F

VIETNAM NATIONAL UNIVERSITY HO CHI MINH CITY
UNIVERSITY OF ECONOMICS AND LAW
FACULTY OF INFORMATION SYSTEMS

REPORT

DATA ANALYTICS
BI SOLUTIONS FOR AWC

Lecturer: Mr. Nguyen Van Ho

PhD. Ho Trung Thanh

Mr. Nguyen Van Tuyen

Group 5 – K19406C

1. Phạm Quốc An   K194060841
2. Nguyễn Khánh Hưng   K194060848
3. Trần Diệu Linh   K194060851
4. Nguyễn Phương Nam   K194060854
5. Nguyễn Thị Phương Thi   K194060872

Ho Chi Minh City, May 26th 2021

ACKNOWLEDGEMENTS

COMMITMENT

This is the research work of all members of group 5.

The research results and project conclusions may not be absolute, but the team is committed to being honest and not duplicating any sources. During project implementation, group 5 consulted a number of related documents for further expansion. All documents have been cited and recorded according to regulations

Ho Chi Minh City, 2022

Group 5 – K19406C

TASK TABLE

Group 5

No.	Name	Task	%Completed
1	Phạm Quốc An	-Team leader. -Write description, report -Evaluate projects, work, results. -Comment and edit data. -Writing chapter 1,6,7	100%
2	Nguyễn Khánh Hưng	-Implement tools like power BI, excel -Analyze data, draw graphs -Writing chapter 2,3,6 -Create presentation	100%
3	Nguyễn Thị Phương Thi	-Monitor the team's progress. -Define KPIs. -Identify the main goals of the business. - -In charge of the ETL section. -Write a description of the Chapter 5 sections. -Writing chapter 4,5.	100%
4	Trần Diệu Linh	-Word synthesis -Check the job content -In charge of presentation content and layout. -Writing chapter 2	100%
5	Nguyễn Phương Nam	-Building a data	100%

		warehouse -Building KPIs system. -Data analysis with MDX and OLAP techniques. -Technical analysis. -Writing chapter 4,5.	

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

No.	Abbreviation	Explanation
1	AWC	Adventure Works Cycles
2	BI	Business Intelligence
3	DB	Database
4	EIM	Enterprise Information Management
5	ERP	Enterprise Resource System
6	ETL	Extracts - Transforms - Load
7	IT	Information Systems
8	KPI	Key Performance Indicator
9	MDX	MultiDimensional eXpressions
10	OLAP	Online Analysis Processing
11	OLE	Object Linking and Embedding
12	OLTP	On-line transactional processing
13	RFPs	Request For Proposal
14	ROI	Return On Investment
15	SQL	Structured Query Language
16	SSAS	SQL Server Analysis Services
17	SSIS	SQL Server Integration Services
18	SSDT	SQL Server Data Tool
19	SSIS	SQL Server Integration Services
20	XML	EXtensible Markup Language

CHAPTER 1. INTRODUCTION

1.1. Business case for the project

In this project, our team will work mainly with a data source from Microsoft's virtual company, Adventure Work Company, which is made for learning and research purposes. Overall, AWC is a multinational company specializing in bicycle products. In addition, they not only sell bicycles, but also provide accessories, clothing and components. As for components, Adventure Works sells brakes, chains, derailleurs, and more. Many of them are manufactured by suppliers, so AdventureWorks stands as a reseller.

There are 2 business models in Adventure Works: retail bicycle shops and internet sales for individual customers. To run the business operations, Adventure Works has a total of 290 employees. Adventure Works' customers include more than 700 stores and more than 19,000 individuals worldwide, and its suppliers are numbered around 100 supply companies.

Adventure Works Cycles is looking to expand the market by targeting sales to the best customers, expanding the product offering through an external website, and reducing costs through lower product costs.

In this project, our team will be in charge of the management of the "purchase" segment. We will handle the paperwork related to the areas of purchase, supplies and materials. Procuring materials from different suppliers, creating and tracking company orders, besides that we also take on the role of paying suppliers on time, ensuring quantities are Import is correct, the quality of imported goods must be good. Our "purchasing" stage is the stage that is considered to be the "starting" stage, so it needs precision, meticulousness and is an extremely important stage.

1.2. Objectives of the project

1.2.1. General Objective

The main goal is to provide the company with accurate indicators, reports and warnings for the company, thereby helping the company to make the most optimal strategic goals by building, deploying and operating business intelligence solutions.

1.2.2. Specific Objectives

All members understand and participate in data analysis in the purchasing segment.

The team will be familiar with and working with BI tools such as Excel, Power BI, SQL server, Visual Studio.

Our team will give a full report of the indicators, detailing what is going on inside the company and the next directions for further growth for the company.

All members participate in the report.

1.3. Research Objects

- Company: virtual company called Adventure Works Company with fictitious figures.

- Product and services providing: Bicycle and Components.

- Database Our team will redo the data warehouse only for the Purchase segment based on the available data of AWC in 2014

- Microsoft SQL Server Management: 2018.

- Power BI: 2018.

1.4. Scope of the project

The main scope of the group is the purchasing department and everything related to it. In addition, the team will use the available data of AWC company to analyze, then make optimal decisions. So the main goal will be to plan, analyze, and evaluate the information based on what we have.

1.5. Value and desired outcome of the project

In terms of value, our team thinks that the value we bring is the experience, our team has, is and will learn more things from this project. And our project will also be a document for future generations to refer to and perfect in the future.

1.6. Structure of project

Chapter 1: Introduction - Overview of the business case, objective, and research scope of the project.

Chapter 2: Theoretical basis - Introduce and describe the theory to carry out this project with the information related to BI, Data Warehouse, processes, and Business Data Analytics techniques to build a Business Intelligence support decision-making system and solution.

Chapter 3: Requirements analytics and introduction to BI solution - Analyze business requirements, technical requirements, and BI solutions needed for this project.

Chapter 4: Building data warehouse and integrating data - Detail about the building data warehouse process for the purchasing module as well as integrate data from the Adventureworks2014 data sample into our data warehouse.

Chapter 5: Data analytics - Use the SSAS technology and the MDX and OLAP technique to make a presentation for the information from the Data Warehouse and visualize them to help the manager make a better decision.

Chapter 6: Visualization and Forecasting - Visualize the requirements and significations of the report according to data in the data warehouse we built and using some BI tools to conclude statistical reports. Chapter 7: Conclusion and future works - Present the results of the topic, the limitations of the topic, the practical meaning, and the next development direction of the topic.

CHAPTER 2. THEORETICAL BASIS

2.1. Overview about BI

BI is a set of processes, tools, technologies and methodologies that combine data from various data sources and make the single source of data available at the right time so that timely, better informed and strategic decisions can be made.

2.1.1. What is BI?

Business intelligence (BI) leverages software and services to transform data into actionable insights that inform an organization's strategic and tactical business decisions. BI tools access and analyze data sets and present analytical findings in reports, summaries, dashboards, graphs, charts and maps to provide users with detailed intelligence about the state of the business.

The term business intelligence often also refers to a range of tools that provide quick, easy-to-digest access to insights about an organization's current state, based on available data.

2.1.2. BI Architecture

Architectural BI can be deployed in the on-premises data center or the cloud. In both cases, it contains a set of software cores that collectively support the 22 different phases of the BI process, from data acquisition, integration, storage, and analytics to data visualization, Information delivery and use of BI data in business decision-making. The cores include the following:

- Source systems
- ETL process
- Data modeling
- Data warehouse
- Enterprise information management (EIM)
- Appliance systems
- Tools and technologies

2.1.3. Advantage of BI in enterprises

2.1.3.1. Visual and Analytics

These are the 2 central features of the BI reports that iERP deploys, administrators can track their data easily, replacing numbers on dry reports into charts to help Administrators

quickly grasp the direction of the activities of their business, especially in terms of revenue, expenses…With just drag and drop, you can build intuitive reports, analyze and aggregate them into dashboards.

Our intelligent reporting software connects with many different data simultaneously, creating a central data analysis center across departments (sales, finance and accounting, human resources, warehouse), or corporate affiliates, and easily share these analytics with colleagues.

2.1.3.2. Discovery

You can easily build scenarios to answer every question asked by an administrator, such as if shipping costs increased by 2%, how would it affect profits and then the index? How is profit on revenue affected; As well as a 3% increase in the exchange rate, how will it affect the raw materials as well as the profit of the business.

BI software system helps administrators easily operate business data along with the data of the industry they are operating in, helping managers quickly understand the market, understand customers and own product

2.1.3.3. Map

Using the analysis of data displayed on the geographical map makes it very easy for you to recognize the analytical indicators on the map such as revenue in the area, the increase/decrease in revenue of the area, coverage (coverage) of products in geographical areas, etc. Very quickly the system, BI software immediately provides administrators with all important information.

2.1.3.4. Mobility

The BI system is compatible with Android and iOS to help administrators always have information and the ability to analyze data anytime, anywhere.

2.1.3.5. Facebook, Twitter and social networks

Directly analyze the metrics of Facebook fan pages, Twitter twitters along with Twitter details, and with direct access to Google Analytics to help administrators navigate the market, more deeply with product features to better serve the needs of customers in each market

2.1.3.6. Integrate with CHEAP language

The ability to analyze data on statistical data systems, and you can develop the R language and with the drag and drop of visuals and data analysis helps you quickly get a lot of perspectives different from the statistics.

The BI system connects administrators with data, enabling administrators to put their management ideas into analysis and data in order to find opportunities for their business

activities, improve productivity and improve business performance. competitiveness and optimize the assets of your business.

2.1.4. BI Strategy for Business

Your BI strategy needs to first align with your business goals and vision. With business intelligence, data can drive transformation in your organization.

Step 1: Choose a sponsor

Step 2: Choose your BI platform

Step 3: Identify the key stakeholders and get them involved

Step 4: Assemble your BI team

Step 5: Define the scope of BI

Step 6: Prepare your data infrastructure

Step 7: Develop a business intelligence roadmap

2.2. ETL Process

2.2.1. What is ETL?

ETL, which stands for extract, transform and load, is a data integration process that combines data from multiple data sources into a single, consistent data store that is loaded into a data warehouse or other target system.

2.2.2. Why do we need ETL?

ETL tools break down data silos and make it easy for your data scientists to access and analyze data, and turn it into business intelligence. In short, ETL tools are the first essential step in the data warehousing process that eventually lets you make more informed decisions in less time.

2.2.3. ETL Process

The 5 steps of the ETL process are: extract, clean, transform, load, and analyze. Of the 5, extract, transform, and load are the most important process steps.

- Extract: Retrieves raw data from an unstructured data pool and migrates it into a temporary, staging data repository

- Clean: Cleans data extracted from an unstructured data pool, ensuring the quality of the data prior to transformation.

- Transform: Structures and converts the data to match the correct target source

- Load: Loads the structured data into a data warehouse so it can be properly analyzed and used

- Analyze: Big data analysis is processed within the warehouse, enabling the business to gain insight from the correctly configured data.

2.3. Data warehouse and Data mart

2.3.1. What are Data warehouse and Data mart?

A data mart is a subset of a data warehouse oriented to a specific business line. Data marts contain repositories of summarized data collected for analysis on a specific section or unit within an organization, for example, the sales department.

A data warehouse is a large centralized repository of data that contains information from many sources within an organization. The collated data is used to guide business decisions through analysis, reporting, and data mining tools.
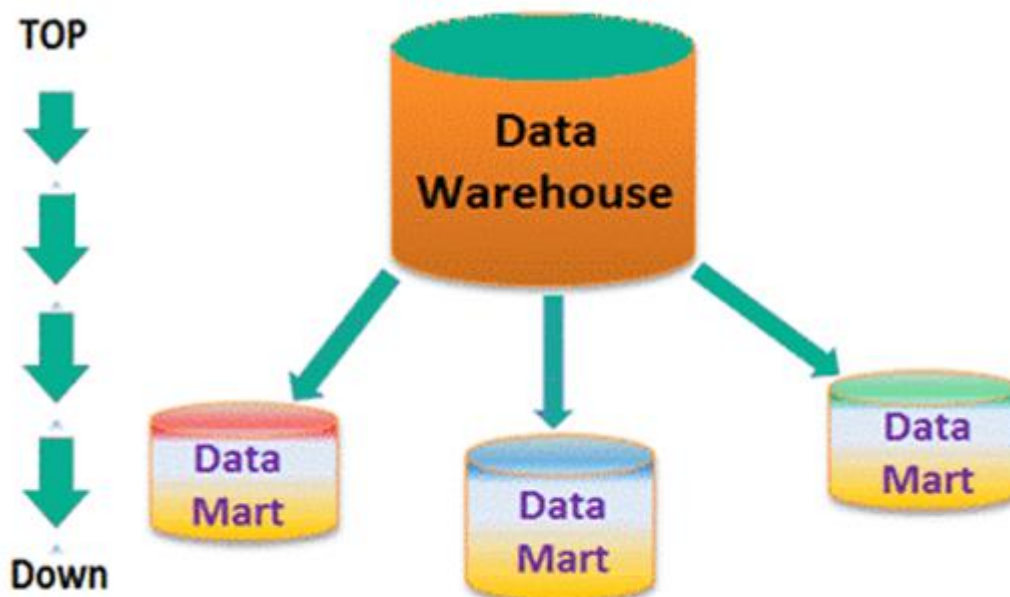

Figure 2-1. Data warehouse and Data mart

2.3.2. Who needs Data warehouse and Data mart?

DWH (Data warehouse) is needed for all types of users like:

- Decision makers rely on large amounts of data

- Users use complex processes to obtain information from multiple data sources.

- Used by people who want simple technology to access data

- For those who want a systematic approach to decision making, it is essential

- If users want fast performance on large amounts of data needed for reports, grids or charts, then Data Warehouse is very useful.

- If you want to discover 'hidden patterns' of data streams and groups then data warehousing is the first step

2.3.3. Advantages and disadvantages of Data warehouse

2.3.3.1. Advantages of Data warehouse

- Speedy Data Retrieving

How many times have you had a piece of information that you wanted to retrieve, but forgot where it was placed? Once you input your information into your DW, you will never lose track of this data again. By conducting a quick search, you'll be able to find the statistic and further analyze it — without having to waste time in your search.

- Error Identification & Correction

Many of us have the natural tendency to skip over dotting the I's and crossing the T's when inputting raw information, but data warehouses help eliminate user oversight. Before loading data, your data warehouse makes it a point to show you inconsistencies and correct them. This is extremely helpful for those who may be careless or hasty when it comes to gathering data.

- Easy Integration

Your DW also adds value to operational business applications like CRM systems when the warehouse is successfully integrated. The complexity of its makeup allows a data warehouse to translate the information in a more simple, digestible format to the point where your team members can easily understand what's been placed in front of them.

2.3.3.2. Disadvantages of Data warehouse

- Time Consuming Preparation

While a major part of a data warehouse's responsibility is to simplify your business data, most of the work that will have to be done on your part is inputting the raw data. Now, while the job the DW does for you is helpful and extremely convenient, this is the most work you'll have to manually perform, as the DW performs many other functions for you.

- Difficulty in Compatibility

Depending on the system you currently have in place, the use of data warehouse technology could likely require a helping hand from an independent BI team. With the intricacies of operating systems, software and programs, it can be difficult for a business

owner to figure out how to properly make use of their data warehouse. Especially since the costs of these tools are investments in your business, you'll want to ensure that your system is working exactly the way you intend it to.

- Maintenance Costs

One of the pros and cons of your DW is its ability to consistently update. This is great for the business owner who wants the best and latest features however these upgrades don't usually come cheap.

Including regular maintenance for your system, you can expect to shell out more than your initial investment should you want to have the latest technology at your fingertips.

- Limited Use Due to Confidential Information

If you have sensitive data that should only be viewable from a certain staff member, your DW's use will be limited. In order to maintain the security of your current system, less usage could eventually decrease the overall value of your data warehouse.

No matter your needs or concerns, our specialists at Business Impact look forward to helping you make the right decision when it comes to selecting the right BI solution for your company. Contact us today to learn more about how we can help your organization get the most out of its business intelligence solution.

2.3.4. Snowflake and Star schemas

2.3.4.1. What is a Star Schema?

Star Schema in a data warehouse, in which the center of the star can have one fact table and a number of associated dimension tables. It is known as star schema as its structure resembles a star. The Star Schema data model is the simplest type of Data Warehouse schema. It is also known as Star Join Schema and is optimized for querying large data sets.
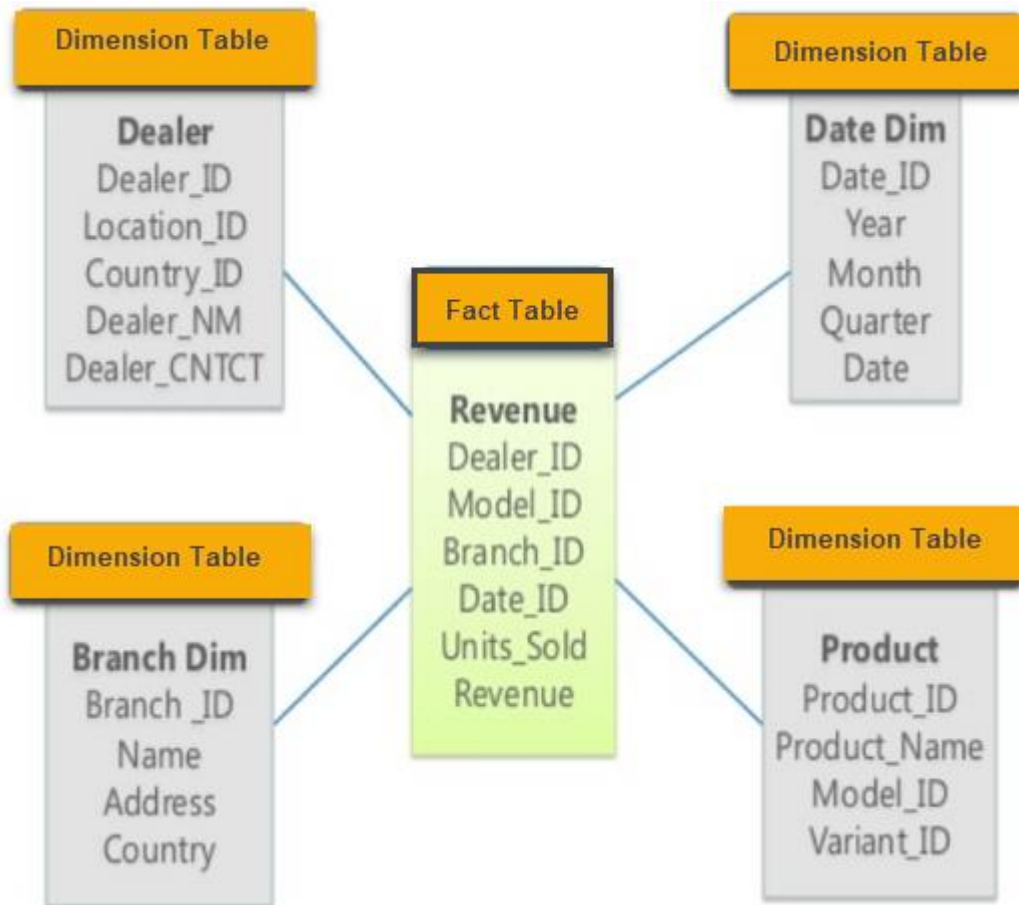
Figure 2-2. Star Schema

2.3.4.2. What is a Snowflake Schema?

Snowflake Schema in a data warehouse is a logical arrangement of tables in a multidimensional database such that the ER diagram resembles a snowflake shape. A Snowflake Schema is an extension of a Star Schema, and it adds additional dimensions. The dimension tables are normalized which splits data into additional tables.
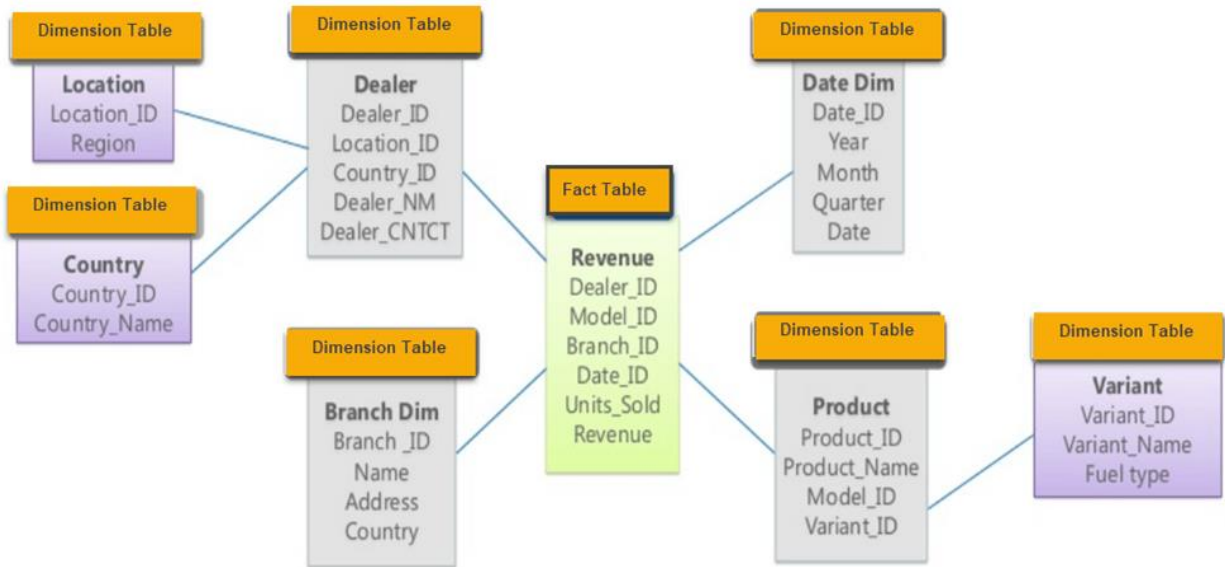
Figure 2-3. Snowflake Schema

2.4. KPIs

2.4.1. KPIs Definition

Key performance indicators (KPIs) refer to a set of quantifiable measurements used to gauge a company's overall long-term performance.

KPIs specifically help determine a company's strategic, financial, and operational achievements, especially compared to those of other businesses within the same sector.

2.4.2. The advantages and disadvantages of KPIs

2.4.2.1. The advantages of KPIs

- Measurable Results

As the sole goal of KPI would be to monitor progress, it shows real benefits in the shape of amounts, metrics, or data.

The worker, staff, or company could quickly quantify or track the development of their goal and know which part of this job requires more attention.

- Alignment

For a large organization with a high number of workers, it might get tough to keep an eye on everyone's progress.

In this circumstance, KPI helps everybody remain aligned to the target since it makes the outcomes accessible to everybody involved with the job.

This helps everybody stay inspired as nobody would love to see their titles or advancement marked red. Furthermore, it ensures everybody works in precisely the same direction.

- Future Strategies

Tracking the progress via KPIs can enable the supervisors to redesign or alter their plans depending on their prior target functionality.

As KPIs help the organization understand everybody's capacity, performance indicator, and productivity, it creates strategy or establishes future objectives.

- Rewards

Any worker intends to work better and hard towards getting an increase or bonuses for their hard work.

With KPIs, every individual gets an opportunity to prove themselves and aid the supervisors to see the advancement and benefit accordingly.

Along with this, it helps workers to monitor their performance and enhance themselves.

2.4.2.2. The disadvantages of KPIs

- Decrease in Quality

With the prime focus on getting results for short-term objectives, there's an excellent likelihood of workers losing attention to the standard of the job.

On account of the setting of financial objectives, there's a trend for metrics gaining additional weight instead of the credibility of the endeavor.

- Short-term Oriented

KPIs help attain short-term targets but might be disadvantageous in attaining long-term objectives. Many resources quantify success in one area that can't be acceptable.

- Standardization

Since the goals are somewhat more result-oriented, there might be a prospect of reducing the degree of creativity of their workers.

Because of this, it discourages workers from executing or considering innovative strategies.

- Loyalty

Since KPIs only demonstrate the advancement levels, it becomes hard to monitor the essence of the job. Consequently, it could impact the loyalty between the company and the customer.

Because of this, the company might lose them and weaken the bond between these.

In conclusion, KPIs are beneficial for short-term aims, whereas; they might not have to be valuable for long-term objectives. It's critical to look at all of the factors before contemplating a KPI for your company.

2.4.3. Categories of KPIs (Purchasing, Production, Sales, or HR)

- Purchase order cycle time

- Cost of purchase order

- Purchases in time and budget

- Purchase order accuracy

- Emergency purchase ratio

- Cost reduction

- Cost avoidance

- Procurement ROI

2.5. MDX language for analyzing multidimensional data and OLAP

2.5.1. What is MDX language?

MDX is a query language designed for OLAP databases, as SQL is a query language for relational databases. MDX is essentially an extension to SQL used for queries and script access to multidimensional data. MDX queries access data stored in a SQL Server Analysis Server cube by bringing back facts related to dimensions.

2.5.2. OLAP technique

OLAP (online analytical processing) is a computing method that enables users to easily and selectively extract and query data in order to analyze it from different points of view. OLAP business intelligence queries often aid in trends analysis, financial reporting, sales forecasting, budgeting and other planning purposes.

2.5.3. MDX method and structure

SELECT clause is used to specify the dimension of the resulting file.

FROM clause specifies the source data (cube) used to get the data.

WHERE clause is used to define the husk data, for the purpose of filtering out the top of the data.

WITH clause: allows the calculation of the named sets during the processing of the SELECT and WHERE clauses.

2.6. Azure Analysis Service (Optional)

Azure Analysis Services is a new preview service in Microsoft Azure where you can host semantic data models. Users in your organization can then connect to your data models using tools like Excel, Power BI and many others to create reports and perform ad-hoc data analysis.

CHAPTER 3. REQUIREMENTS ANALYTICS AND INTRODUCTION TO BI SOLUTION

3.1. Business processes (Purchasing)

3.1.1. Purchasing department

A company's purchasing department may have a large influence on its ability to reach its strategic and daily operational goals. Its ability to acquire enough materials while lowering costs can allow a company to raise its profits, lower its expenditures and achieve growth. Purchasing departments can use their contacts with suppliers to improve the overall quality of a business' product and lower risks in a company's operations by effectively managing these relationships. On the operational level, purchasing departments ensure businesses receive everything they need for projects. This can include raw materials for manufacturing, supplies for employees, office spaces or technology. When purchasing departments provide these goods efficiently, it can ensure production remains high and products and services reach customers on time.

At Adventure Works Cycles, the purchasing department plans to purchase supplies and parts that are used to manufacture bicycle parts based on the organization's business goals…The purchasing department often maintains relationships with important bicycle vendors. They can provide benefits for longtime suppliers to improve their relationship and avoid competition for bicycle resources. In some cases, procurement departments may build relationships with secondary suppliers to mitigate risk in case they have to change their primary supplier. In addition, the purchasing department can monitor their operations and make the necessary changes. They can evaluate the quality of bikes their company receives and return or replace any items that are not up to their standards.

3.1.2. The purpose of Purchasing

3.1.2.1. Needs and Supplier Analysis

The starting point for a purchasing strategy is to evaluate the performance of the existing business, the purchasing department will review the company's growth trajectory and come up with a plan to make the business better and/or What are the cost savings and purchase costs per department, team, or job function? In addition, the purchasing department will analyze the supplier's market to see if the company is using the right supplier, at the right price, to meet its business needs.

3.1.2.2. Award Supplier Contracts

When the company works with a supplier that the supplier does not provide enough requirements on quality costs, production capacity, delivery schedule, it will cause great loss to the company. Therefore, the purchasing department will play a role in reviewing

before deciding to sign a contract with a supplier. In large companies. The department can also make decisions about whether to manufacture products in-house.

3.1.2.3. Supplier Selection and Relationships

For large companies, their list of suppliers is extensive and the role of the purchasing department's equipment is to manage and maintain these relationships. Work closely with suppliers which means you can share your knowledge about school changes, new products and technologies, or other elements that can help you stay ahead of the competition.

3.1.2.4. Ordering and Inventory Control

If the inventory quantity is too high, in addition to the goods stored for a long time, company will be damaged and lose quality, making it difficult to compete with competitors in the market, the purchasing department or goods are constantly bombed purchasing department will have systems in place that trigger orders whenever a certain amount of inventory is available obtain. For those using a stock management system, the minimum inventory and order quantities are usually predefined and automatically ordered by the software.

3.1.2.5. Compliance and Quality Control

Quality control is an essential part of the procurement process. The purchase department needs to continually inspect the quality, performance and reliability of the supplier to ensure they do not lapse into complacency. One essential role of the purchasing department is to analyze and measure performance data to ensure that suppliers are achieving the desired outcomes, in accordance with the company's procurement strategy. For example, the department might measure:

- The percentage of products delivered on time.

- The number of suppliers used and how much product they supply.

- Supplier availability.

- Lead times.

- Product defect rates.

These metrics enable the purchase department to assess how well suppliers are fulfilling the company's requirements, how well they respond to urgent demand, and whether the company is over-relying on just one or two key suppliers which could leave the company vulnerable if the supplier goes bust.

3.1.3. Purchasing process

Figure 3-1. Purchasing process steps

Step 1. Needs Analysis

At this stage, the company recognizes and documents a need for goods or services to solve a particular problem. The purchase department describes the need to be met, and works with others to determine how best to do so.

Step 2. Purchase Requisition to Purchase Order

Purchase requisitions are generated by the individual or group or department requesting goods or services to the purchasing department or purchasing manager. The purchase request contains full details about the items or services to be received. Purchase requisitions below the set budget threshold are automatically updated to purchase orders and sent to the preferred supplier for that item or service.

Step 3. Purchase Order Review and Approval

Approved purchase orders are sent to accounting to verify the funds exist in the appropriate budget to cover the requested goods and services.

Step 4. Requests for Proposal

POs that receive budget approval are returned to the procurement department and, as required, used to create requests for proposal (RFPs), also known as requests for quotation, or RFQs.

Step 5. Contract Negotiation and Approval

The vendor is awarded a contract, which will be further finalized before signing to ensure optimal terms and conditions and to ensure mutual agreement for both parties. Once the contract is signed, the purchase order is a legally binding agreement between the buyer and the seller.

Step 6. Shipping and Receiving

The supplier delivers the goods or services within the agreed time frame. When the goods are delivered to the buyer, they carefully review the goods and services to ensure that they have received what was promised and notify the supplier of any problems.

Step 7. Three-Way Matching

Three-way-matching is the comparison of shipping documents/packing slips with the original purchase order and the invoice issued by the supplier. This comparison is used to ensure all the information related to the transaction is accurate.

Step 8. Invoice Approval and Payment

Successfully matched orders are approved for payment. Any modifications or additional charges may require another layer of approvals before payment can be issued. Once approved, payment is issued to the vendor.

Step 9. Accounting Records Update

Completed orders are recorded in the company's books, and all documents related to the transaction are securely stored in a centralized location.

3.2. Data source and challenges

3.2.1. Data source

The data in the AdventureWorks2019 version of the database pertains to the years 2010-2019. It includes two Microsoft SQL Server databases. The first is an online transaction processing (OLTP) database, which is rich in structure, content, and variety. The second is a data warehouse, which is useful for online analytical processing (OLAP) and data mining, as well as teaching data warehouse concepts and structures.

The OLTP database consists of 69 tables grouped into five schemas related to Adventure Works' business model: Sales, Purchasing, Production, Human Resources, and Person. The database (in its raw state) contains data of almost 20,000 people (employees, customers, store contacts, vendor contacts, and general contacts). It also contains data of over 31,000 sales transactions to customers and over 4000 purchasing transactions from suppliers. Several advanced data types are demonstrated in the AdventureWorks OLTP database, including bitmapped product photographs, XML documents, and hierarchy id fields for representing hierarchical data relationships
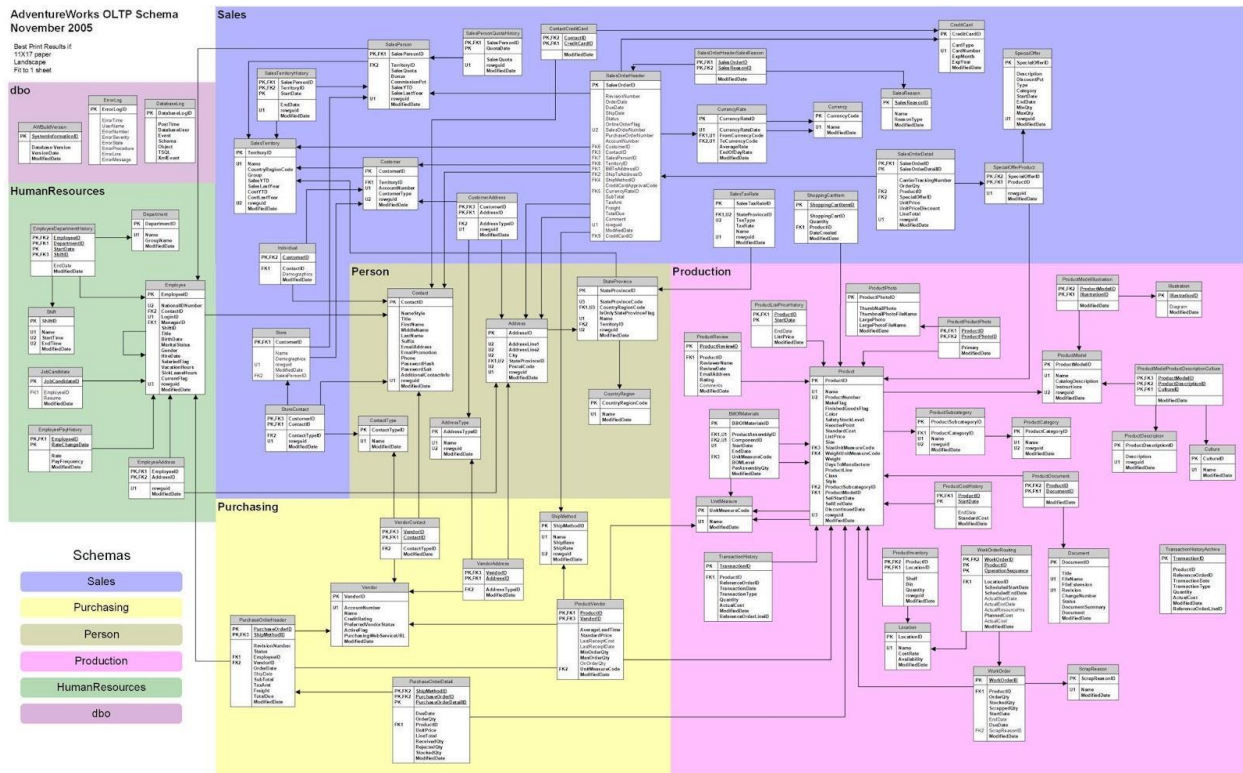
Figure 3-2. AdventureWorks OLTP Schema

3.2.2. Challenges

- Information is easily lost, poor security: documents and records arising in the processing process are not closely linked, decentralized use and monitoring easily create vulnerabilities and are at risk of being stolen. take advantage of the company.

- Unclear information: product details are not stored seamlessly and there is a lack of information about the quality of the product. From that, it can be seen that the information processing process is time-consuming and cumbersome, but it is not effective

- Information is not continuously circulated between departments

3.3. Business requirements analysis of Purchasing

3.3.1. Find the right source (compare products)

Product prices: from the data sources stored in the data warehouse, our team begins to compare and thereby make import decisions in accordance with the company's budget, optimizing the work. Enter the quantity of goods through stability, price changes…

Product quality: based on the company's data for products, from which to compare and find products of good quality, filter those products and prioritize importing products such as: So through features, durability, fit, perceived quality…

33

3.3.2. Find the right supplier.

Supplier quality: During the import process, we do not only import from one source, because it will not be enough for our company to supply to our customers, so we have to import from other suppliers together. Of course, the quality of suppliers will vary, so we will compare the supplies from the stored data, which will help the company make a decision to choose the suppliers. The most reasonable offer through the credit level of the supplier. We also need the supplier's information to be clear through the information about the supplier's address, contact method to find out. Figure out how to maintain contact with suppliers.

For example:

- We need supplier information (address, email. phone…)

- Preference for good suppliers (credit rating, preferred vendor status…)

3.3.3. Solve cost problems in the process of importing goods from suppliers

There is a comparison between the estimated cost and the actual cost. Once an order is placed, there may be costs associated with confirming the delivery date, modifying the order, or other communications regarding the status of the order. Therefore, the cost of the order may change, so we will need the data of the orders, from which to compare and derive the difference, the costs incurred when placing the order. to make it easier to control order costs. Once the supplier completes the order and sends it to your receiving department, the order needs to be checked to confirm it's of the right quality and quantity. If there is a problem, the PO may incur additional costs related to returns, additional shipping, order errors, and more.

3.3.4. Optimizing time

Order processing time will be faster, thanks to the collected data, thereby making decisions to shorten the time of shipping, exchange, and handover of products to other parties. This helps to increase the company's profits, saving time. Sometimes there will be orders that are delayed or delayed due to unexpected cases and incidents, from data on specific dates, hours, and times, thereby making more accurate processing decisions.

3.3.5. Handling inventory and damaged products.

In the process of stocking goods to provide to the sales department, there will be damaged goods or goods that the customer has canceled the order from, from the data we have, we will process to exchange or return the goods or reduce the order quantity. row. Regarding inventory products, the data warehouse will make statistics and make a decision to import less similar items to avoid overstocking or may not import more if necessary. As for damaged items, the data will be statistical and notified to return the goods to the supplier.

3.4. IT requirements Analysis (IT & Infrastructure)

A software requirement is a capability needed by the user to solve a problem or to achieve an objective. In other words, requirement is a software capability that must be met or possessed by a system or system component to satisfy a contract, standard, specification, or other formally imposed documentation. Ultimately, what we want to achieve is to develop quality software that meets customers' real needs on time and within budget.

3.4.1. Analysis software and analysis support software

Requirement analysis helps organizations to determine the actual needs of stakeholders. At the same time, it enables the development team to communicate with stakeholders in a language they understand (like charts, models, flow-charts,) instead of pages of text, using diagramming software like Diagrams. Use SQL server and BI tool, Tableau for data storage and analysis.

3.4.2. Meet data security needs

Data of a business is very important, if leaked to outside will seriously affect the company. Therefore, when using software and tools to analyze and store data, it is necessary to have good information security and safety.

CHAPTER 4. BUILDING DATA WAREHOUSE AND INTEGRATING DATA

In this chapter, we will present the process of building, designing and organizing the data warehouse of the purchasing module as well as extracting data from the data system of Adventureworks2014.

4.1. Designing Data Warehouse

4.1.1. Bus Matrix

With Adventureworks2014 data schema, we design a bus matrix for purchasing modules for analysis including 10 steps comparing with 6 tables Dimensions to see what they have in common.

Dimension Process	Time	Product	Vendor	Ship Method	Location	Employee
Prepare specification, check budget	X	X			X	
Prepare purchase requisition	X	X			X	X
Check and authorize requisition	X	X	X			X
Obtain quotations	X	X	X			
Select supplier	X	X	X			
Issue purchase order	X	X	X			X

Receive goods	X	X	X	X		X
Receive & check invoice	X	X	X			X
Pay supplier invoice	X	X	X			
Enter payment in cash book	X	X				

Table 4.1. Bus Matrix for purchasing module

4.1.2. Master Data

Master data is the set of identifiers that provides context about business data such as location, customer, product, asset, etc. It is the core data that is absolutely essential for running operations within a business enterprise or unit. Following is the master data for the main purchasing module that we need for analysis.

No.	Objects	Description
1	Product	Products information
2	Vendor	Information about suppliers of products and materials for the company
3	Employee	Employee information to perform
4	ShipMethod	Product shipping method information
5	Location	Inventory location information to control
6	Address	Street address information for vendors

4.1.3. Transaction Data

Transactional data is information that is captured from transactions. The transaction data used to build Fact tables.

No.	Objects	Description
1	PurchaseOrderHeader	General purchase order information
2	PurchaseOrderDetail	A detailed description of each item line in the purchase order
3	ProductVendor	Information on import and export quantities, prices, and product costs of suppliers
4	ProductInventory	Information AWC's inventory to manage incoming shipments.

Table 4.3. The transaction data used to build Fact tables

4.1.4. Mapping

From the business requirements mentioned in chapter 3 combined with Adventureworks2014 data, from there, my team analyzed and linked the source data to the target data, mapping atomic data units from the two data units. Various data are used in the purchasing module data warehouse.

- DimProduct: Information about products

Data Warehouse		Data Source		Data Type	Allow Nulls	Rule
Table	Column Name	Table	Column Name			
Dim Product	ProductKey (PK)			int		Auto Increment
	ProductID (NK)	Purchase.Product	ProductID	int		From Source

	Name		Product Name	nvarchar (50)		From Source
	Product Number		Product Number	int		From Source
	Safety Stock Level		Safety Stock Level	smallint		From Source
	Standard Cost		Standard Cost	money		From Source
	List Price		List Price	money		From Source
	Color		Color	Nvarchar (15)		From Source
	Size		Size	Nvarchar (5)		From Source
	Style		Style	Nchar (2)		From Source
	Weight		Weight	decimal (8,2)		From Source

- DimShipMethod: Information about Ship method.

Data Warehouse		Data Source		Data Type	Allow Nulls	Rule
Table	**Column Name**	**Table**	**Column Name**			
DimShipMethod	Ship Method Key (PK)			int		Auto Increment
	ShipMethodID (NK)	Purchase. Ship Method	Ship Method ID	int		From Source
	ShipMethodName		Name	nvarchar (50)		From Source
	ShipBase		ShipBase	Money		From Source
	ShipRate		ShipRate	Money		From Source

Table 4.5. The table describes the data mapping DimShipMethod

- DimEmployee: Information about Employee

Data Warehouse		Data Source				

Table	Column Name	Table	Column Name	Data Type	Allow Nulls	Rule
Dim Employee	Employee Key (PK)			int		Auto Increment
	Employee ID (NK)	Purchase. Employee	Employee ID	int		From Source
	National IDNumber		NationalID Number	nvarchar (15)		From Source
	Gender		Gender	Nchar (1)		From Source
	JobTitle		JobTitle	Nvarchar (50)		From Source
	BirthDate		BirthDate	Date		From Source
	SickLeave Hours		SickLeave Hours	Smallint		From Source
	Vacation Hours		Vacation Hours	Smallint		From Source

Data Warehouse		Data Source		Data Type	Allow Nulls	Rule
	AddressID	Person. Address	AddressID	Int		From Source

Table 4.6. The table describes the data mapping DimEmployee

- DimVendor: Information about Vendor

Data Warehouse		Data Source		Data Type	Allow Nulls	Rule
Table	**Column Name**	**Table**	**Column Name**			
Dim Vendor	Vendor Key (PK)			Int		Auto Increment
	VendorID (NK)	Purchase. Vendor	Business EntityID	int		From Source
	Vendor Name		Name	Nvarchar (50)		From Source
	AccountNumber		Account Number	Nvarchar (15)		From Source
	Credit Rating		Credit Rating	tinyint		From Source
	Preferred Vendor Status		Preferred Vendor Status	Bit		From Source

City		City	Nvarchar (30)		From Source
Address Line1	Person. Address	Address Line1	Nvarchar (60)		From Source
Address Line2		Address Line2	Nvarchar (60)		From Source
Postal Code		PostalCode	Nvarchar (15)		From Source

Table 4.7. The table describes the data mapping DimVendor

- DimTime: Information about Time

Data Warehouse		Data Source		Data Type	Allow Nulls	Rule
Table	**Column Name**	**Table**	**Column Name**			
DimTime	TimeKey			Int		Auto Increment
	Date			Datetime		
	Year			Int		

	Month			Int	
	Day			Int	
	Datename			Nvarchar (30)	
	DayNumber OfWeek			Int	
	DayNumberOfMonth			Int	
	DayNumberOfYear			Int	
	WeekNumberOfYear			Int	
	Quarter			Int	

Table 4.8. The table describes the data mapping DimTime

- DimLocation: Information about Location

Data Warehouse	Data Source			

Table	Column Name	Table	Column Name	Data Type	Allow Nulls	Rule
Dim Location	Location Key (PK)			int		Auto Increment
	Location Name (NK)	Production.Location	Location Name	nvarchar (50)		From Source
	CostRate		CostRate	small money		From Source
	Availability		Availability	Decimal (8,2)		From Source

Table 4.9. The table describes the data mapping DimLocation

- FactPurchaseKey: Detail information about Purchase Order

Data Warehouse		Data Source		Data Type	Allow Nulls	Rule
Table	Column Name	Table	Column Name			
Fact Purchase Key	Fact Purchase Key (PK)			int		Auto Increment
	Status	Purchase. Purchase OrderHeader	Status	Nvarchar (15)		From Source

	ShipDate		ShipDate	Int		From Source
	Order Date		OrderDate	Int		From Source
	Freight		Freight	Money		From Source
	TaxAmt		TaxAmt	money		From Source
	EmployeeKey (NK)			int		Look up fromDim Employee
	Vendor Key			Int		Look up fromDim Vendor
	Ship Method Key			Int		Look up from DimShip Method

	Product Key			Int		Look up fromDim Product
	LineTotal	Purchase. Purchase OrderDetail	LineTotal	Money		From Source
	ReceivedQty		Received Qty	Int		From Source
	RejectedQty		Rejected Qty	Int		From Source
	Stocked Qty		Stocked Qty	Int		From Source
	OrderQty		OrderQty	Int		From Source

Table 4.10. The table describes the data mapping FactPurchaseKey

- FactProductVendorKey: Detail information about Product Vendor

Data Warehouse		Data Source		Data Type	Allow Nulls	Rule
Table	Column Name	Table	Column Name			
FactProduct	Fact Product			int		Auto Increment

VendorKey	Vendor Key (PK)					
	Product Key			Int		Look up from Dim Product
	Vendor Key			Int		Look up fromDim Vendor
	Average LeadTime	Purchase .Product Vendor	Average LeadTime	Int		From Source
	Standard Price		Standard Price	Money		From Source
	Last Receipt Cost		Last Receipt Cost	Money		From Source
	Last Receipt Date		Last Receipt Date	int		From Source
	OnOrderQty		OnOrderQty	Int		From Source

| | MaxOrderQty | | MaxOrderQty | Int | | From Source |
| | MinOrderQty | | MinOrderQty | Int | | From Source |

Table 4.11. The table describes the data mapping FactProductVendorKey

- FactProductInventoryKey: Detail information about Product Inventory

Data Warehouse		Data Source		Data Type	Allow Nulls	Rule
Table	**Column Name**	**Table**	**Column Name**			
FactProductInventory Key	Product Inventory Key (PK)			int		Auto Increment
	Product Key			Int		Look up from Dim Product
	Location Key			Int		Look up from Dim Location
	Shelf	Product. Product	Shelf	Nchar (10)		From Source

49

		Bin	TinyInt		From Source
Bin	Inventory				
Quantity		Quantity	Int		From Source

Table 4.12. The table describes the data mapping FactProductInventoryKey

4.1.5. Data Warehouse model

After converting the data of Adventureworks2014 to the data warehouse to be analyzed, my team described the relationship of the data tables through the star model. This model simplifies the data and presents it clearly, easily, fully representing the Dim, Fact tables.

In it, the FactProductVendor table will show identifiable data, compare the quality between suppliers, and the FactPurchase table will show data that can determine the cost, time, and price in the process. purchase process, and finally, the FactProductInventory table will show data that can determine the status of the warehouse to supply and ship products to. The Dimensions Table provides manager-specific information during the purchase process such as DimVendor, DimProduct, DimEmployee, DimTime, and DimShipMethod.

Objective	Description
DimProduct	Information about Product
DimShipMethod	Information about Ship Method
DimEmployee	Information about Employee
DimVendor	Information about the Vendor
DimTime	Information about the Time
DimLocation	Information about Location

FactPurchaseKey	Information about Purchase Order
FactProductVendorKey	Information about the Product Vendor
FactProductInventoryKey	Information about the Product Inventory

Table 4.13. Describes the tables in the Data Warehouse model

All the Dimension and Fact tables will be described in the diagram below:
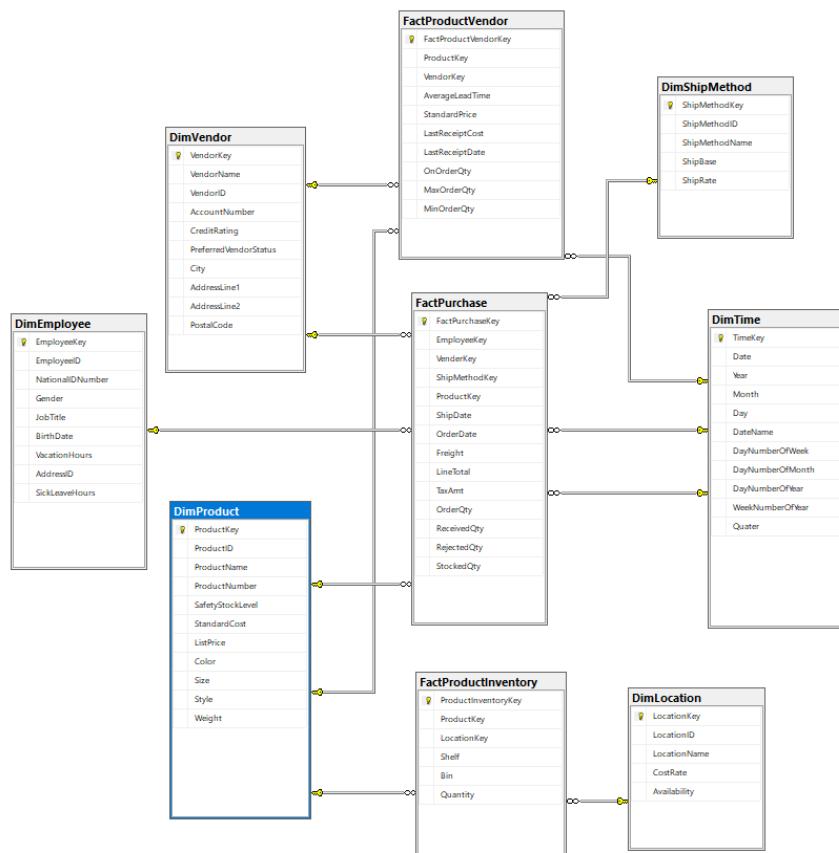
Figure 4-1. Data warehouse diagram

4.1.6. Relationships in the model data warehouse schema

After having the mapping data, we link the fact tables, dim together into relationships. And described in the following table:

No.	Relationship	Relationship Type	Description
1	DimProduct - FactPurchaseKey	1 - n	One product can have one or more purchase order, each purchase order belongs to only one product
2	DimEmployee - FactPurchaseKey	1 - n	One employee can have one or more purchase order, each purchase order belongs to only one employee
3	DimShipMethod - FactPurchaseKey	1 - n	One ship method can have one or more purchase order, each purchase order belongs to only one ship method
4	DimVendor - FactPurchaseKey	1 - n	One vendor can have one or more purchase order, each purchase order belongs to only one vendor
5	DimTime - FactPurchaseKey	1 - n	One time can have one or more purchase order, each purchase order belongs to only one time
6	DimProduct - FactProductVendorKey	1 - n	One product can have one or more product vendor, each product vendor belongs to only one product
7	DimTime - FactProductVendorKey	1 - n	One time can have one or more product vendor, each product

			vendor belongs to only one time
8	DimVendor – FactProductVendorKey	1 - n	One vendor can have one or more product vendor, each product vendor belongs to only one vendor
9	DimProduct – FactProductInventoryKey	1 - n	One product can have one or more product inventory, each product inventory belongs to only one product
10	DimLocation - FactProductInventoryKey	1 - n	One location can have one or more product inventory, each product inventory belongs to only one location

Table 4.14. The relationship in the Data Warehouse model

4.2. ETL processes

4.2.1. Dimension Table's ETL Process

In this section, we will demonstrate the specific ETL process for Dimension tables by using the DimVendor table as an example:

Step 1: Setup the data flow
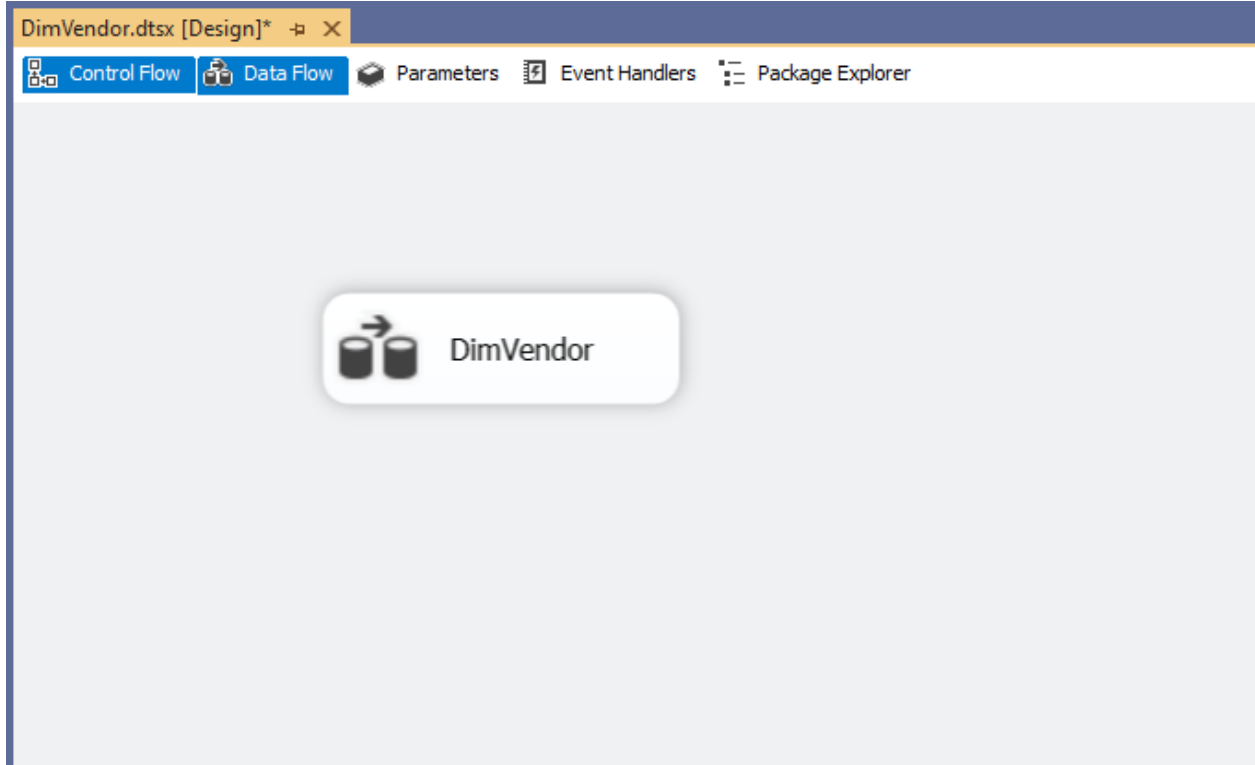
Figure 4-2. Setup the data flow

Step 2: Setup the OLE DB source
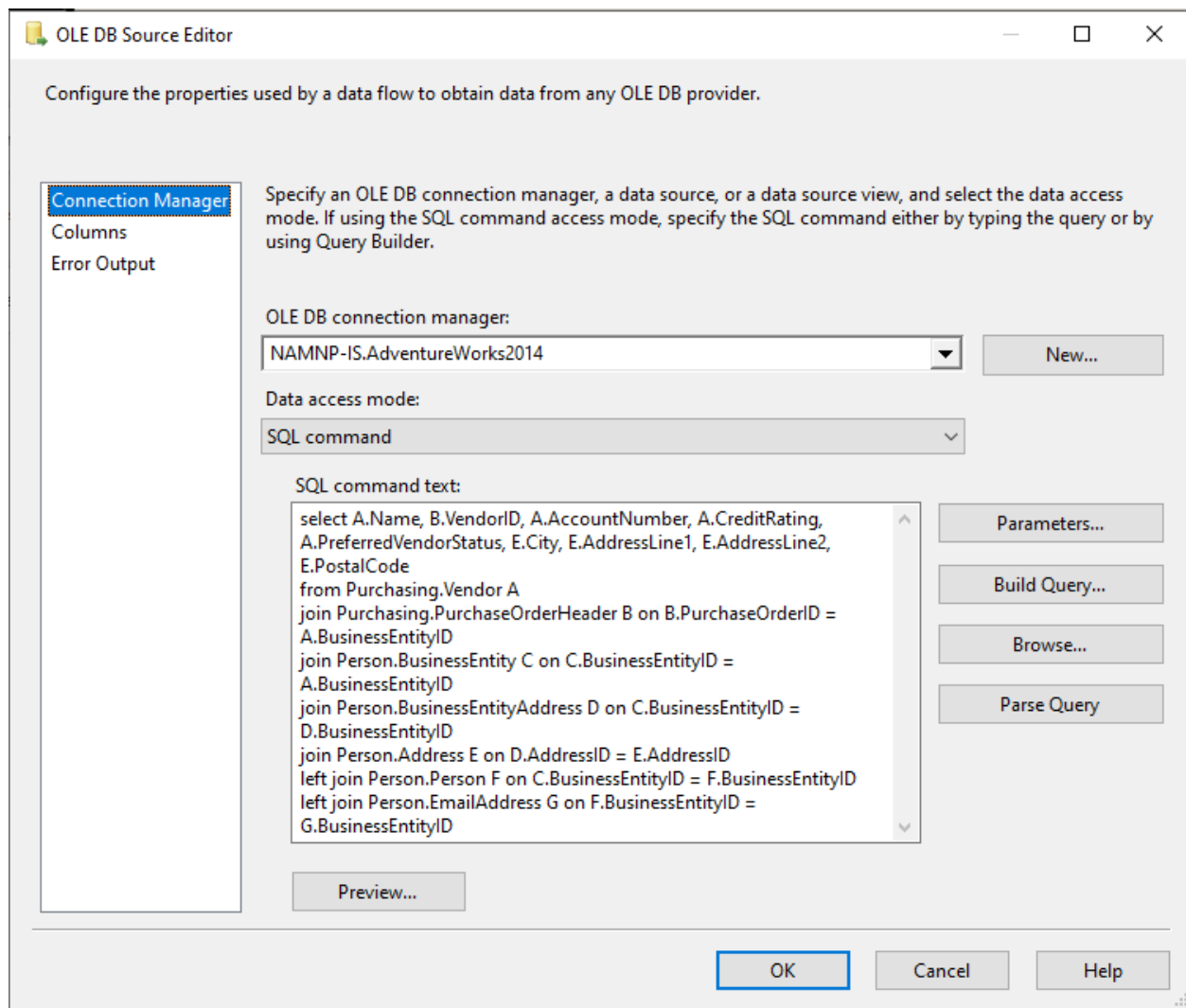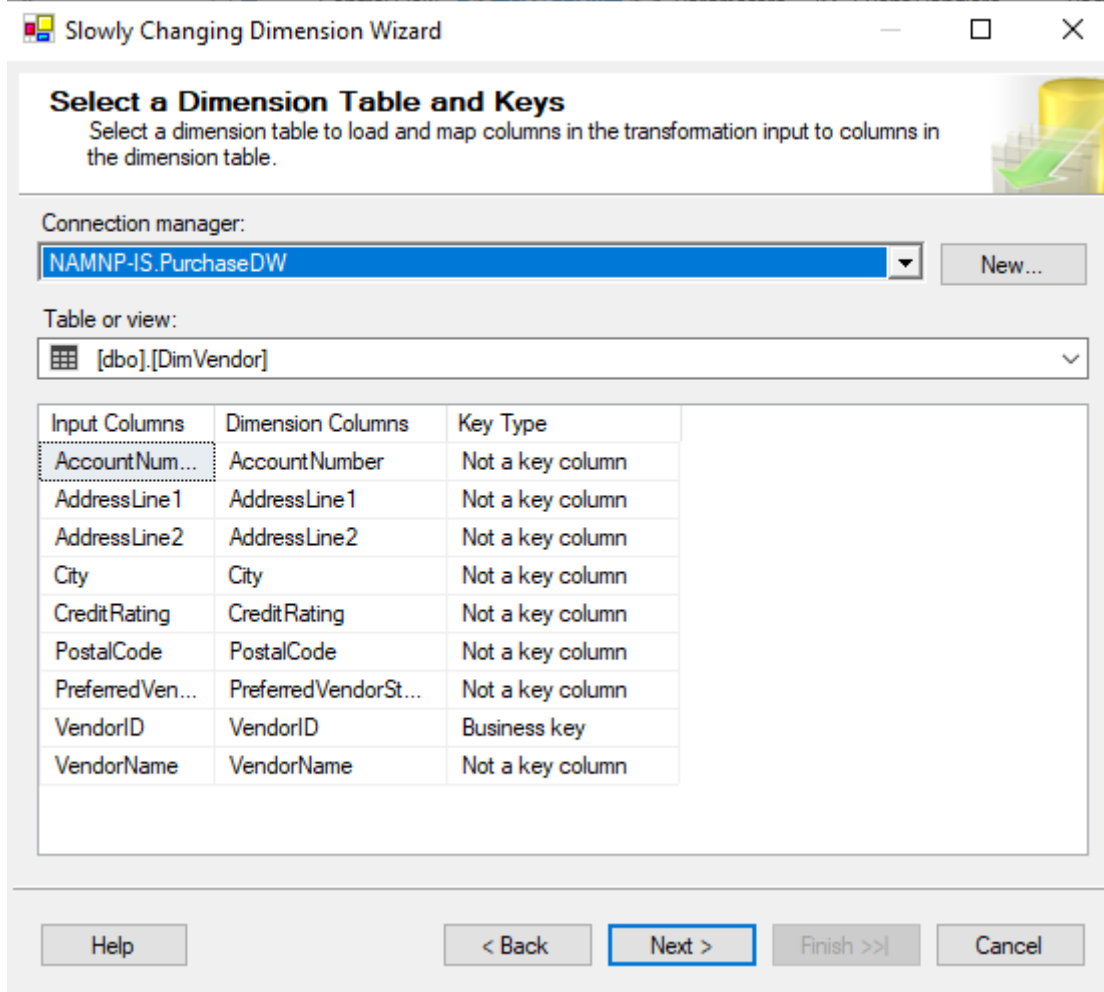
Figure 4-3. Setup the OLE DB source

Step 3: Select column and business key

Figure 4-4. Select column and business key

Step 4: Select column type

Figure 4-5. Select column type

Step 5: Select the OLE DB destination
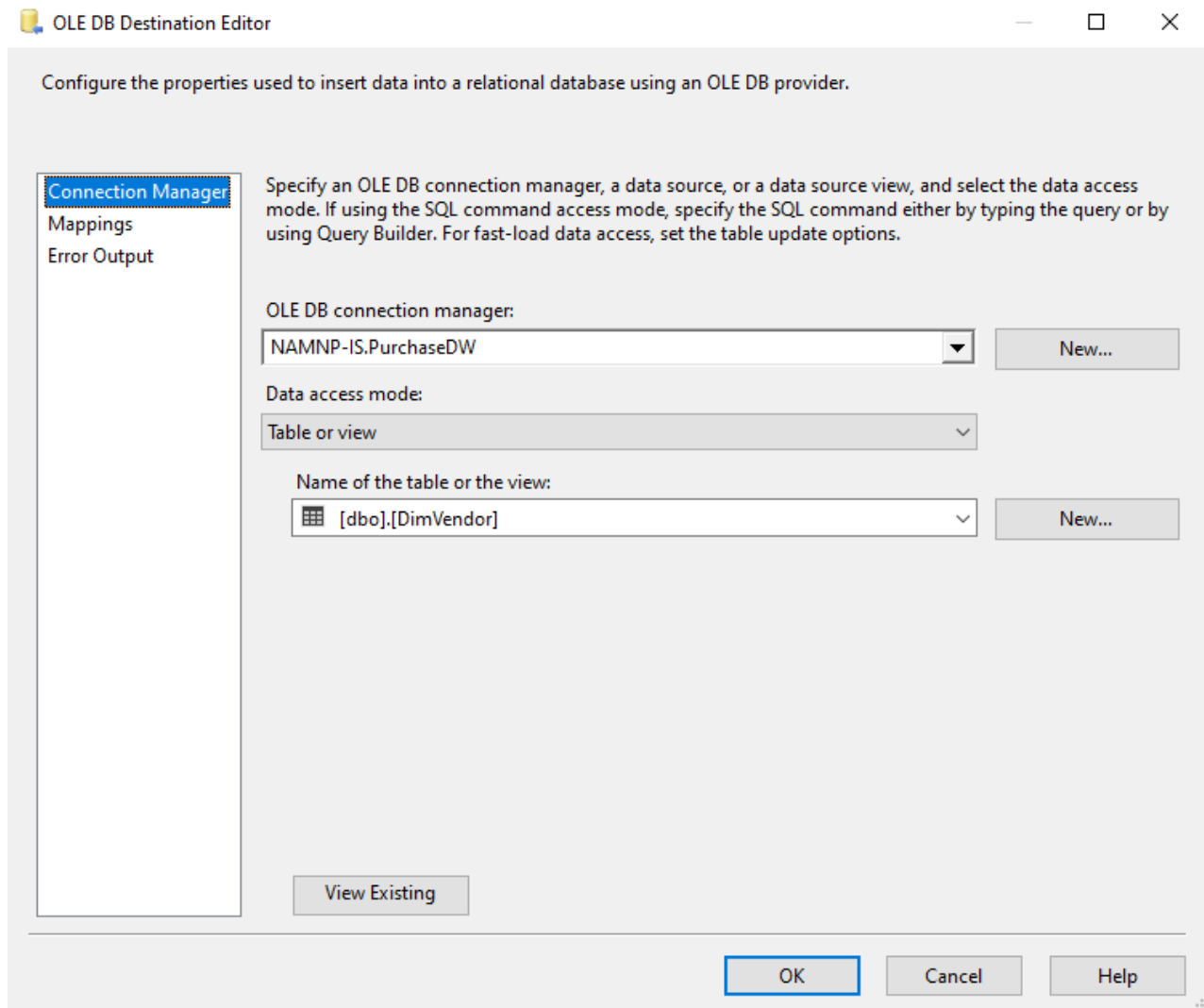
Figure 4-6. Select the OLE DB destination

Step 6: Loading the data into the destination.
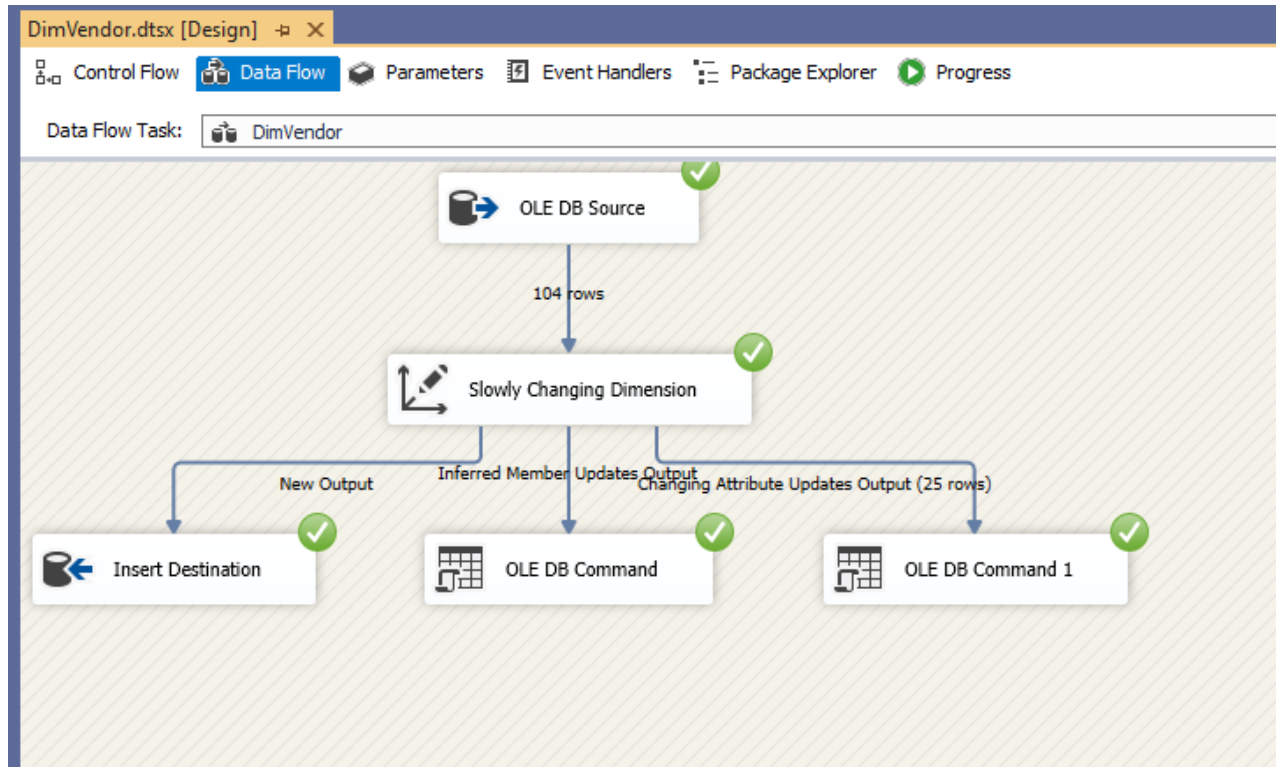
Figure 4-7. Loading the data into the destination.

4.2.2. Fact Table's ETL Process

For the Fact table, we will use the FactPurchase table to represent the specific step in the ETL process.

Step 1: Setup data flow
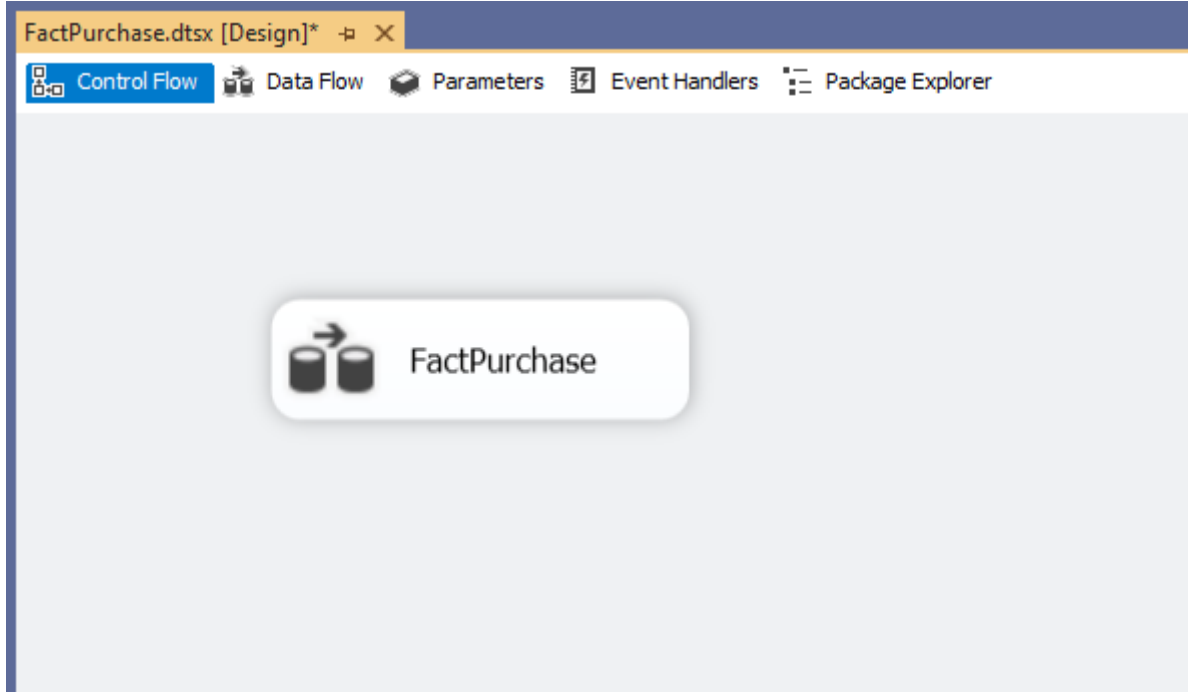
Figure 4-8. Setup data flow

Step 2: Setup data warehouse
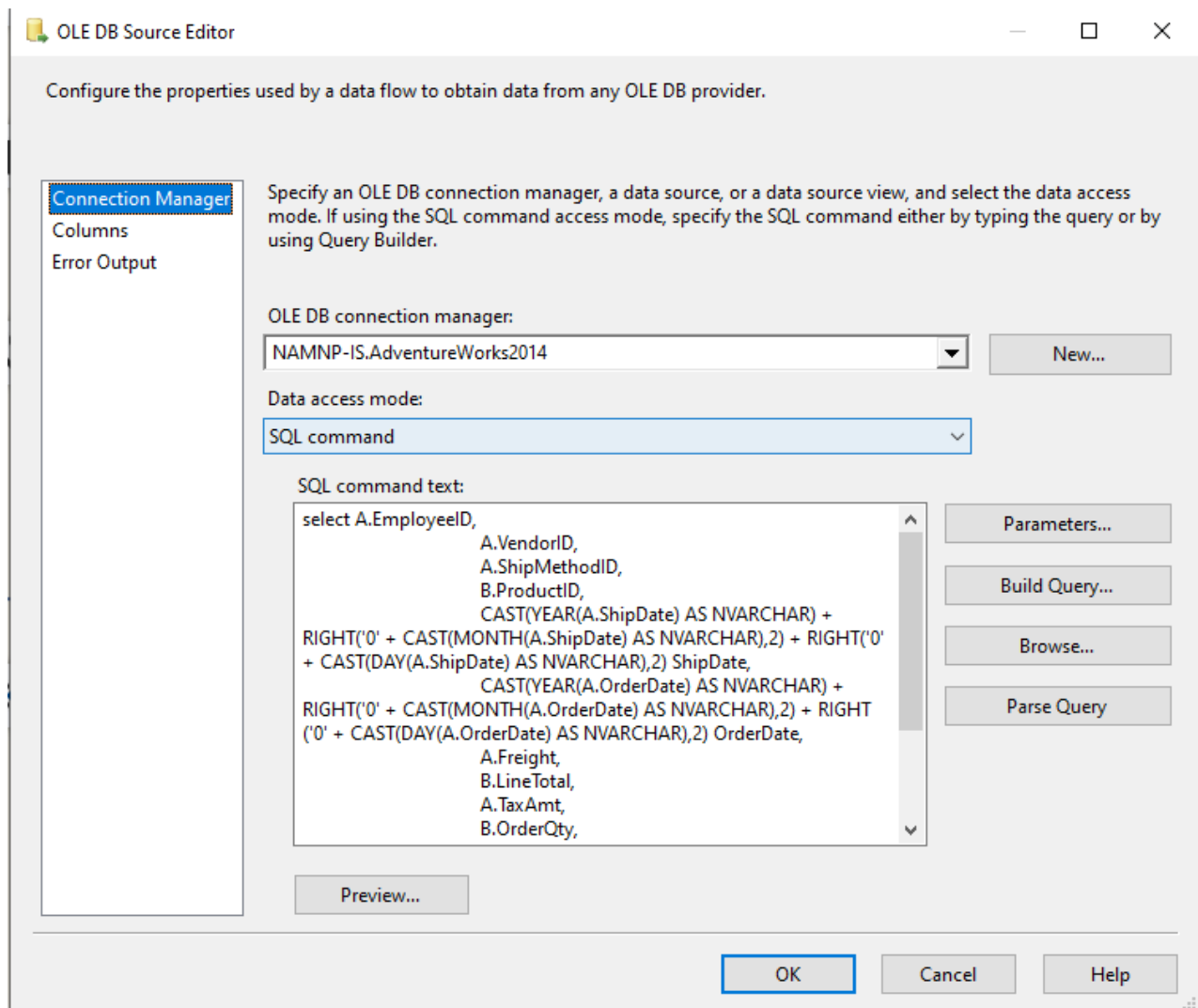
Figure 4-9. Setup data warehouse

Step 3: Select table lookup
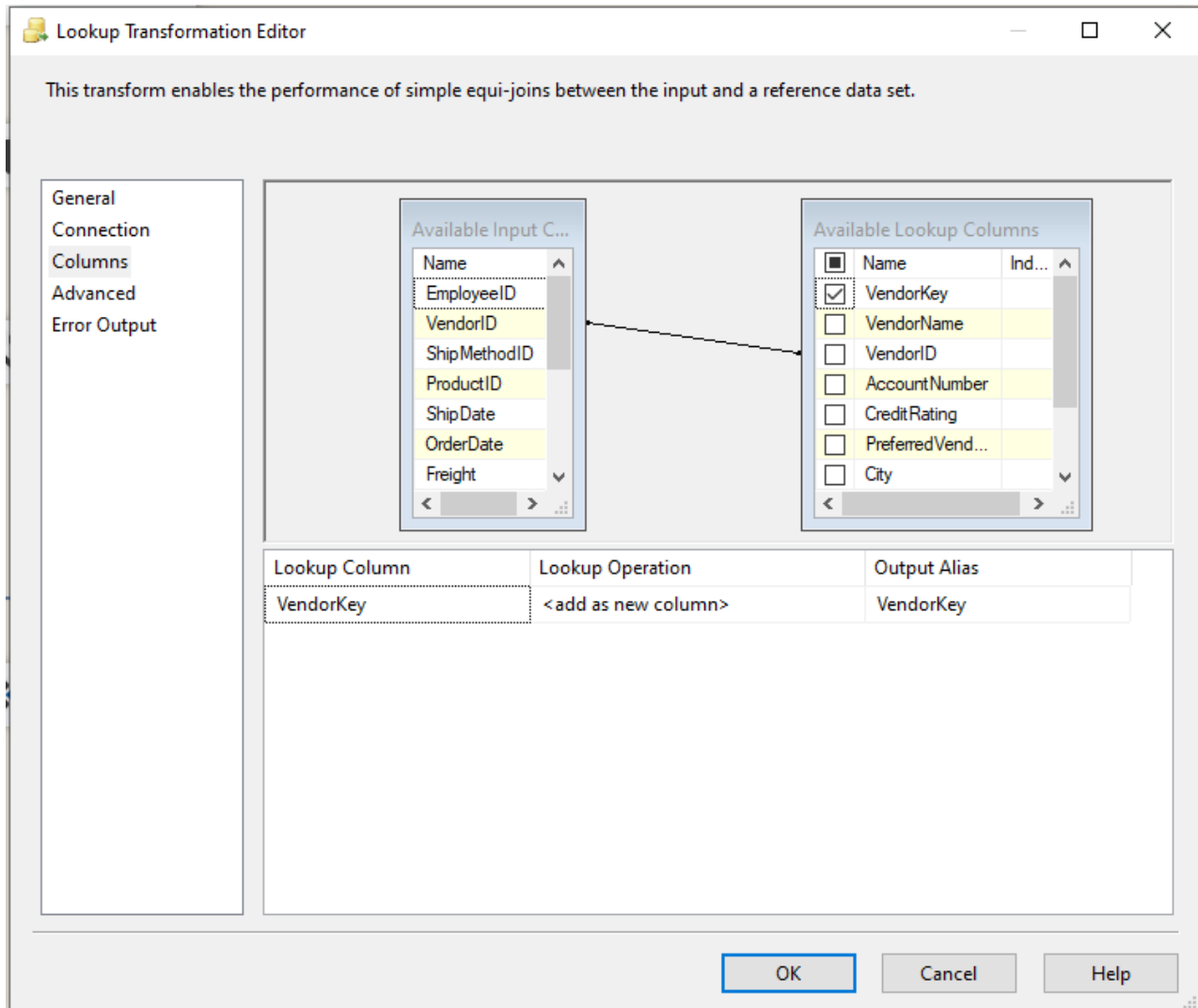
Figure 4-10. Select table lookup

Step 4: Sort data based on the lookup key:
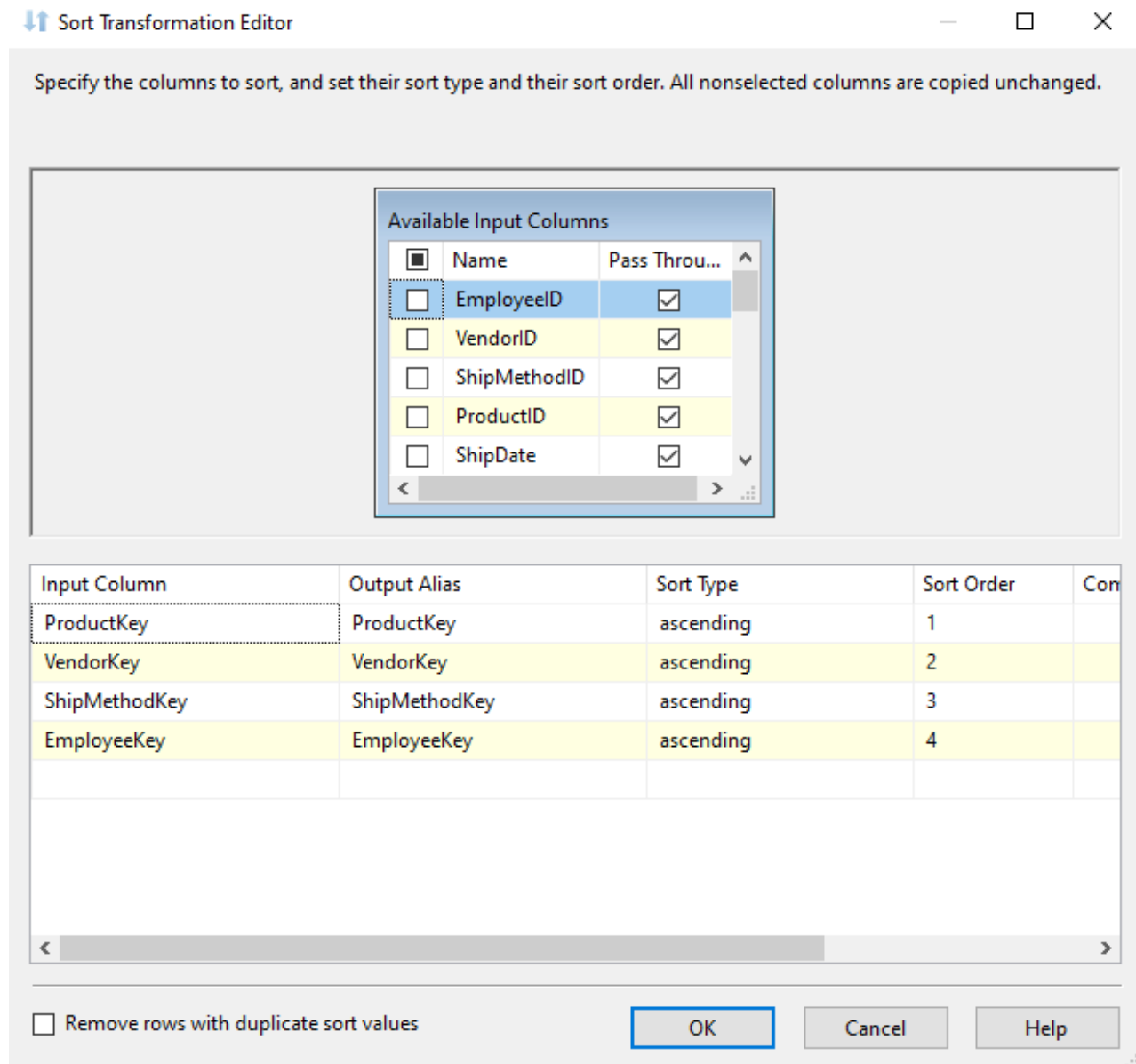
Figure 4-11. Sort data based on the lookup key

Step 5: Setup OLE DB source
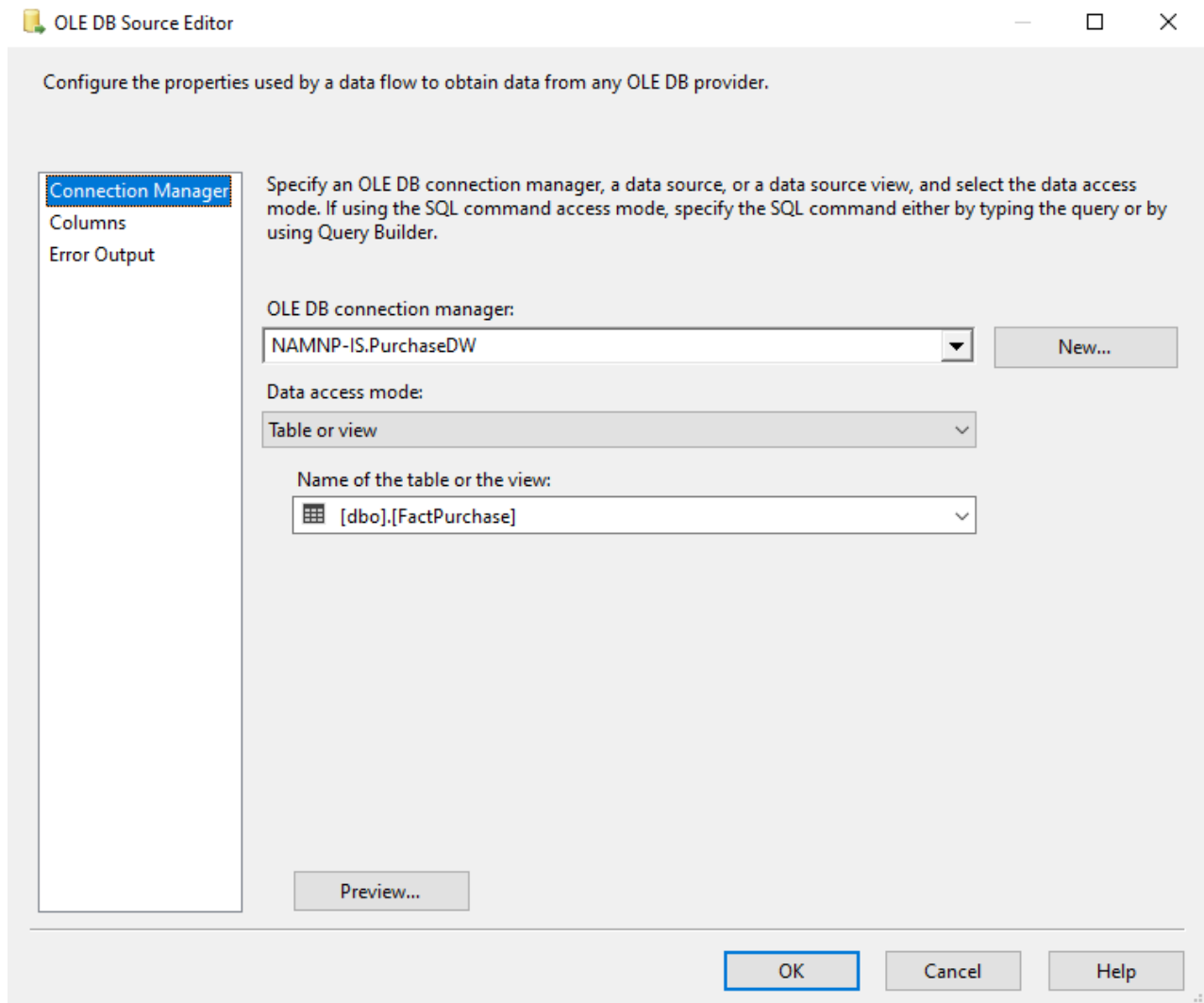
Figure 4-12. Setup OLE DB source

Step 6: Sort data on DB bases on the lookup key

Figure 4-13. Sort data on DB bases on the lookup key

Step 7: Merger - synchronize data

Figure 4-14. The relationship in the Data Warehouse model

Step 8: Setup conditional split
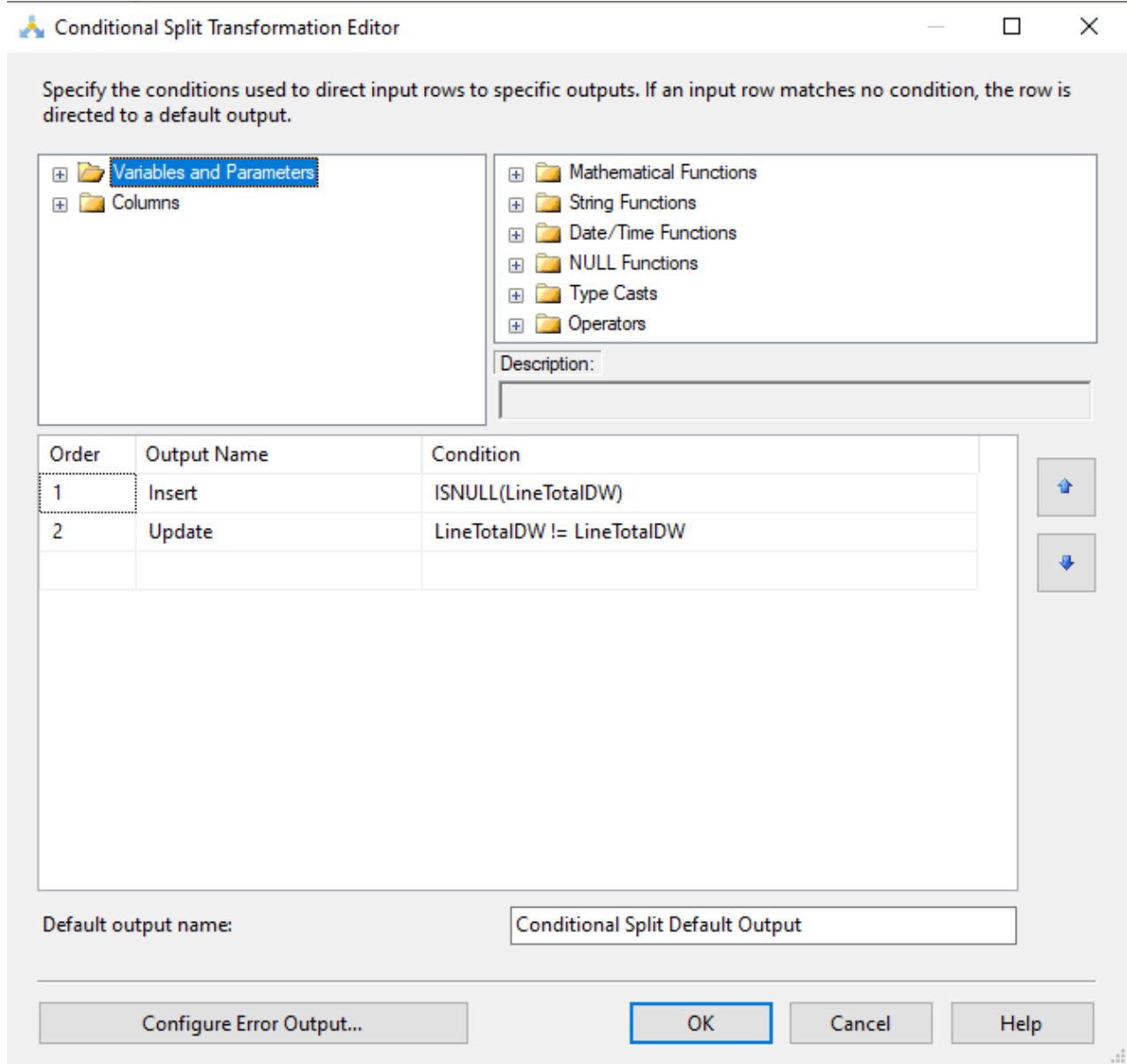
Figure 4-15. Setup conditional split

Step 9: Setup OLD DB destination
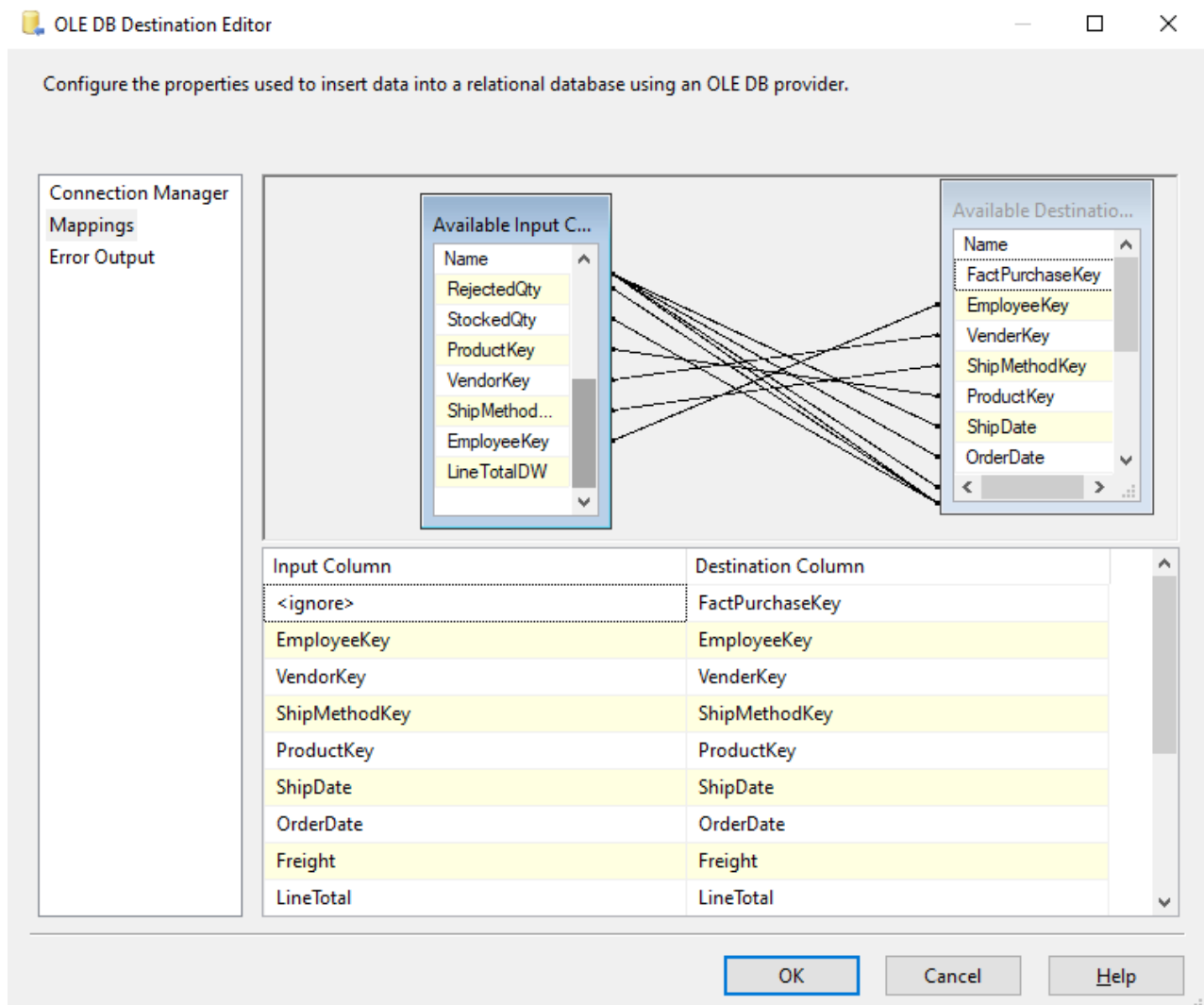
Figure 4-16. Setup OLD DB destination

Step 10: Setup syntax update function
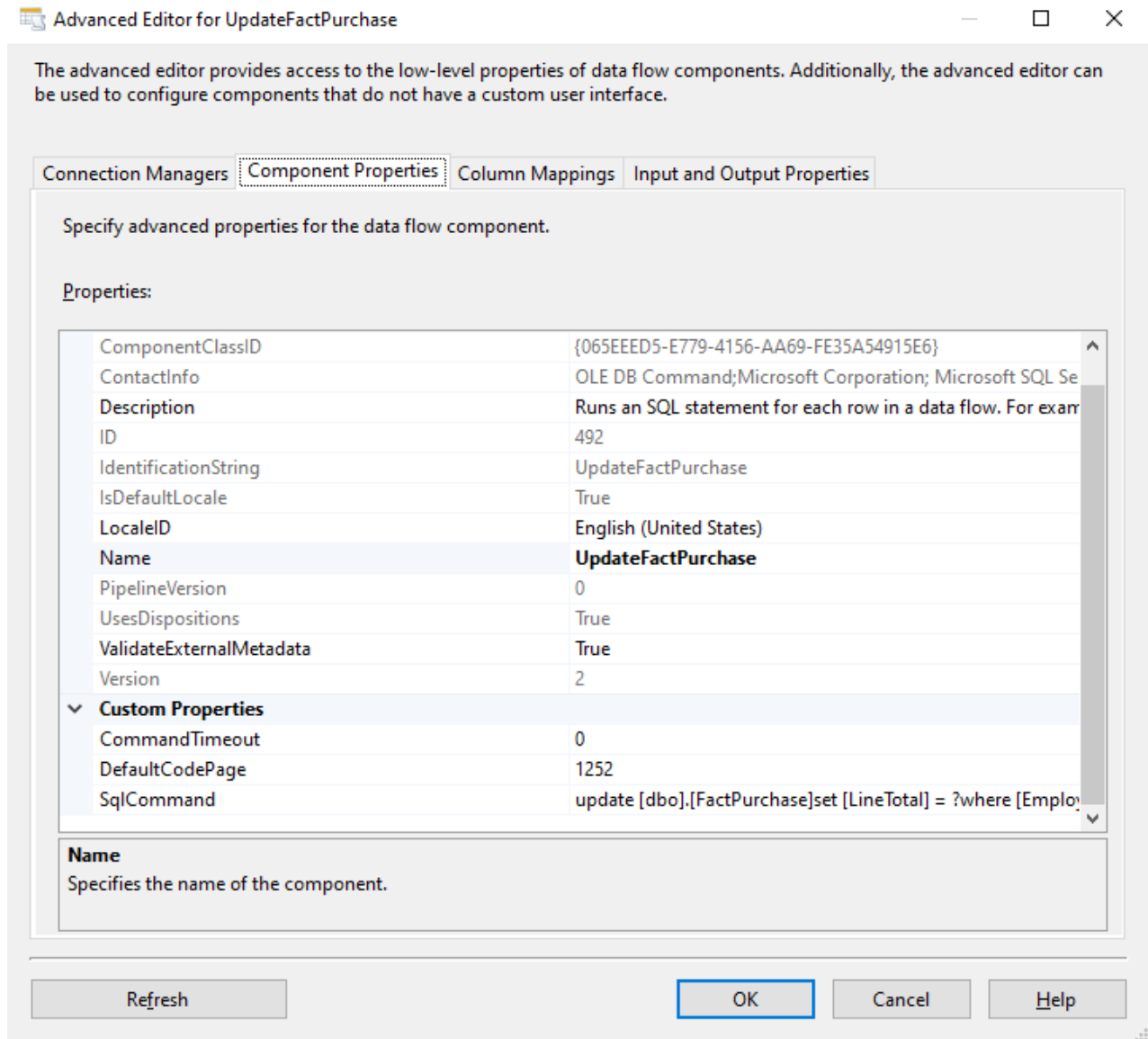
Figure 4-17. Setup syntax update function

Step 11: Load data from the data source into the fact table.
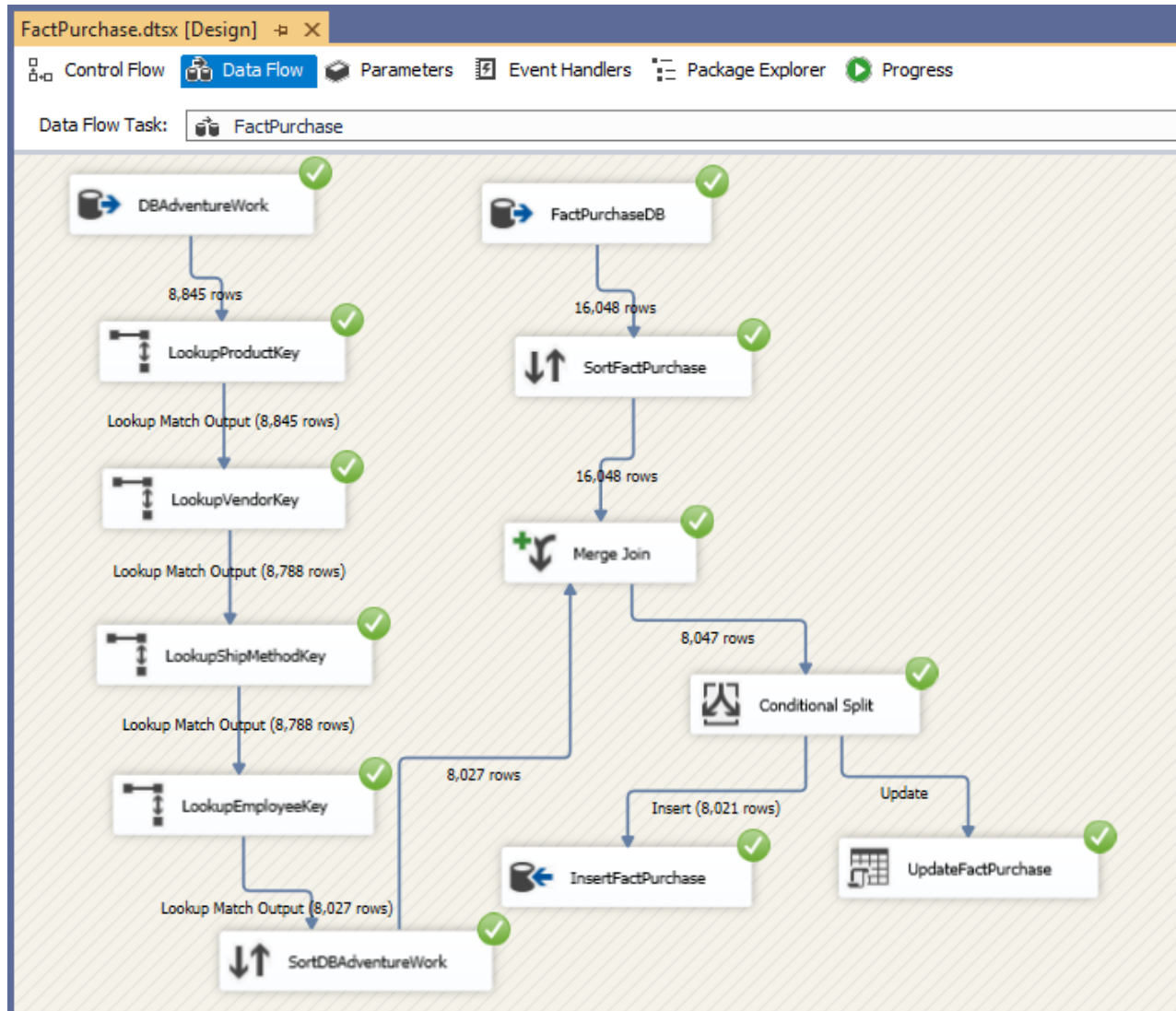
Figure 4-18. Load data from the data source into the fact table

Complete all the steps and here is a summary table of connections between dimension and fact tables and integrating data from the data source into our Data Warehouse to conduct analyst.
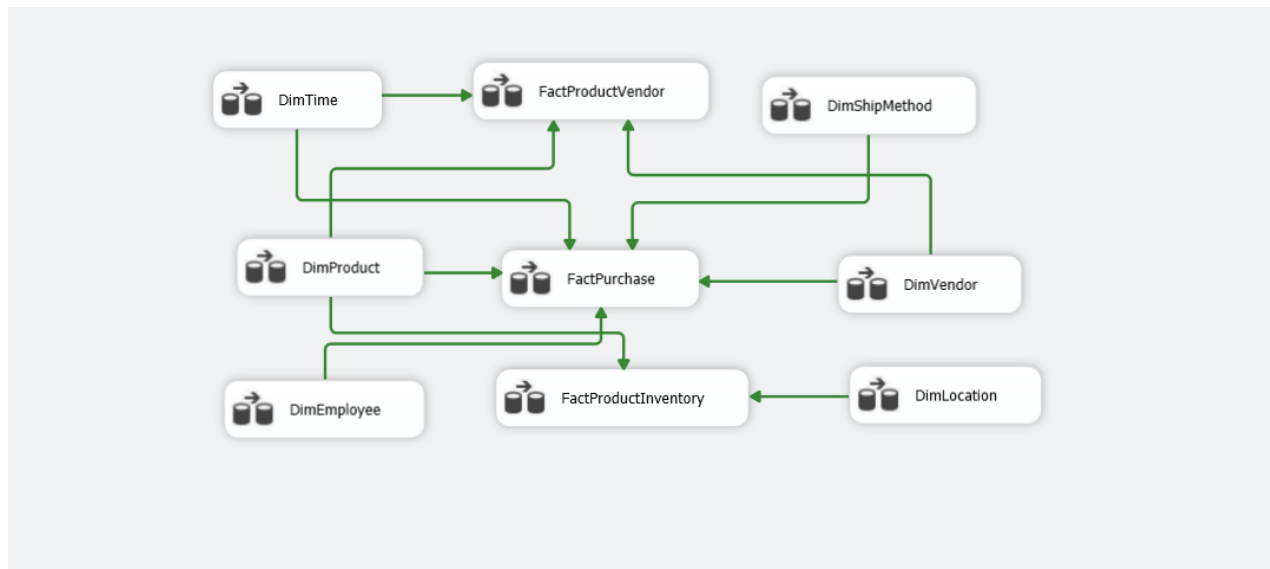
Figure 4-19. Connections between dimension and fact tables

CHAPTER 5. DATA ANALYTICS

In this chapter, we will build SSAS, MDX and OLAP technique and use to analysis the Datawarehouse.

5.1. Data analytics with SSAS technology

Step 1: Create Data Sources.
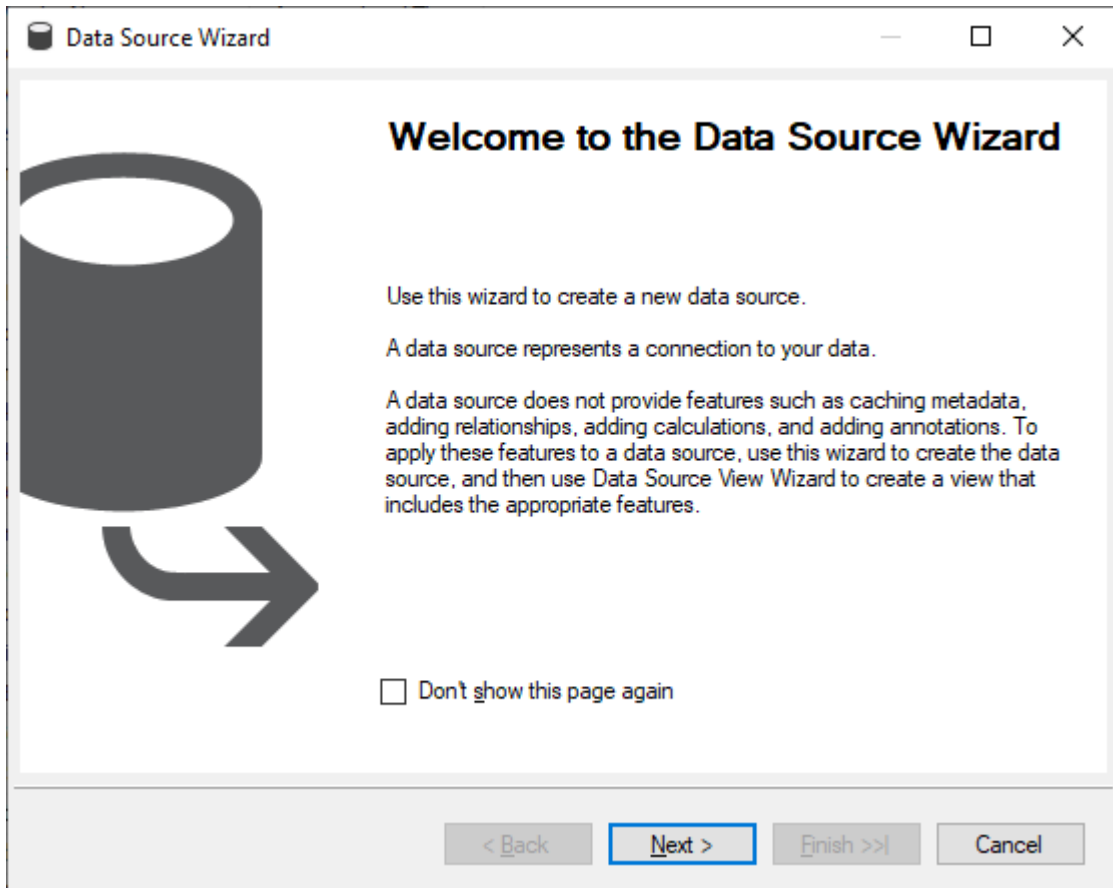
Figure 5-1. Create Data Sources

Step 2: Setup data source
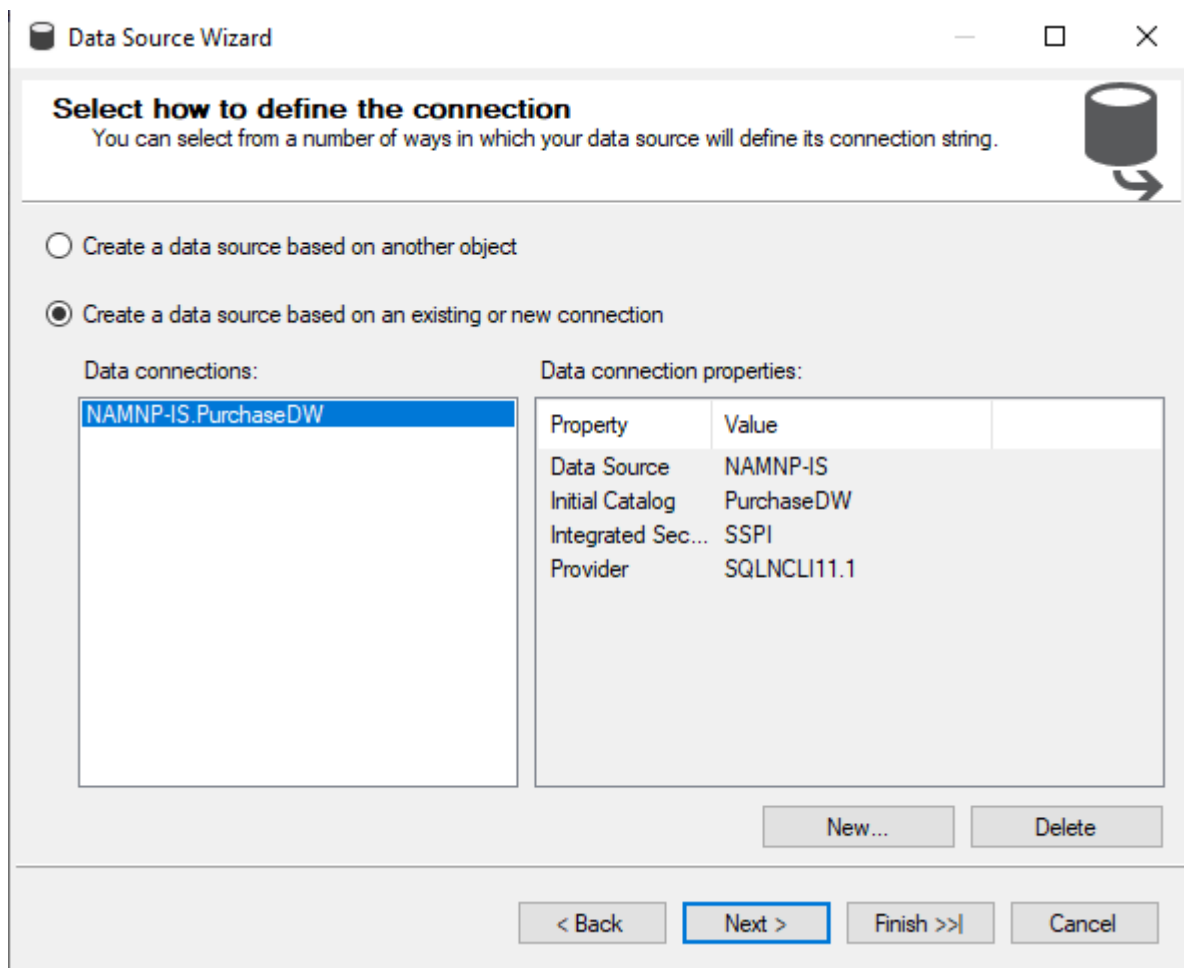
Figure 5-2. Setup data source

Step 3: Select your data warehouse
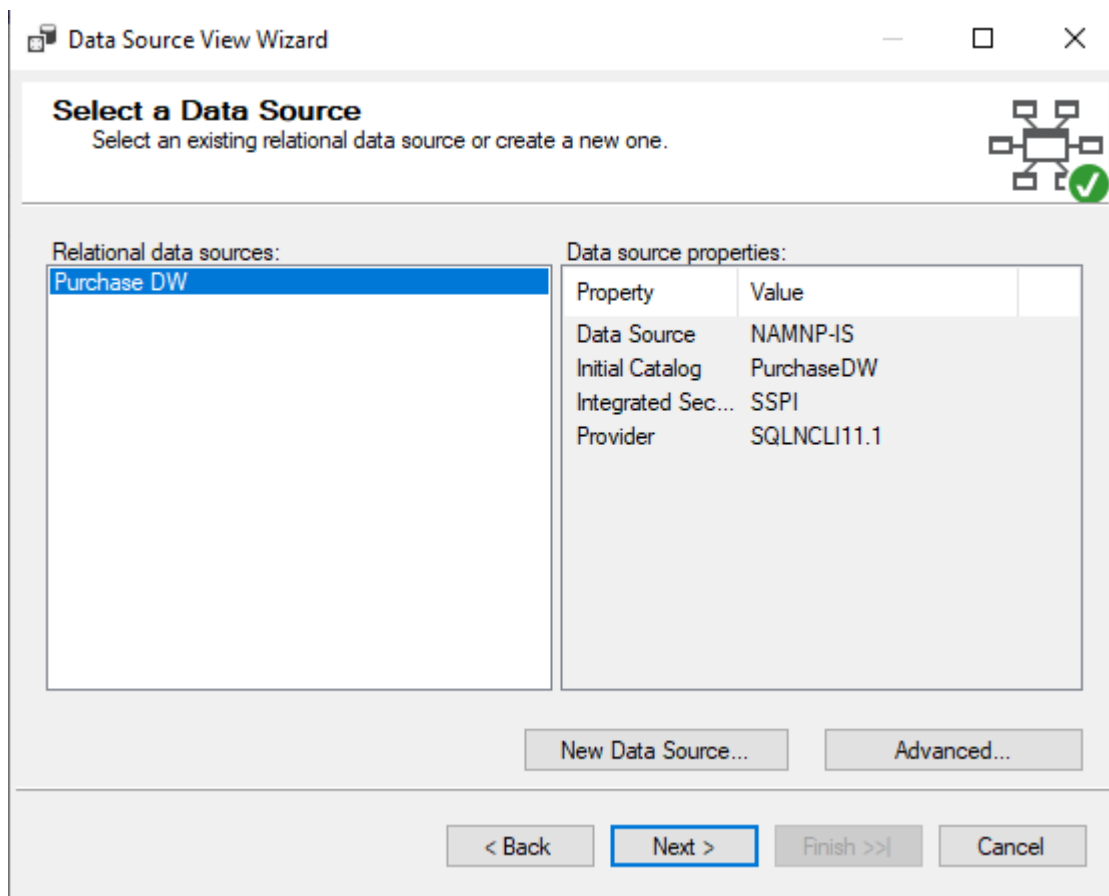
Figure 5-3. Select your data warehouse

Step 4: Select tables and views
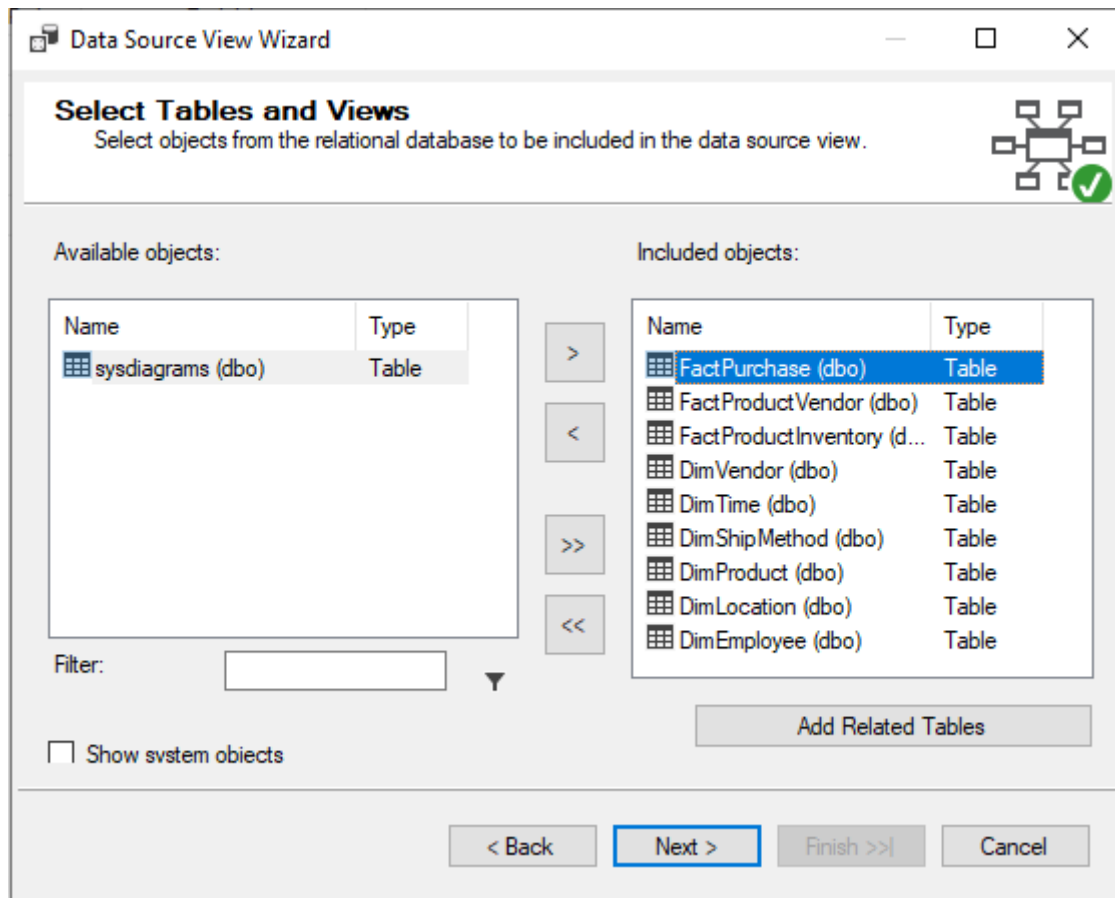
Figure 5-4. Select tables and views

Step 5: Create a Cube Wizard from the data source view
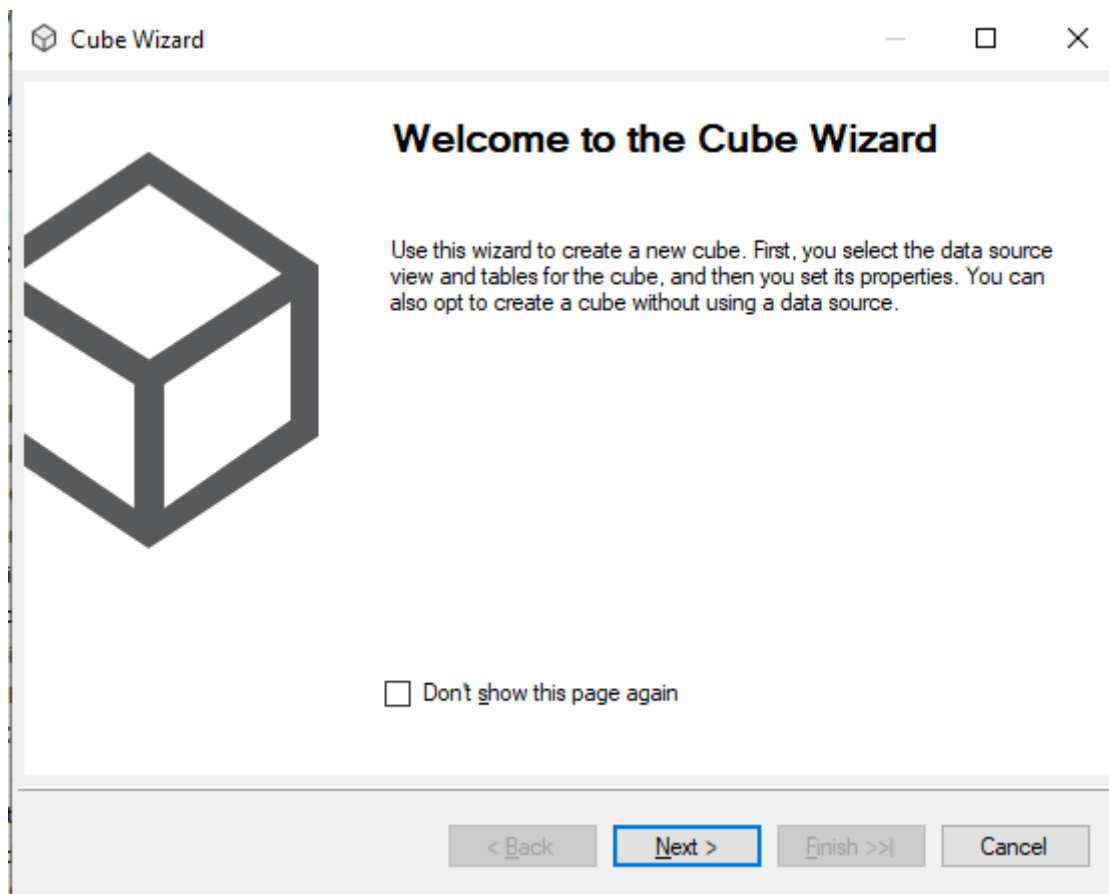
Figure 5-5. Create a Cube Wizard from the data source view

Step 6: Select the measure to use in the Cube Wizard.
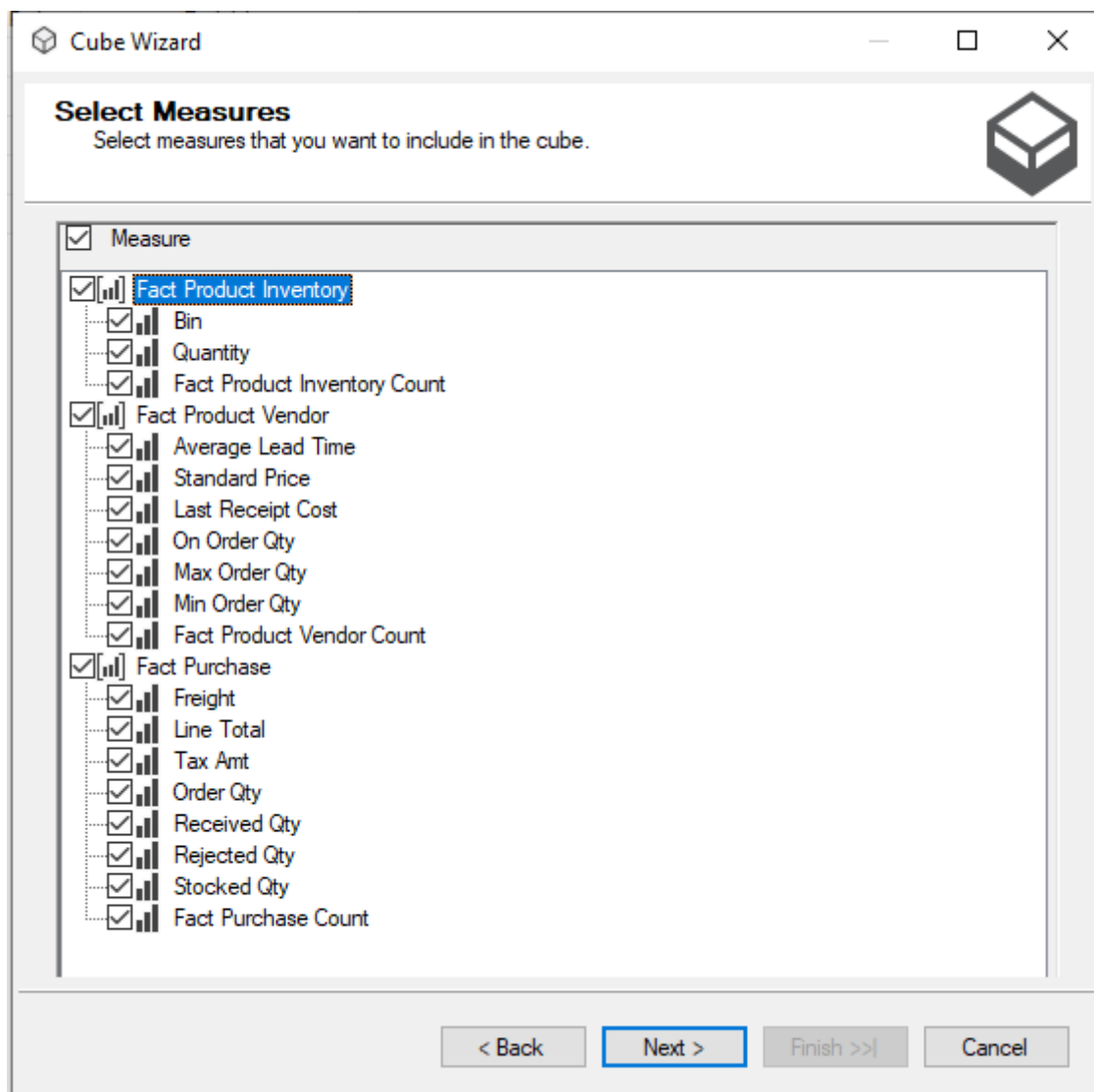
Figure 5-6. Select the measure to use in the Cube Wizard.

Step 7: Select the dimension to use in the Cube Wizard
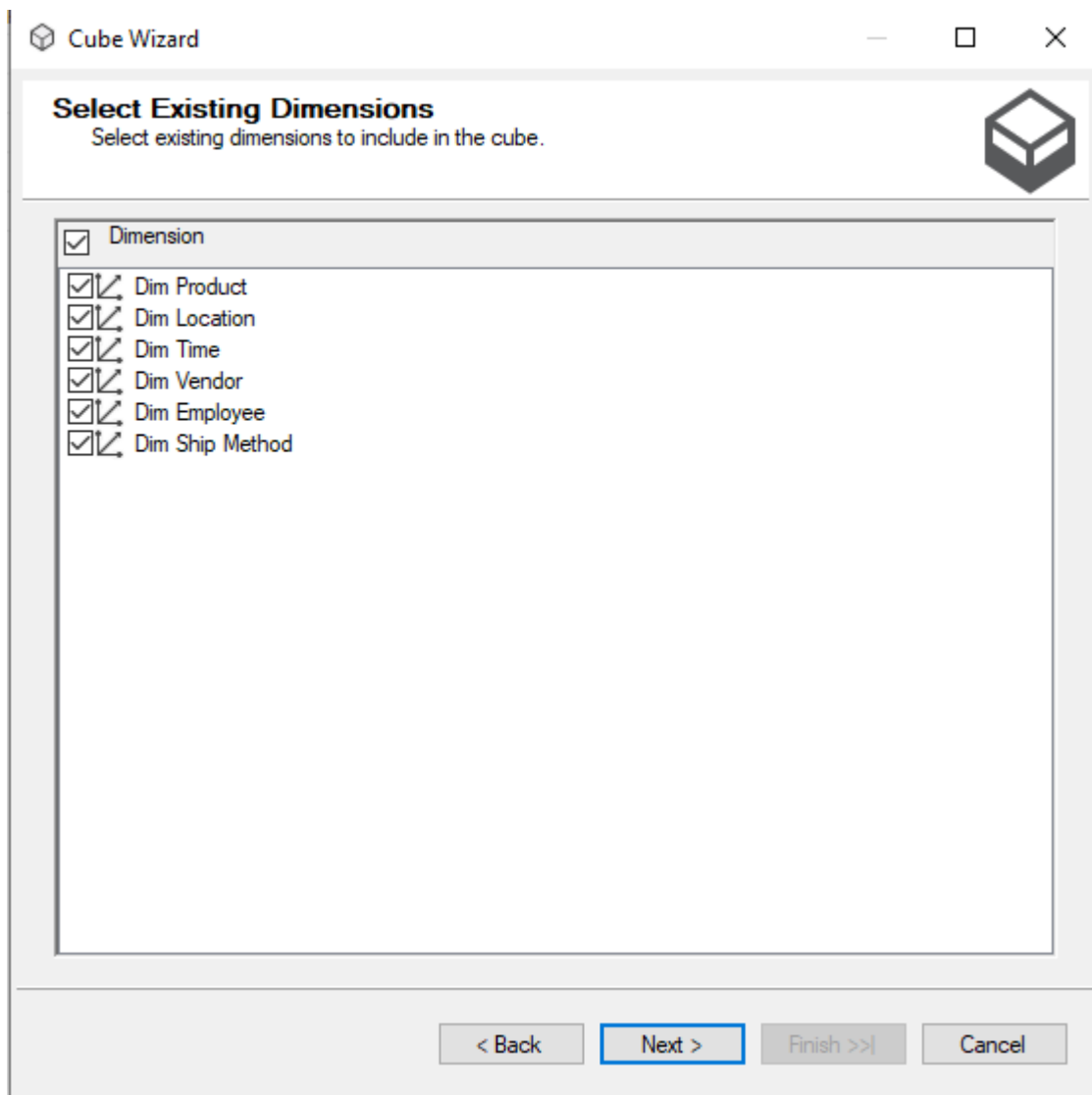
Figure 5-7. Select the dimension to use in the Cube Wizard

Step 8: Confirm all the measures and dimensions in the Cube Wizard, and then click "Finish".
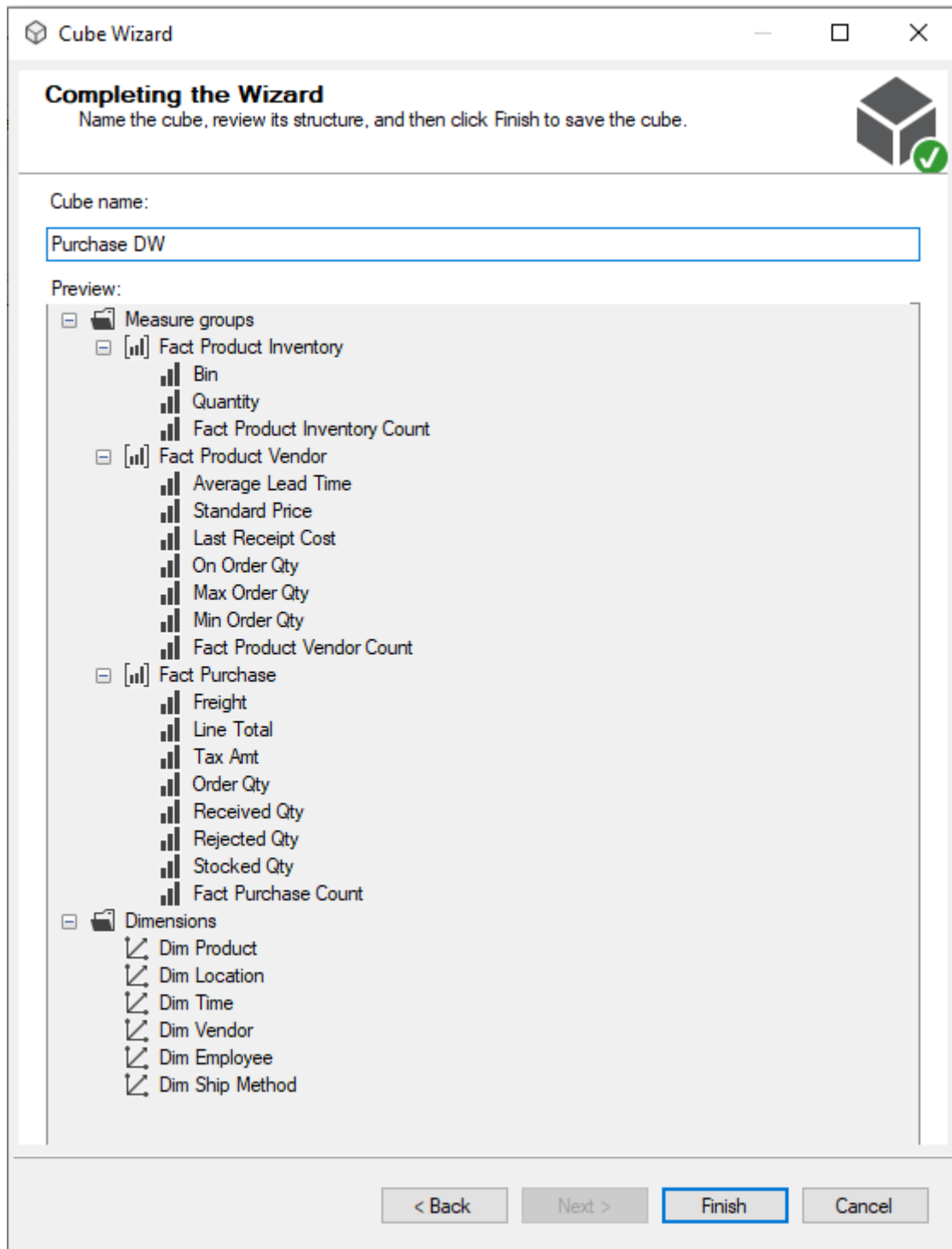
Figure 5-8. Confirm all the measures and dimensions in the Cube Wizard, and then click "Finish".

Complete all the steps and here is your Cube Wizard interface
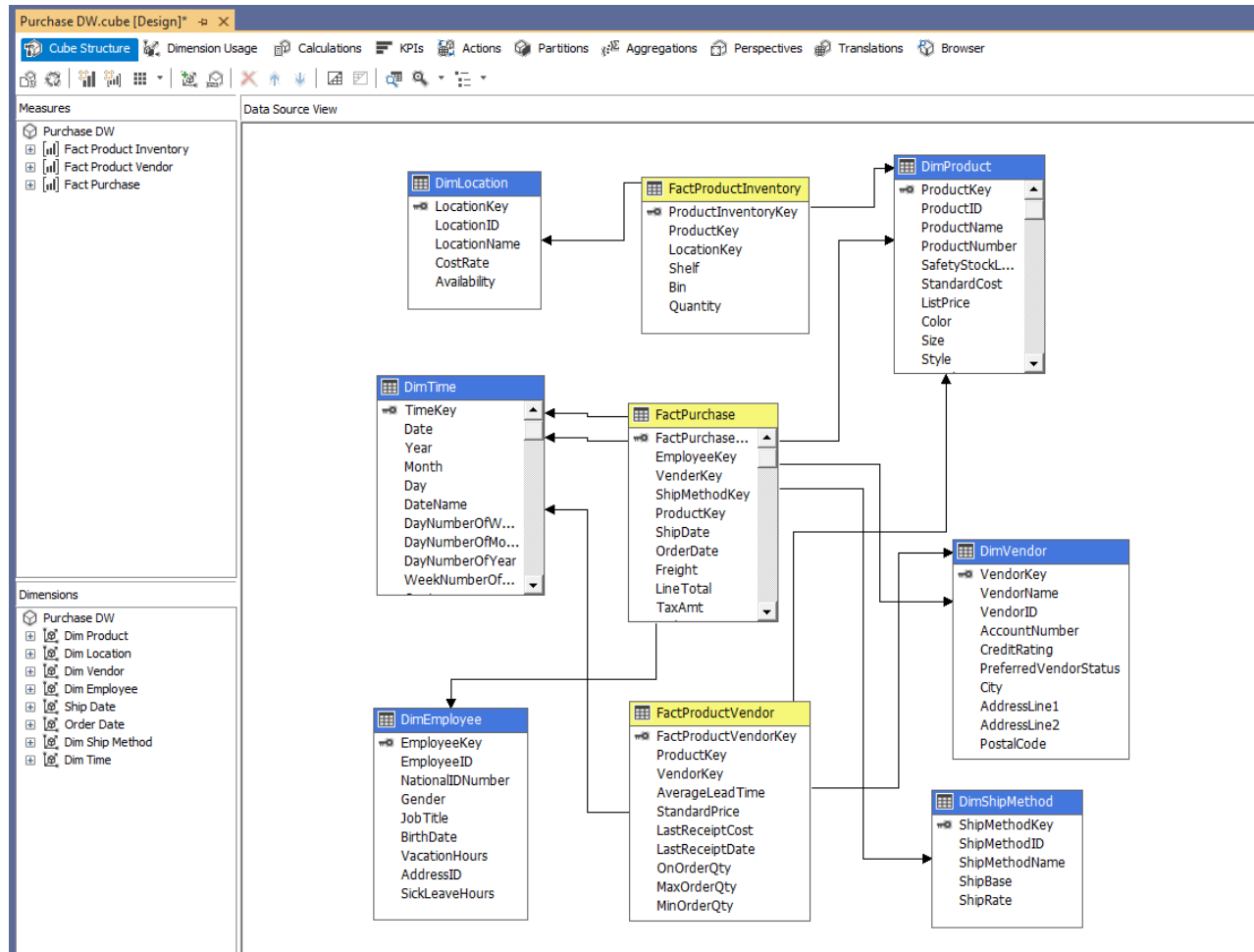
Figure 5-9. Complete all the steps and here is your Cube Wizard interface

5.1.1. Analysis with SSAS

5.1.1.1. Compare shipping costs of different shipping methods
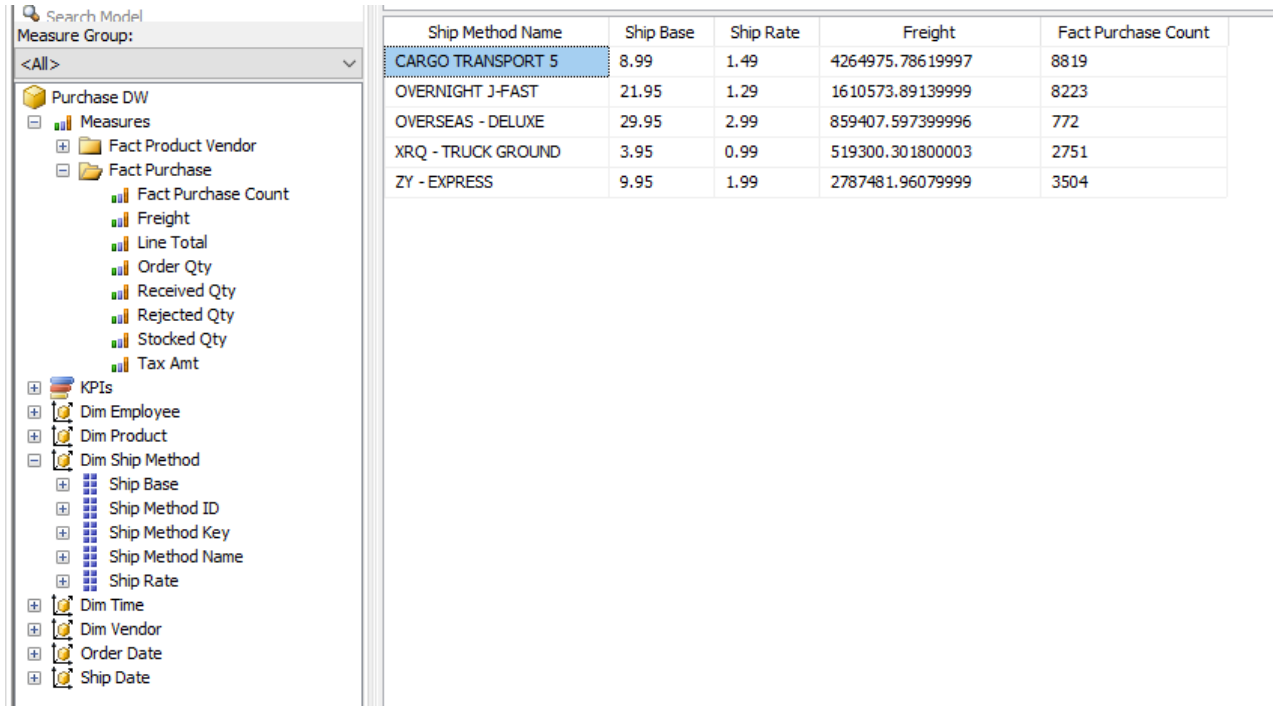
Figure 5-10. Compare shipping costs of different shipping methods

This figure shows the purchase cost of different shipping methods. For businesses, choose a good shipping method and an acceptable cost to balance each other. The 2nd Lowest Ship Base is cargo transport 5 has an 8.99$ in the list and this method is the most used. The lowest used method is overseas – deluxe has 29.95$ due to the highest cost of all methods. Therefore, we have controlled and considered the best and lowest cost shipping method to save costs for businesses.

5.1.1.2. Aggregate order quantity of suppliers

Vendor Name	Min Order Qty	Max Order Qty	On Order Qty
A. Datum Corporation	3423018	18215925	5223114
American Bicycles and Wheels	3423018	18215925	5223114
American Bikes	3423018	18215925	5223114
Anderson's Custom Bikes	3423018	18215925	5223114
Aurora Bike Center	3355900	17858750	5120700
Beaumont Bikes	3423018	18215925	5223114
Bicycle Specialists	3355900	17858750	5120700
Bike Satellite Inc.	3423018	18215925	5223114
Bloomington Multisport	3355900	17858750	5120700
Burnett Road Warriors	3423018	18215925	5223114
Business Equipment Center	3355900	17858750	5120700
Capital Road Cycles	3355900	17858750	5120700
Carlson Specialties	3355900	17858750	5120700
Chicago City Saddles	3355900	17858750	5120700
Circuit Cycles	3355900	17858750	5120700
Compete, Inc.	3355900	17858750	5120700
Competition Bike Training Systems	3355900	17858750	5120700
Consumer Cycles	3355900	17858750	5120700
Continental Pro Cycles	3423018	18215925	5223114
Crowley Sport	3423018	18215925	5223114
Custom Frames, Inc.	3423018	18215925	5223114
Electronic Bike Co.	3423018	18215925	5223114
Electronic Bike Repair & Supplies	3423018	18215925	5223114
Expert Bike Co	3423018	18215925	5223114
Federal Sport	3423018	18215925	5223114
First Rate Bicycles	3423018	18215925	5223114
Fitness Association	3423018	18215925	5223114
GMA Ski & Bike	3355900	17858750	5120700
Green Lake Bike Company	3355900	17858750	5120700
Hill Bicycle Center	3355900	17858750	5120700
Hill's Bicycle Service	3355900	17858750	5120700
Holiday Skate & Cycle	3423018	18215925	5223114
Hybrid Bicycle Center	3423018	18215925	5223114
Illinois Trek & Clothing	3423018	18215925	5223114
Indiana Bicycle Center	3355900	17858750	5120700

Figure 5-11. Aggregate order quantity of suppliers

Analysis the maximum quantity, the minimum quantity should be ordered and the quantity currently on order, rely on it to determine the reliability of the supplier and control the quantity of products ordered.

5.1.1.3. Compare the quality of the employee's work

Figure 5-12. Check the quality of the employee's work

Analyze the number of employee's order participation to reward employees for their hard work. Employees who participate in many orders, the amount of products supplied to the company is large.

5.1.2. Building KPIs system

KPIs Quality: This metric is calculated and compares the quality status of the suppliers in 2014. Reviewing and reconsidering the quality status, then conducting an analysis of why items were rejected, is from internal or suppliers. Screening out bad suppliers

Value expression:

[Measures]. [QualityCheck]=[Measures].[Rejected Qty]/[Measures].[Received Qty]

Figure 5-13. Set value expression

Goal Expression: 0.1.

Status expression:

case

when KPIVALUE("QualityCheck")/KPIGOAL("QualityCheck") >0.9 then -1

when KPIVALUE("QualityCheck")/KPIGOAL("QualityCheck") <=0.9

and KPIVALUE("QualityCheck")/KPIGOAL("QualityCheck") >0.8 then 0

else 1

Trend expression:

case

whenISEMPTY(PARALLELPERIOD([OrderDate].[Year].[Year],1,[OrderDate].[Year])) then 0

when   [Measures].   [QualityCheck]   >   (PARALLELPERIOD([Order Date].[Year].[Year],1, [Order Date].[Year]),[Measures].[QualityCheck]) then -1

when   [Measures].   [QualityCheck]   =   (PARALLELPERIOD([Order Date].[Year].[Year],1, [Order Date].[Year]),[Measures].[QualityCheck]) then 0

else 1

end

KPI

Name:

QualityCheck

Associated measure group:

<All>

Value Expression

```
[Measures].[QualityCheck]
```
✔ No issues found   Ln: 1   Ch: 26   SPC   CRLF

Goal Expression

```
0.1
```
✔ No issues found   Ln: 1   Ch: 4   SPC   CRLF

Status

Status indicator: Traffic light

Status expression:

```
case
when KPIVALUE("QualityCheck")/KPIGOAL("QualityCheck") >0.9 then -1
when KPIVALUE("QualityCheck")/KPIGOAL("QualityCheck") <=0.9
and KPIVALUE("QualityCheck")/KPIGOAL("QualityCheck") >0.8 then 0
else 1
```
✔ No issues found   Ln: 6   Ch: 4   SPC   CRLF

Trend

Trend indicator: Standard arrow

Trend expression:

```
case
when ISEMPTY(PARALLELPERIOD([Order Date].[Year].[Year],1, [Order Date].[Year] )) then 0
when [Measures].[Fact Purchase Count] > (PARALLELPERIOD([Order Date].[Year].[Year],1, [Order Date].[Year] ), [Measures].
    [Fact Purchase Count]) then 1
when [Measures].[Fact Purchase Count] = (PARALLELPERIOD([Order Date].[Year].[Year],1, [Order Date].[Year] ), [Measures].
```
✔ No issues found   Ln: 6   Ch: 4   SPC   CRLF

Figure 5-14. Set KPI value

Row Labels	QualityCheck Goal	QualityCheck Status	QualityCheck Trend
A. Datum Corporation	0.1		⬆
American Bicycles and Wheels	0.1		⬇
American Bikes	0.1		➡
Anderson's Custom Bikes	0.1		⬇
Aurora Bike Center	0.1		➡
Beaumont Bikes	0.1		⬆
Bicycle Specialists	0.1		⬆
Bike Satellite Inc.	0.1		⬆
Bloomington Multisport	0.1		➡
Burnett Road Warriors	0.1		⬇
Business Equipment Center	0.1		⬆
Capital Road Cycles	0.1		⬆
Carlson Specialties	0.1		⬆
Chicago City Saddles	0.1		⬆
Circuit Cycles	0.1		⬆
Compete, Inc.	0.1		⬇
Competition Bike Training Systems	0.1		⬇
Consumer Cycles	0.1		⬆
Continental Pro Cycles	0.1		⬆
Crowley Sport	0.1		⬆
Custom Frames, Inc.	0.1		⬆
Electronic Bike Co.	0.1		⬇
Electronic Bike Repair & Supplies	0.1		⬇
Expert Bike Co	0.1		⬇
Federal Sport	0.1		⬆
First Rate Bicycles	0.1		⬆
Fitness Association	0.1		⬇

Figure 5-15. Quality status of suppliers in 2014

5.2. Data analytics with MDX and OLAP technique

5.2.1. Top 5 employees with the highest number of times to order from the vendor in 2014

```
select topcount ([Dim Employee].[Employee ID].children,5,[Measures].[Fact Purchase Count]) on 1,
{[Measures].[Fact Purchase Count]} on 0
from [Purchase DW]
where [Order Date].[Year].&[2014]
```

	Fact Purchase Count
257	2103
251	1803
253	1701
261	1668
255	1473

Figure 5-16. Top 5 employees with the highest number of times to order from the vendor in 2014

In 2014, the manager analyzed how well the employees in this module worked and then rewarded the employees who met or exceeded the target.

5.2.2. Top 5 vendors with the fastest average lead times on total products in 2014

```
select bottomcount ([Dim Vendor].[Vendor Name].children,5,[Measures].[Fact Product Vendor Count]) on 1,
{[Measures].[Fact Product Vendor Count],[Measures].[Average Lead Time]} on 0
from [Purchase DW]
where [Order Date].[Year].&[2014]
```

	Fact Product Vendor Count	Average Lead Time
WestAmerica Bicycle Co.	23000	447350
West Junction Cycles	23000	447350
Sports House	23000	447350
Sport Fan Co.	23000	447350
Speed Corporation	23000	447350

Figure 5-17. Top 5 vendors with the fastest average lead times on total products in 2014

This measure helps companies calculate the average number of days in supplier lead time on purchased materials, lead time from a supplier can entail higher inventory levels,

higher inventory carrying costs, and associated logistics costs. Lead time is one of the most important measures in inventory control, reducing lead time in manufacturing can increase output and revenue. From there, compare and evaluate suppliers with each other, choose the supplier with the shortest average lead time.

5.2.3. Compare the standard price of a product between different vendors in 2014

Figure 5-18. Compare the standard price of a product between different vendors in 2014

In the purchasing module, the price also plays a very important part, here, our team has some calculations to compare the standard prices between suppliers, from which we can give the best prices. decide to choose a supplier, or as a strategy to cut costs.
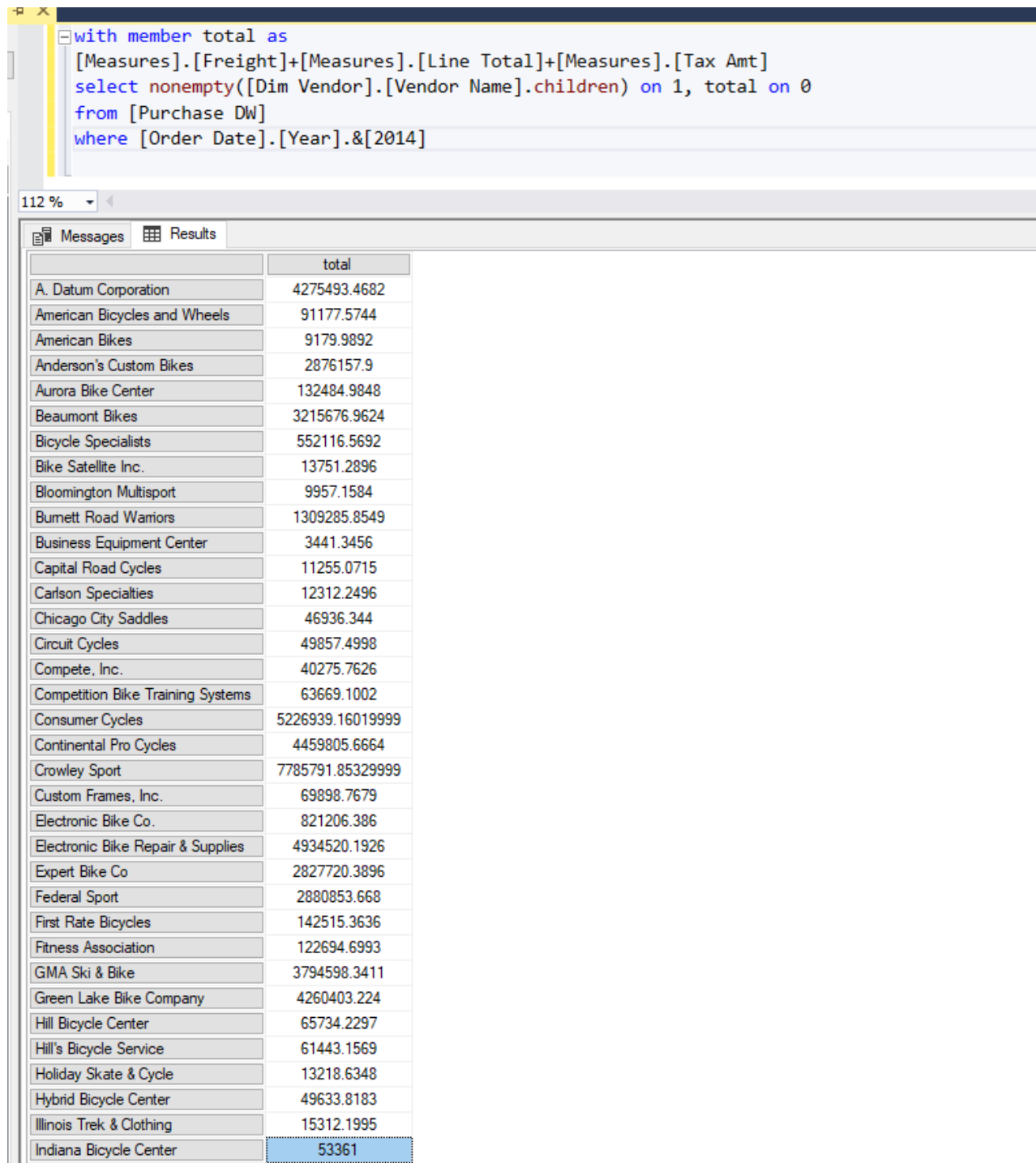
5.2.4. List of quantity of goods received, rejected goods and rate of rejected goods based on 2014 data

```
with member RateOfRejectedGoods as [Measures].[Rejected Qty]/[Measures].[Received Qty]
select nonempty([Dim Vendor].[Vendor Name].children) on 1,
{[Measures].[Rejected Qty],[Measures].[Received Qty],RateOfRejectedGoods} on 0
from [Purchase DW]
```

112 %

Messages   Results

	Rejected Qty	Received Qty	RateOfRejectedGoods
A. Datum Corporation	4695	176319	2.66278733432018E-02
American Bicycles and Wheels	18	2313	7.78210116731518E-03
American Bikes	9	414	2.17391304347826E-02
Anderson's Custom Bikes	4614	150084	3.07427840409371E-02
Aurora Bike Center	18	2349	7.66283524904215E-03
Beaumont Bikes	1722	74670	2.30614704700683E-02
Bicycle Specialists	3126	73020	0.042810188989318
Bike Satellite Inc.	27	423	6.38297872340425E-02
Bloomington Multisport	0	828	0
Burnett Road Warriors	4530	74670	6.06669345118521E-02
Business Equipment Center	18	405	4.44444444444444E-02
Capital Road Cycles	9	405	2.22222222222222E-02
Carlson Specialties	9	405	2.22222222222222E-02
Chicago City Saddles	522	16020	3.25842696629213E-02
Circuit Cycles	18	1341	1.34228187919463E-02
Compete, Inc.	288	8145	3.53591160220995E-02
Competition Bike Training Systems	36	1413	2.54777070063694E-02
Consumer Cycles	4776	166992	2.86001724633515E-02
Continental Pro Cycles	7830	150165	5.21426430925981E-02
Crowley Sport	4776	272019	1.75575970796158E-02
Custom Frames, Inc.	18	1773	1.01522842639594E-02
Electronic Bike Co.	984	76320	1.28930817610063E-02
Electronic Bike Repair & Supplies	8322	176892	4.70456549759175E-02
Expert Bike Co	360	96915	3.71459526389104E-03
Federal Sport	4944	153465	3.22158146808719E-02
First Rate Bicycles	18	6903	2.60756192959583E-03
Fitness Association	2880	76320	3.773584905660381E-02
GMA Ski & Bike	4857	209811	2.31494058938759E-02
Green Lake Bike Company	8079	150165	5.38008190989911E-02
Hill Bicycle Center	18	1431	1.25786163522013E-02
Hill's Bicycle Service	0	1674	0
Holiday Skate & Cycle	0	414	0
Hybrid Bicycle Center	18	1584	1.13636363636364E-02
Illinois Trek & Clothing	3	411	7.2992700729927E-03
Indiana Bicycle Center	17994	341079	5.27561063565919E-02
Inline Accessories	18	1431	1.25786163522013E-02
Inner City Bikes	4941	150165	0.032903805813605

Figure 5-19. List of quantity of goods received, rejected goods and rate of rejected goods based on 2014

In this table, we have listed the data related to the number of imported goods, returned goods and the rate of returned goods, thereby giving an objective assessment of the quality of the goods we import. , we will consider ratios that are too high, because if the ratio is too high, the product has to be re-evaluated for quality. In addition, there are products with a zero return rate, which is very good for the company.

5.2.5. Total purchase price list for each vendor

```
with member total as
  [Measures].[Freight]+[Measures].[Line Total]+[Measures].[Tax Amt]
  select nonempty([Dim Vendor].[Vendor Name].children) on 1, total on 0
  from [Purchase DW]
  where [Order Date].[Year].&[2014]
```

112 %

Messages Results

	total
A. Datum Corporation	4275493.4682
American Bicycles and Wheels	91177.5744
American Bikes	9179.9892
Anderson's Custom Bikes	2876157.9
Aurora Bike Center	132484.9848
Beaumont Bikes	3215676.9624
Bicycle Specialists	552116.5692
Bike Satellite Inc.	13751.2896
Bloomington Multisport	9957.1584
Burnett Road Warriors	1309285.8549
Business Equipment Center	3441.3456
Capital Road Cycles	11255.0715
Carlson Specialties	12312.2496
Chicago City Saddles	46936.344
Circuit Cycles	49857.4998
Compete, Inc.	40275.7626
Competition Bike Training Systems	63669.1002
Consumer Cycles	5226939.16019999
Continental Pro Cycles	4459805.6664
Crowley Sport	7785791.85329999
Custom Frames, Inc.	69898.7679
Electronic Bike Co.	821206.386
Electronic Bike Repair & Supplies	4934520.1926
Expert Bike Co	2827720.3896
Federal Sport	2880853.668
First Rate Bicycles	142515.3636
Fitness Association	122694.6993
GMA Ski & Bike	3794598.3411
Green Lake Bike Company	4260403.224
Hill Bicycle Center	65734.2297
Hill's Bicycle Service	61443.1569
Holiday Skate & Cycle	13218.6348
Hybrid Bicycle Center	49633.8183
Illinois Trek & Clothing	15312.1995
Indiana Bicycle Center	53361

Figure 5-20. Total purchase price list for each vendor

Regarding this table, we will list the total amount that we have spent to import goods from our suppliers in 2014. This table is mainly just a list to give an overview of the capital that the company has spent.

CHAPTER 6. VISUALIZATION AND FORECASTING OR PREDICTIVE MODEL

6.1. Report and dashboard systems (structure)

After having a data warehouse on purchasing segmentation, our group proposed tools to make statistics, compare and visualize data such as Power BI, Excel tools. Because these tools are quite easy to access and in the process of learning, we have also been taught about these tools.
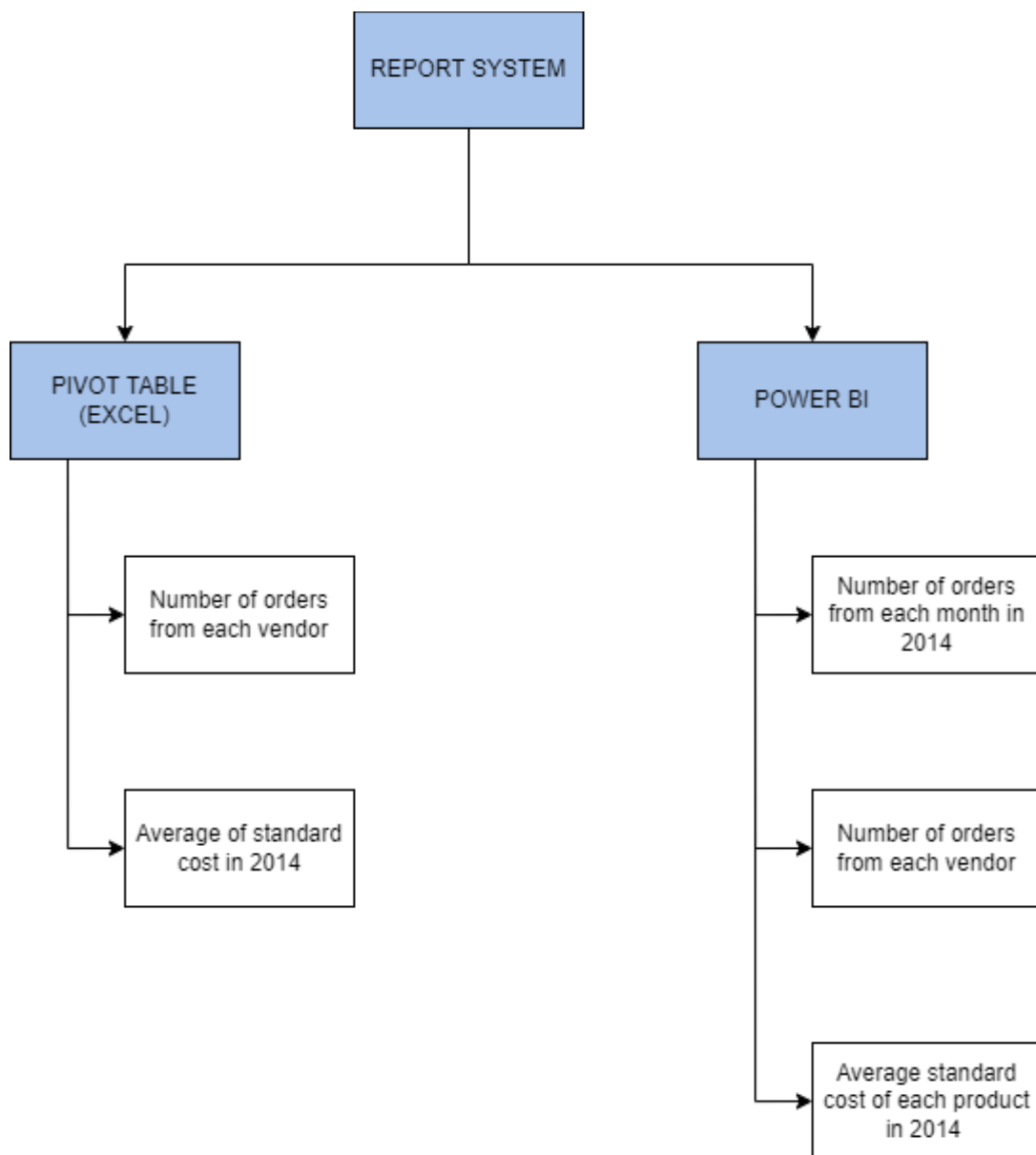
Figure 6-1. Report and dashboard systems

6.2. Data analysis with Power BI

6.2.1. Received quantity entered by month in 2014.
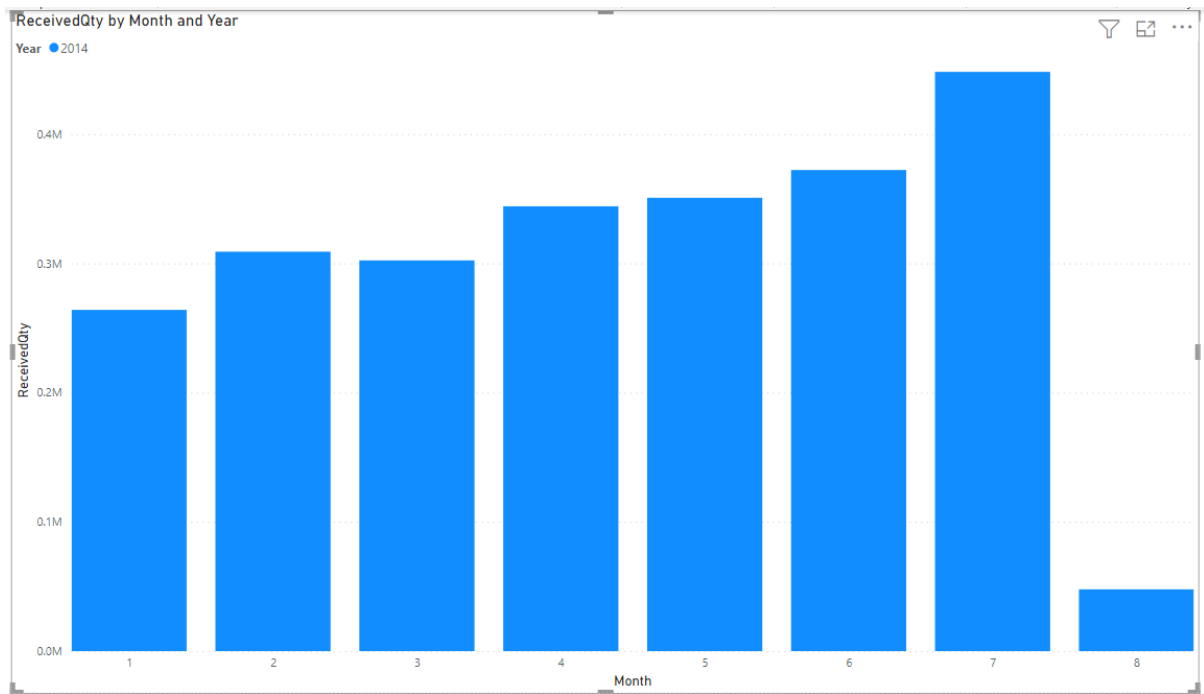
Figure 6-2. Received quantity entered by month in 2014

This is a list of the number of orders imported by month in 2014, we can see that in 2014, the company only imported orders from January to August, and the number of orders was high. Especially in July, the lowest number of orders entered is in August.

6.2.2. Number of orders received from each vendor in 2014.

Figure 6-3. Number of orders received from each vendor in 2014

The table above shows us the number of orders imported from each different supplier in 2014, from which we can somewhat evaluate the main supplier.

6.2.3. Average estimated cost of each product in 2014
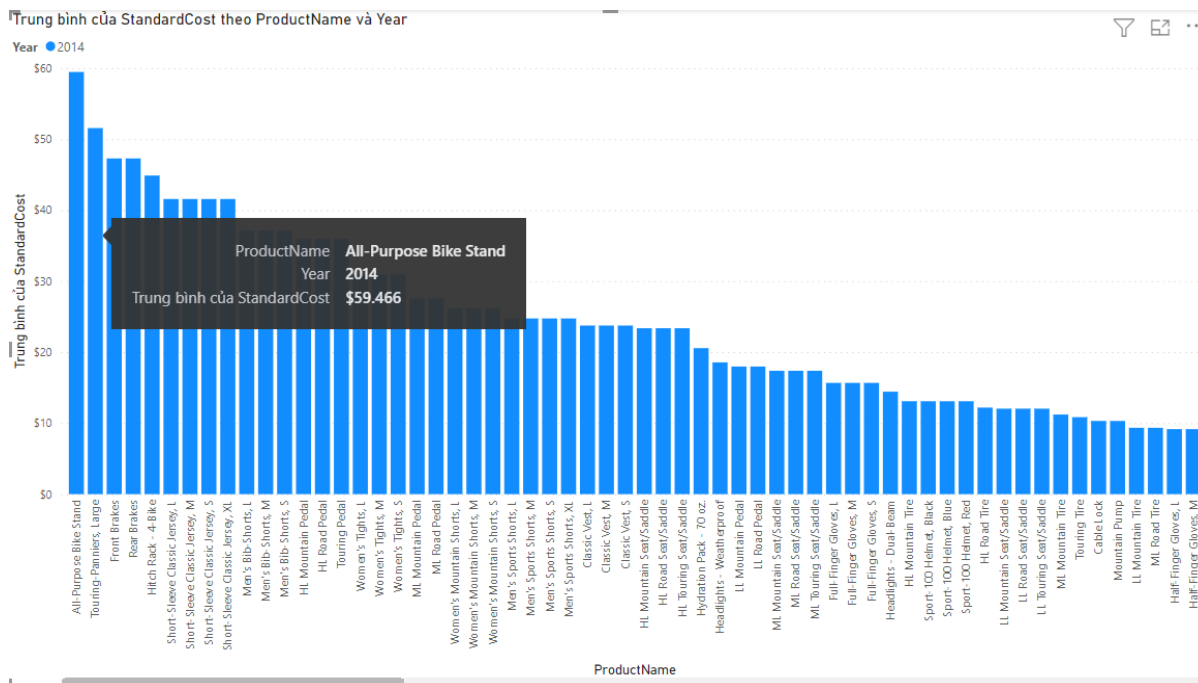
Figure 6-4. Average estimated cost of each product in 2014

We calculate the average price of each product in 2014, to assess which products will account for the high cost, which products will account for the low cost, thereby making strategic assessments of pricing for the company.
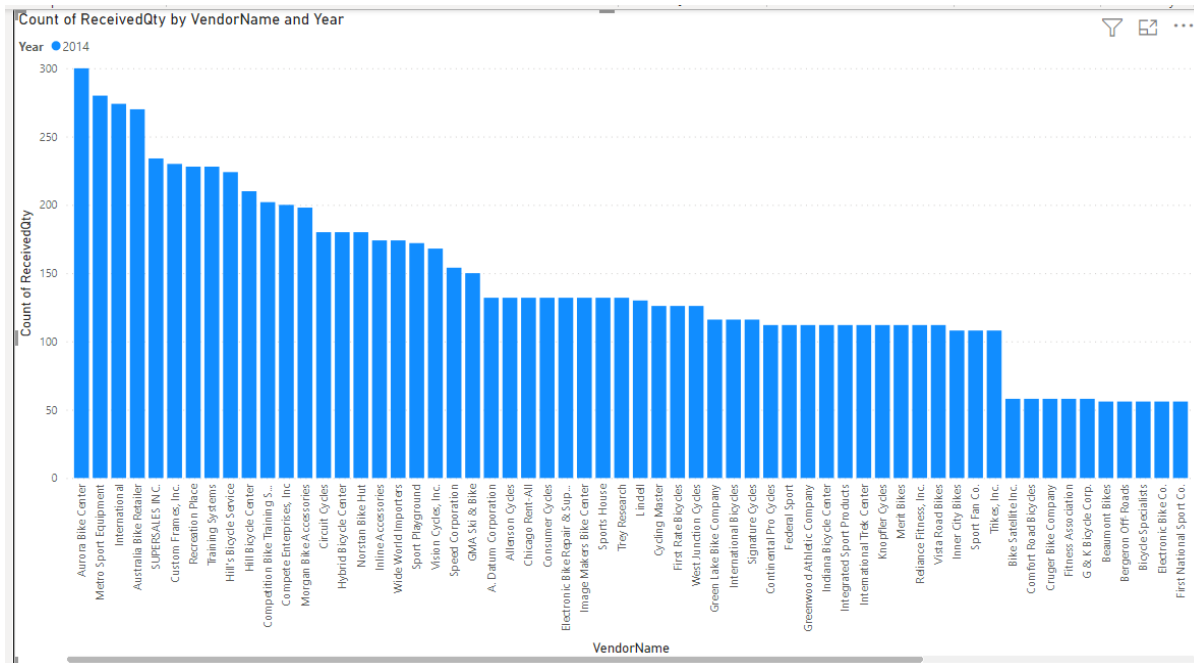
6.2.4. Compare between the safety stock level and the quantity of inventory in 2014.
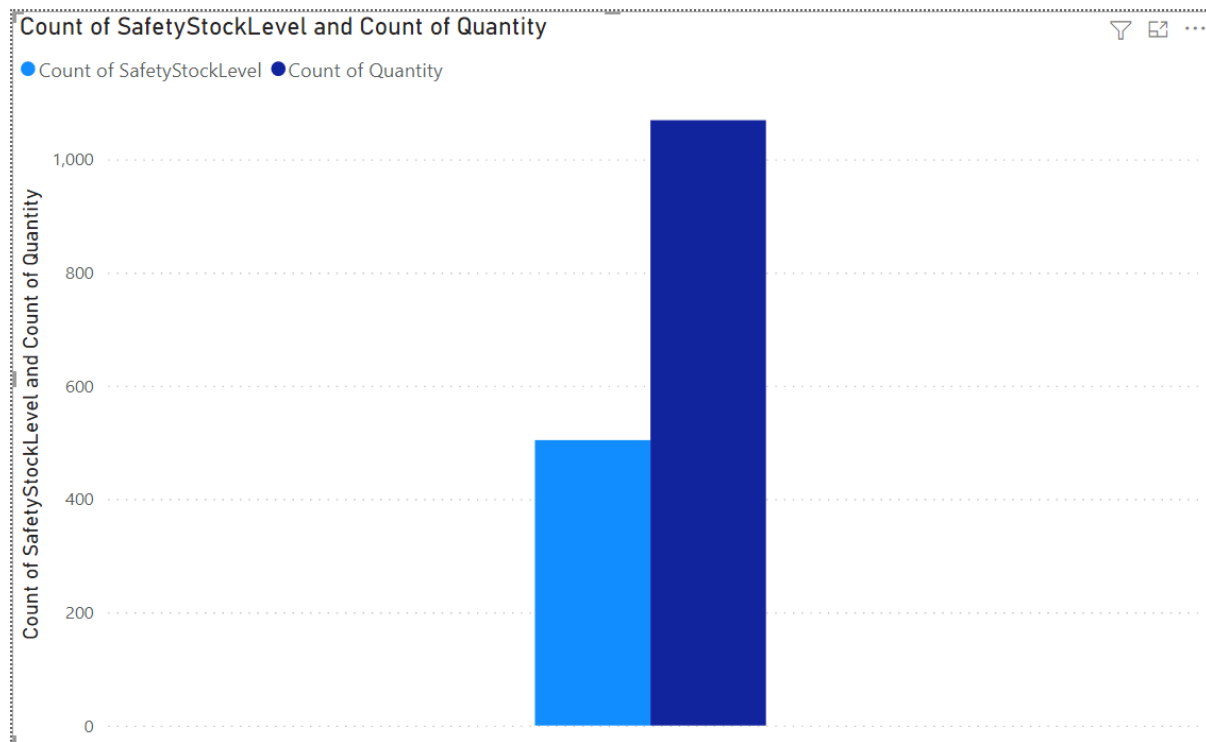
Figure 6-5. Compare between the safety stock level and the quantity of inventory in 2014

We calculate the average price of each product in 2014, to assess which products will account for the high cost, which products will account for the low cost, thereby making strategic assessments of pricing for the company.

6.2.5. Number of products that each employee entered in 2014.
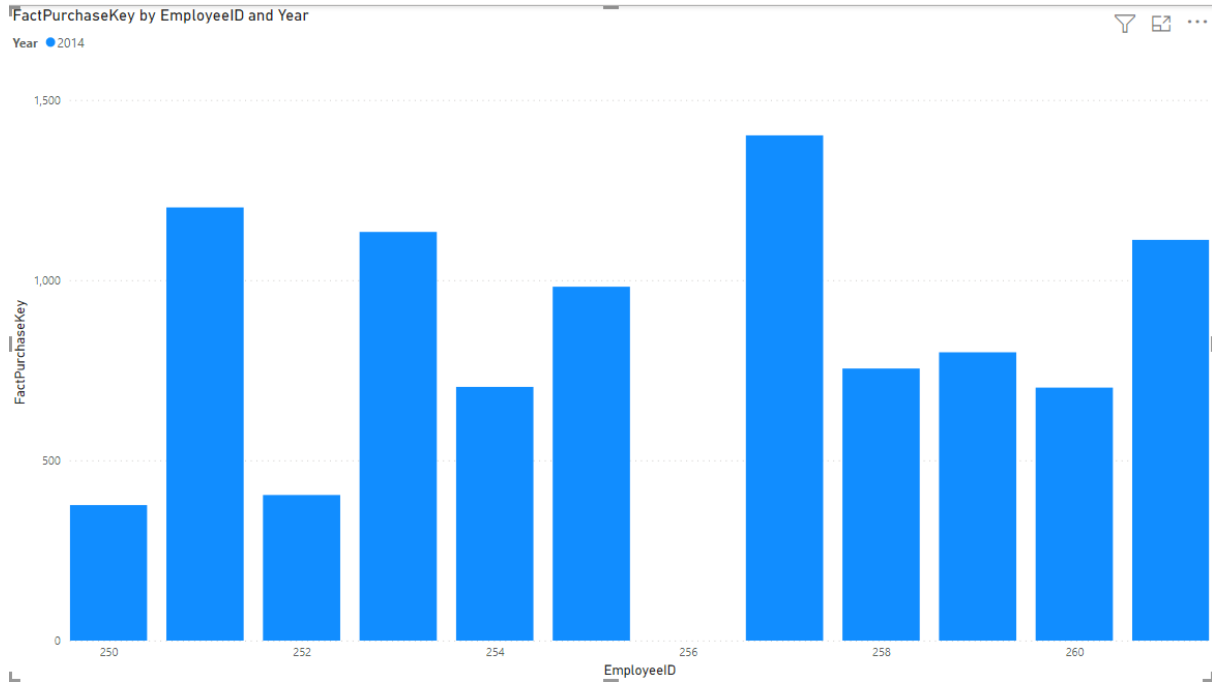
Figure 6-6. Number of products that each employee entered in 2014

We compare the quality of work among employees and know the number of goods they have imported in 2014 of each employee, thereby giving an assessment of their work quality, which employees High-productivity employees, low-productivity employees.

6.2.6. Compare standard price and last receive costs of suppliers
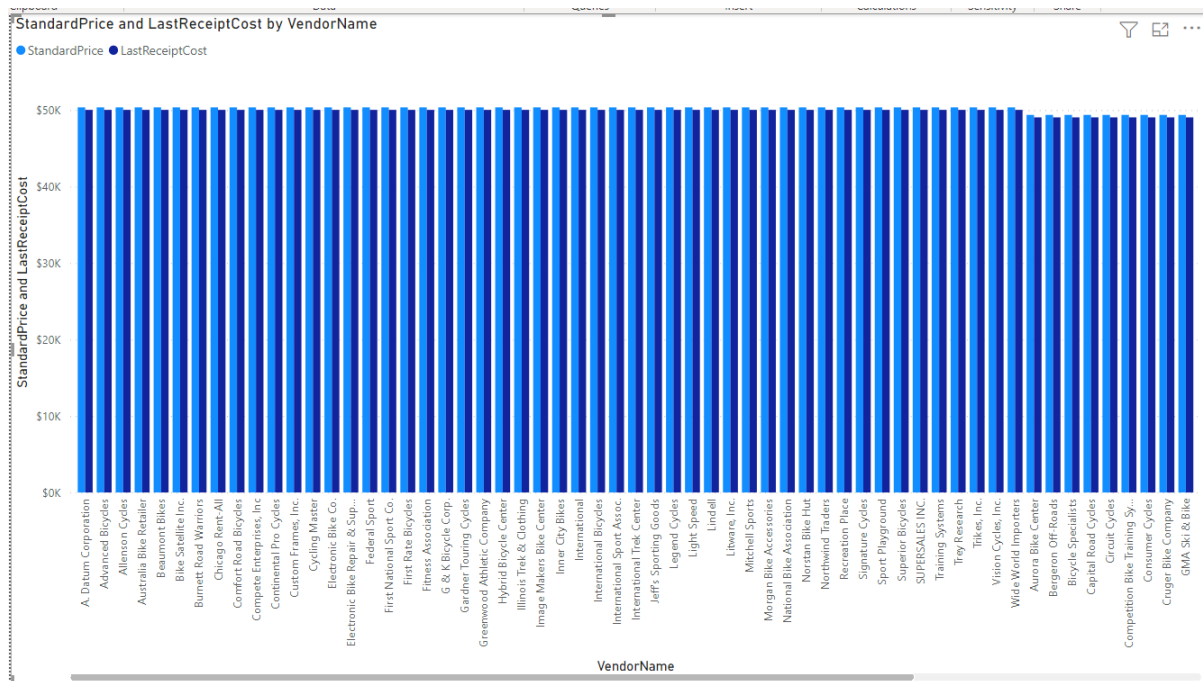
Figure 6-7. Compare standard price and last receive costs of suppliers

In terms of price, capital, this is also a very important thing for any company, so we made a comparison between the estimated price and the price paid for the products we import. From there, seeing the difference between the expected price and the actual payment, this will more or less help the company take the right steps to cut costs if necessary.

6.3. Data analysis with the Pivot Table tool in Microsoft Excel

6.3.1. Number of orders received from each vendor in 2014.
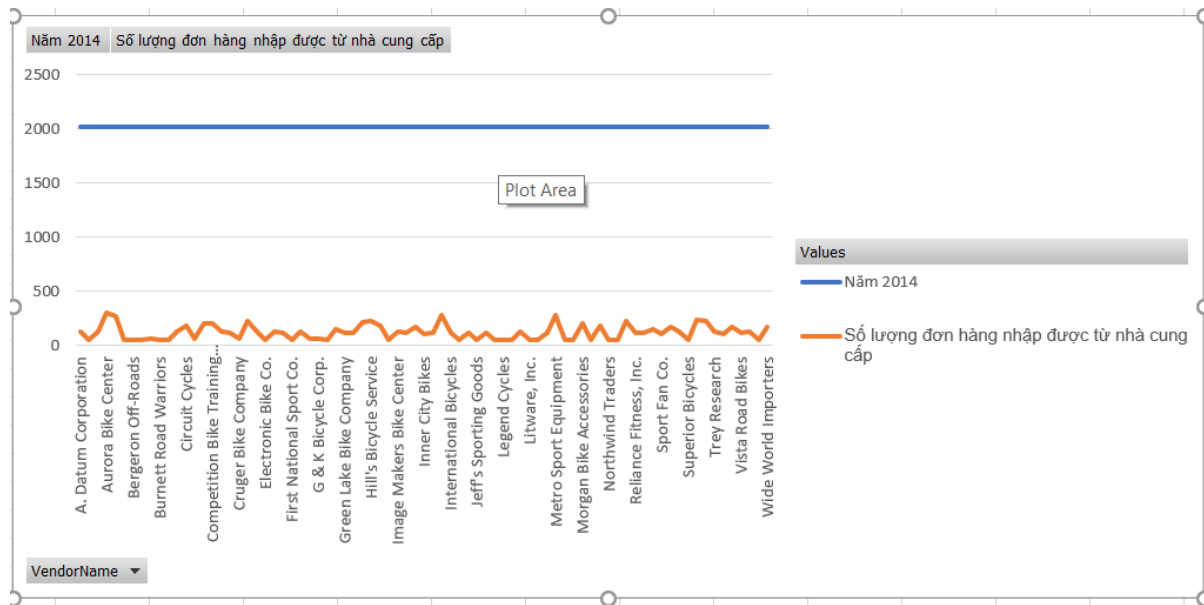
Figure 6-8. Number of orders received from each vendor in 2014

The above table shows the same as table 6.2.2 but we use the Pivot Table tool. We hope that from one data source, we can analyze different tools, gain broad perspectives and more overview, thereby contributing to making better decisions in the future.

6.3.2. Average estimated cost of each product in 2014
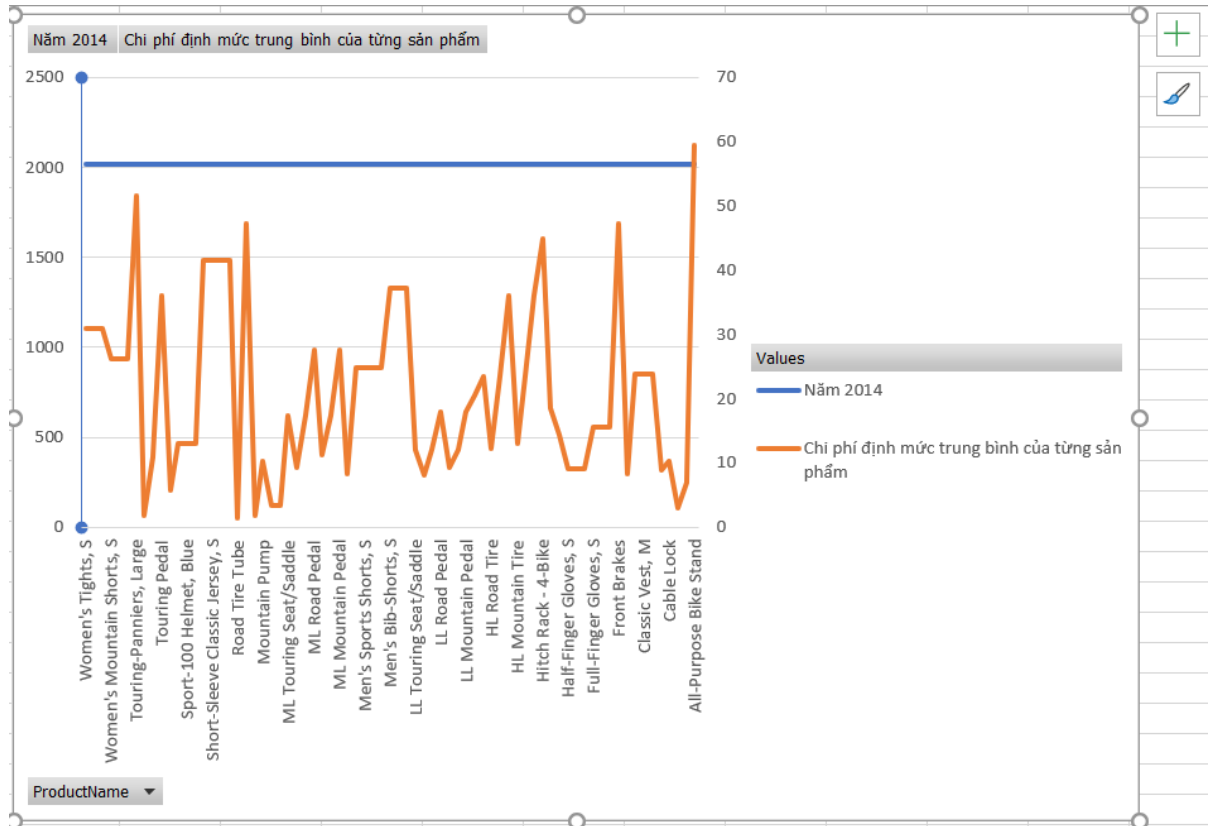
Figure 6-9. Average estimated cost of each product in 2014

This is a table of the estimated cost of each product according to the line chart, we can more or less see the differences of each different product. Above are 25 different products with different average estimated costs in 2014. We can easily see that the "All-Purpose Bike Stand" product has an estimated average price of the highest near $60.

6.3.3. Total shipping cost of each vendor in 2014.
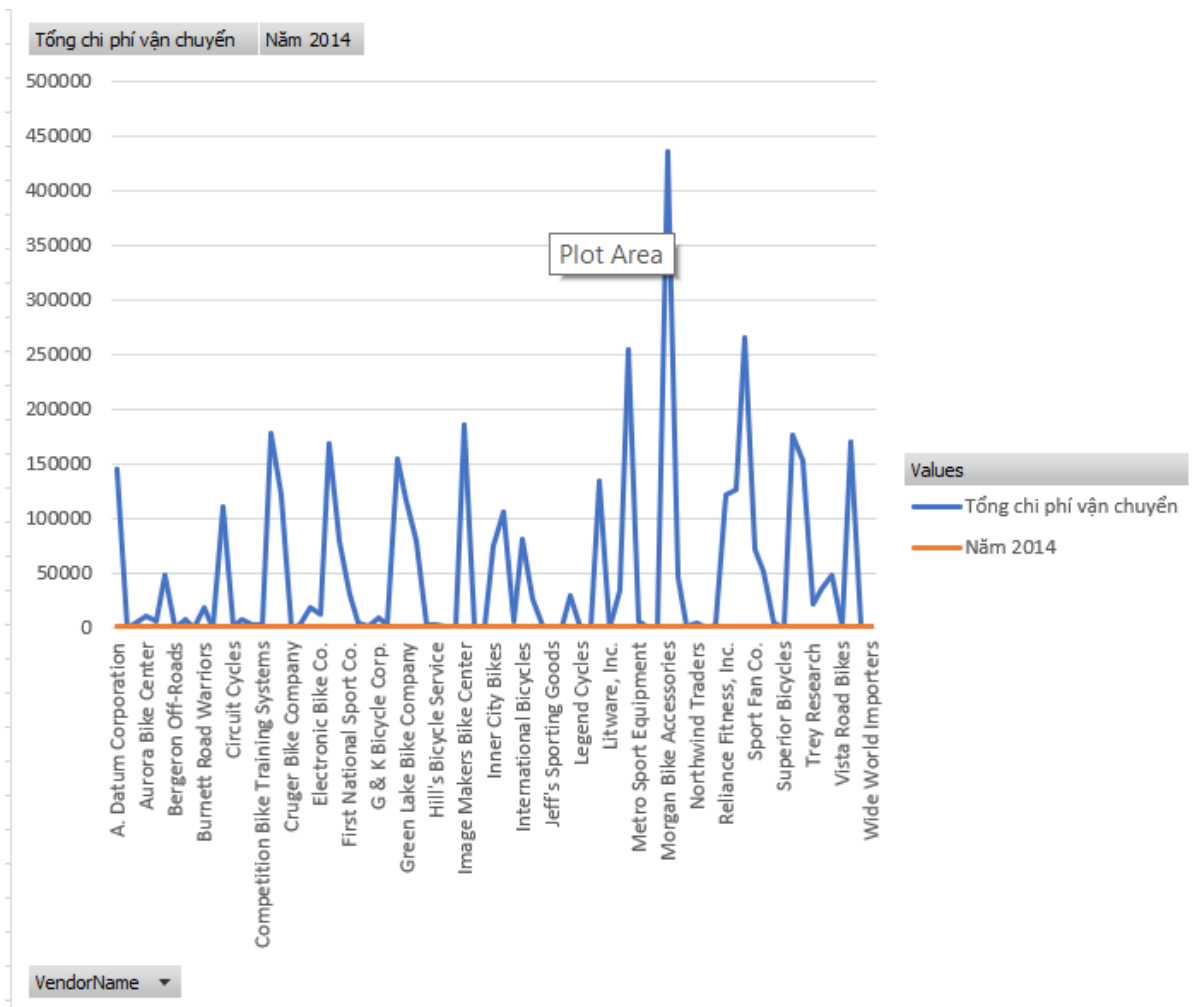
Figure 6-10. Total shipping cost of each vendor in 2014

Shipping cost is also an important factor in the purchasing module, so it also needs to be listed clearly, from which to judge which supplier has good shipping cost, this also contributes to the selection and evaluation of suppliers' quality and saves a considerable amount of money for the company.

6.4.   Evaluation and Discussion

Excel has always been a very useful tool, with a layout of simple, easy-to-understand charts. Basic requirements can be met. However, this tool has not yet produced detailed and diverse statistics. Therefore, our team has used this tool to visualize the not too complicated parts of Adventure Works, but the above tables have also partly helped readers understand more about the general situation of the company in purchasing modules.

About the Power BI tool, this is a tool that specializes in data visualization through charts. It is highly appreciated by users and is used a lot in large and small businesses.

Data on the Purchasing module is included for both basic and advanced visualization. Compared to Excel's Pivot table, Power BI is superior, it is there to help businesses evaluate more accurately through visualized data.

CHAPTER 7. CONCLUSION AND FUTURE WORKS

7.1. Results

Completing the basic requirements set forth by the course in the process of building a data warehouse on purchasing for AWC company.

Use tools to visualize fully and effectively mined data, making the presentation of information clearer and easier to understand.

Proposing BI solutions to help support business decision making. Solutions to support decision-making include three main elements: data sources, data warehouses and data mining.

Can use data storage and analysis tools such as Power BI, SQL server, SSAS, SSIS ... to support the selection of tools for future study and research.

Research is reported KPIs - an indicator to evaluate the effectiveness and analysis of business activities of enterprises. This is an element of data analysis and business decision making.

7.2. Limitations

There are tools and software that our group has just had access to, both learning and making, so time and expertise are limited.

Struggling and spending a lot of time rebuilding the data warehouse.

Since the amount of work is quite large in a short period of time, perhaps our team will have some errors in the research process.

7.3. Future works

After successfully applied to the Purchase array, it can be extended and applied to all other modules of the Adventure Works database-based business system. In addition, we will also be able to apply business intelligence solutions to other companies based on different databases.

Our team will try to apply more than other analysis tools, which will save more time and also help the team access more technology so that they can have a good foundation later, when traveling. work on real projects.

Process and work, more intervention, deeper into the study of source data and ETL for more accurate and reliable data quality.

Continue to complete and expand the project

REFERENCES

https://creately.com/diagram/example/jo83qjn5e/procurement-process-flowchart.

https://planergy.com/blog/purchasing-process/.

https://azure-microsoft-com.translate.goog/es-es/blog/first-azure-as/?_x_tr_sl=en&_x_tr_tl=vi&_x_tr_hl=vi&_x_tr_pto=sc.

http://bis.net.vn/forums/t/560.aspx.

https://www.techtarget.com/searchdatamanagement/definition/OLAP.

https://www.zoho.com/expense/articles/8-essential-kpis-for-the-purchasing-department.htm.

https://smallbusiness.chron.com/roles-purchasing-department-2270.html.

https://www.solobis.net/limitations-of-key-performance-indicators/

https://www.guru99.com/star-snowflake-data-warehousing.html#:~:text=Comparing%20Snowflake%20 vs%20Star%20 schema,table%20and%20any%20 dimension%20 tables.

https://businessimpactinc.com/blog/the-pros-cons-of-data-warehouses/

https://www.guru99.com/data-warehousing.html#:~:text=DWH%20(Data%20warehouse)%20is%20needed,technology%20to%20access%20the%20data

https://panoply.io/data-warehouse-guide/data-mart-vs-data-warehouse/

https://www.cleo.com/blog/knowledge-base-etl-integration#:text=The%205%20steps%20of%20the,transform%2C%20 load%2C%20and%20analyze.

https://www.matillion.com/resources/blog/the-importance-of-etl-tools-in-data-warehousing

https://www.ibm.com/cloud/learn/etl

https://www.tableau.com/learn/articles/business-intelligence/successful-strategy

https://ierp.vn/loi-ich-business-intelligence-tableau-la-gi/#:~:text=H%E1%BB%87%20th%E1%BB%91ng%20BI%20k%E1%BA%BFt%20n%E1%BB%91i,s%E1%BA%A3n%20c%E1%BB%A7a%20doanh%20nghi%E1%BB%87p%20m%C3%ACnh