 2 \varepsilon t\text{-}2 + ... + \theta q \varepsilon t\text{-}q + \varepsilon t$$

5.3.2. Forecast results

The team applied the SARIMA model to forecast the store's purchasing demand using the FactStockHolding data from January 2, 2013, to May 31, 2016. The forecasted total purchasing demand for June 2016 to August 2016 was found to be 5,467,528,933 - 5,696,486,037 - 5,653,627,293, respectively. From the chart below, we can identify some important trends.

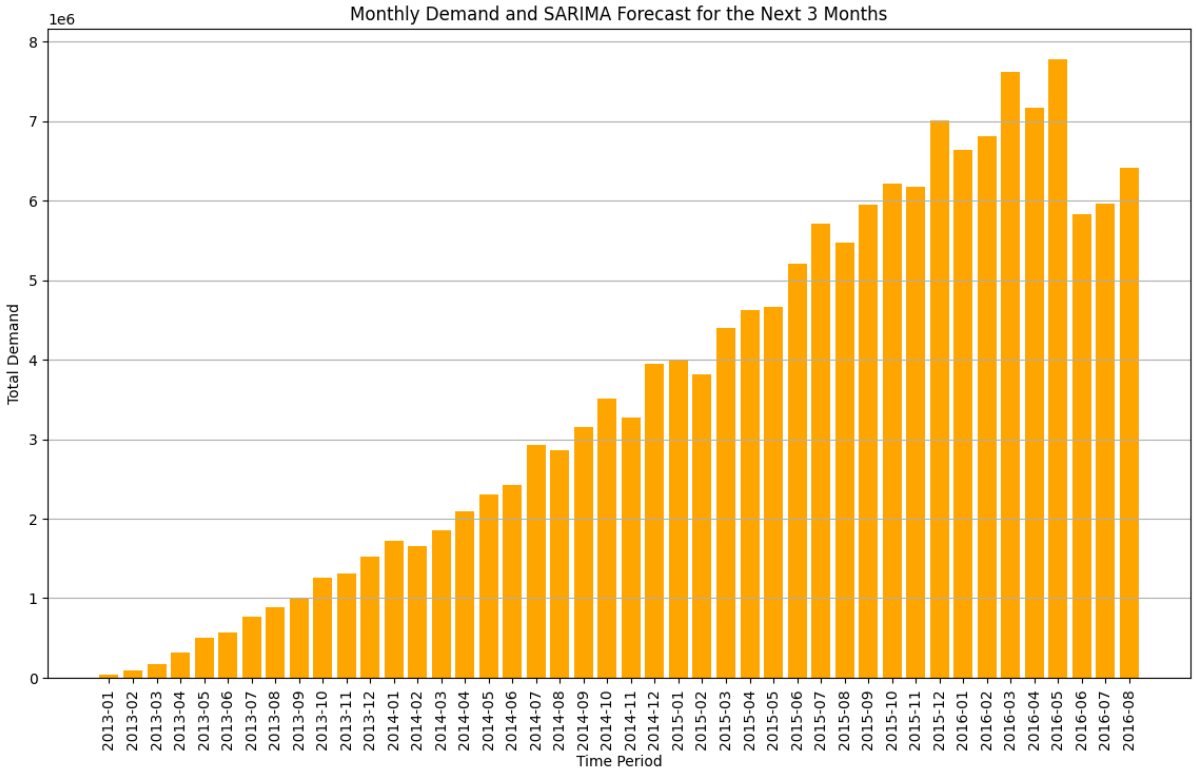

Figure 5.44. Yearly Demand and Forecast for the next 3 months

Source: The author synthesizes suggestions

The forecast for the next three months, specifically June, July, and August 2016, shows that the projected purchasing demand will be lower compared to previous years. This may indicate a temporary downward trend in market demand or could reflect seasonal factors affecting purchasing demand. This is a critical period that the business needs to monitor closely to adjust purchasing plans and inventory levels appropriately, ensuring neither shortage or surplus.

Comparing the years, we see that 2015 had the highest purchasing demand. This could be the result of several positive factors combined, such as effective marketing strategies, the launch of new products, or a sudden increase in market demand. The years 2014 and 2016 had relatively stable demand, while 2013 had the lowest demand, possibly due to the business just starting to implement or the market not being fully developed.

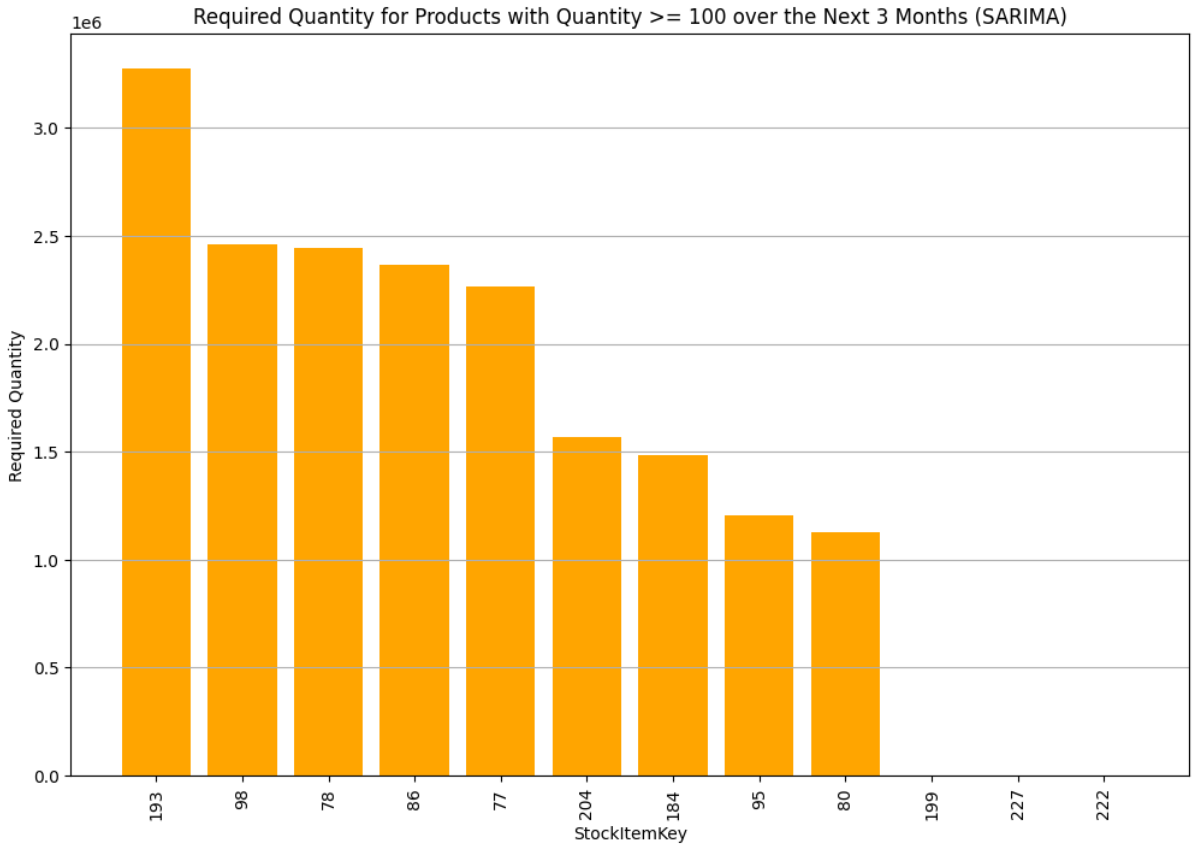

Figure 5.45. Required Quantity for Products over the next 3 months

Source: The author synthesizes suggestions

The second chart focuses on the quantity required for products with a demand of 100 or more units in the next three months. The distribution of product quantities shows that the product with StockItemKey 193 has the highest purchasing demand, significantly surpassing other products. This indicates that this product plays a very important role in the company's product portfolio and needs to be prioritized in purchasing and inventory management plans to ensure there is always enough stock to meet market demand. The next products, such as those with StockItemKey 98, 78, 86,

and 77, also have relatively high purchasing demand but are lower compared to StockItemKey 193.

Focusing on products with a demand of 100 or more units, we see that only a few products reach this threshold. This suggests that a small number of products play a crucial role in the company's total purchasing demand. These products could be key items with high consumption levels and therefore need to be managed more strictly in terms of inventory compared to other products.

5.4. Evaluation and Discussion

Using SSAS (SQL Server Analysis Services) to analyze data provides a comprehensive and detailed process. This process, from creating an SSAS project and setting up data sources to creating cubes and linking columns, helps systematically extract information from complex datasets. However, a notable limitation is that data quality issues can affect the accuracy of the analysis. If the data is inaccurate or incomplete, conclusions from the analysis may be skewed.

Power BI dashboards offer a clear view of Wide World Importer's warehouse performance and inventory management. These visualizations make it easy to track key metrics such as inventory accuracy, warehouse capacity, and processing time. The dashboard on inventory and purchasing management provides actionable insights to optimize inventory management and purchasing strategies. However, while dashboards are useful, they may lack depth in analysis and do not always provide insights into the underlying causes of trends. Incorporating predictive analysis could enhance the ability to forecast future performance.

The SARIMA (Seasonal Auto Regressive Integrated Moving Average) model has provided crucial predictions about the company's purchasing demand. Forecasts for the upcoming months indicate a declining trend compared to previous years, allowing the company to adjust its purchasing and inventory plans accordingly. The SARIMA model is sensitive to historical data and may not fully reflect external factors or sudden changes affecting demand. To improve prediction accuracy, refined forecasting methods and additional contextual data may be needed.

Overall, data analysis tools like SSAS and Power BI, combined with the SARIMA predictive model, have provided valuable insights into company performance and demand. While these tools offer actionable insights, attention must be paid to data quality, contextual analysis, and prediction accuracy. Addressing these issues through additional contextual data and advanced forecasting methods will enhance the overall effectiveness of the data analysis process.

CHAPTER 6: CONCLUSION AND FUTURE WORK

6.1. Result

In this project, we implemented Business Intelligence solutions to enhance the operational efficiency of Wide World Importers. The process involved data collection and processing, data analysis, and visualization of results to support strategic and operational decision-making.

We utilized SQL Server Analysis Services (SSAS) to analyze data from the data warehouse and developed Key Performance Indicators (KPIs) to assess operational efficiency. Additionally, Power BI was used to create interactive dashboards, providing an overview of warehouse and purchasing department performance.

Furthermore, the SARIMA forecasting model provided accurate predictions of future product demand trends, aiding in effective planning and inventory management.

6.2. Limitation

It's crucial to recognize the limitations of this study. Throughout the research, we faced challenges related to data processing. Our limited experience in efficiently and accurately handling data impacted the study's outcomes.

Another key consideration is that the source data came exclusively from the Wide World Importers database. While this is a well-known and commonly used sample database in Business Intelligence research, it might not accurately represent the business environment of a real-world enterprise. Thus, we should be mindful that the data in this study may be incomplete and may not fully capture the actual situation.

Additionally, our study had limitations in both scope and time. We were restricted to using Wide World Importers data from 2013 to 2016, which means the findings may be outdated and not fully relevant to current conditions. To enhance the timeliness and accuracy of the analysis, it would be beneficial to use real-time and continuously updated data.

6.3. Future work

This project has successfully implemented Business Intelligence (BI) solutions to enhance the operational efficiency of Wide World Importers. However, there are several areas where future work could further improve and expand upon the current findings:

- Expanding the scope and data sources from other systems such as ERP, CRM, and other systems to obtain a richer dataset. In addition, we can also search for data sources related to industries and markets to enhance the accuracy and applicability of business decisions.
- Incorporating real-time data integration to enhance the timeliness and accuracy of analysis. This will ensure that decision-making processes are based on the most current data available.
- Integrating more advanced predictive analytics techniques, such as machine learning algorithms, could improve accuracy and provide deeper insights. Exploring techniques like neural networks, gradient boosting, or ensemble models may offer more robust predictions.
- Developing comprehensive training programs and support resources to ensure users can maximize the benefits of the implemented solutions.
- Addressing data quality issues is essential for accurate analysis. Future projects should prioritize implementing robust data cleaning and validation processes to improve the overall quality and reliability of the data used for BI.
- Exploring emerging BI tools and technologies to enhance the capabilities of the current system. This could include investigating the use of cloud-based BI solutions, advanced data visualization tools, or augmented analytics platforms.
- Increasing the applicability of the study by building scenarios and experiments to help managers make more effective business decisions.

REFERENCES

[1] Wibjorn. (n.d.). *Microsoft Learn: Build skills that open doors in your career*. https://learn.microsoft.com/en-us/

[2] Hamed Taherdoost. Different Types of Data Analysis; Data Analysis Methods and Techniques in Research Projects Authors. International Journal of Academic Research in Management (IJARM), 2020, 9 (1), pp.1-9. hal-03741837

[3] Elena, C. (2011). Business intelligence. Journal of Knowledge Management, Economics and Information Technology, 1(2), 1-12.

[4] "IBM:What is business intelligence (BI)?" [Online] . Available: *What Is Business Intelligence (BI)? | IBM*

[5]"Get Ahead with BI: 7 Steps For Implementing Business Intelligence" [Online] . Available: *Get Ahead with BI: 7 Steps For Implementing Business Intelligence (datarundown.com)*

[6] Simitsis, A., Vassiliadis, P., & Sellis, T. (2005, April). Optimizing ETL processes in data warehouses. In 21st International Conference on Data Engineering (ICDE'05) (pp. 564-575). Ieee.

[7] HubSpot. (n.d.). What is ETL & why is it important for business decision making?. Retrieved July 17, 2024, from https://blog.hubspot.com/marketing/what-is-etl

[8] Analytics Vidhya. (2022). A complete guide on building an ETL pipeline for beginners. Access form https://www.analyticsvidhya.com/blog/2022/06/a-complete-guide-on-building-an-etl-pipeline-for-beginners/

[9] Oracle. (n.d.). What is a data warehouse? Retrieved July 17, 2024, from https://www.oracle.com/database/what-is-a-data-warehouse/#:~:text=a%20Data%20Warehouse%3F-,Data%20Warehouse%20Defined,large%20amounts%20of%20historical%20data.

[10] Insightsoftware. (n.d.). Data mart. Retrieved July 17, 2024, from https://insightsoftware.com/encyclopedia/data-mart/

[11] NetSuite. (n.d.). 3 types of data marts and how to get the most from that particular dataset. Retrieved July 17, 2024, from https://www.netsuite.com/portal/resource/articles/data-warehouse/data-mart.shtml#:~:text=3%20Types%20of%20Data%20Marts,most%20from%20that%20particular%20dataset.

[12] Software AG. (n.d.). Schemas for data warehouses: Star, galaxy, snowflake. Retrieved July 17, 2024, from https://www.softwareag.com/en_corporate/blog/streamsets/schemas-data-warehouses-star-galaxy-snowflake.html

[13] GeeksforGeeks. (n.d.). Fact constellation in data warehouse modelling. Retrieved July 17, 2024, from https://www.geeksforgeeks.org/fact-constellation-in-data-warehouse-modelling/

[14] Panoply. (n.d.). Data mart vs. data warehouse. Retrieved July 17, 2024, from https://panoply.io/data-warehouse-guide/data-mart-vs-data-warehouse/

[15] Coursera. (n.d.). Data mart vs. data warehouse. Retrieved July 17, 2024, from https://www.coursera.org/articles/data-mart-vs-data-warehouse

[16] Ellow. (n.d.). Advantages and disadvantages of data warehouse. Retrieved July 17, 2024, from https://ellow.io/advantages-and-disadvantages-of-data-warehouse/

[17] JavaTpoint. (n.d.). Advantages and disadvantages of data warehouse. Retrieved July 17, 2024, from https://www.javatpoint.com/advantages-and-disadvantages-of-data-warehouse

[18] Twin, A. (2024, January 30). *KPIs: What are key performance indicators? types and examples*. Investopedia. https://www.investopedia.com/terms/k/kpi.asp

[19] CONSTANTINESCU, M., & DUMITRACHE, V. I. THE HIDDEN DISADVANTAGES OF PERFORMANCE INDICATORS.

[20] Microsoft. (n.d.). *Data Analysis Expressions (DAX) overview*. Microsoft Learn. Retrieved July 30, 2024, from https://learn.microsoft.com/en-us/dax/dax-overview

[21] Oracle. (n.d.). Lesson 3: Implementing slowly changing dimensions. Oracle. Retrieved July 30, 2024, from https://www.oracle.com/webfolder/technetwork/tutorials/obe/db/10g/r2/owb/owb10gr2_gs/owb/lesson3/slowlychangingdimensions.htm

[22] MashaMSFT. (2023, February 28). *Wide World Importers - Sample database for SQL - SQL Server*. Microsoft Learn. https://learn.microsoft.com/en-us/sql/samples/wide-world-importers-what-is?view=sql-server-ver16#workflow-for-warehouse-stock-items

[23] Fernando, J. (2024, June 27). *Supply Chain Management (SCM): How It Works & Why It's Important*. Investopedia. https://www.investopedia.com/terms/s/scm.asp

[24] Del Vecchio, L. (2024, July 11). *Purchasing Process: definition, steps, and best practices*. Planergy Software. https://planergy.com/blog/purchasing-process/

[25] *Warehousing Process: Steps and role in Fulfillment Efficacy*. (2024, February 1). https://www.inboundlogistics.com/articles/warehousing-process/

[26] Bland, A. (2024, May 2). *Sales Order Processing: management, systems & best practice*. Unleashed Software. https://www.unleashedsoftware.com/blog/sales-order-processing-what-is-it-and-how-do-you-optimise-it

[27] Viblo. (n.d.). Phân tích Requirement và cách Q&A với khách hàng một cách hiệu quả. Retrieved from https://viblo.asia/p/phan-tich-requirement-va-cach-qa-voi-khach-hang-mot-cach-hieu-qua-4P8566rB5Y3

[28] Stobierski, T. (2023, February 21). What is a stakeholder? ProjectManager. Retrieved from https://www.projectmanager.com/blog/what-is-a-stakeholder

[29] International Institute of Business Analysis. (2015). A guide to the business analysis body of knowledge (BABOK Guide) (Version 3.0). International Institute of Business Analysis.

[30] Viện Logistics và Chuỗi Cung Ứng Việt Nam. (n.d.). Quy trình mua hàng cơ bản trong chuỗi cung ứng. Retrieved from https://vilas.edu.vn/quy-trinh-mua-hang-co-ban-trong-chuoi-cung-ung.html

[31] TopCV. (n.d.). Quy trình quản lý kho cần biết. Retrieved from https://www.topcv.vn/quy-trinh-quan-ly-kho-can-biet

[32] ETL Tools. (n.d.). Loading dimensions into the data warehouse. Retrieved from https://www.etltools.org/loading/dimensions.html

[33] Switchboard Software. (n.d.). Five steps of the ETL process. Retrieved from https://switchboard-software.com/post/five-steps-of-etl-process/

[34] Software Testing Help. (n.d.). ETL process in data warehouse. Retrieved from https://www.softwaretestinghelp.com/etl-process-in-data-warehouse/

[35] ETL Tools. (n.d.). *Loading facts*. ETL Tools. https://www.etltools.org/loading/facts.html

[36] CareerLink. (n.d.). Lead time là gì? Ý nghĩa, vai trò và các vấn đề liên quan. Retrieved August 5, 2024, from https://www.careerlink.vn/cam-nang-viec-lam/kien-thuc-kinh-te/lead-time-la-gi-y-nghia-vai-tro-va-cac-van-de-lien-quan

[37] Viindoo. (n.d.). Reorder point. Viindoo. Retrieved August 5, 2024, from https://viindoo.com/vi/blog/quan-tri-doanh-nghiep-3/reorder-point-1771

[38] American Material Handling Systems, Inc. (2023). Warehouse capacity utilization: Maximizing efficiency and productivity. Retrieved from https://www.amsc-usa.com/blog/warehouse-capacity/#:~:text=Warehouse%20capacity%20utilization%20is%20the,utilization%20is%201%2C800%20square%20feet

[39] pmdarima.arima.auto_arima — pmdarima 2.0.4 documentation. (n.d.). https://alkaline-ml.com/pmdarima/modules/generated/pmdarima.arima.auto_arima.html