



Project Report

Datawarehouse

Muhammad Uzair

MSDS - 29414

Project Report: Inventory Management System for Wide World Importers

Project Statement

The objective of this project is to design and implement a star schema-based data warehouse for Wide World Importers to enhance their inventory management and reporting capabilities.

Business Domain

Retail and Wholesale Distribution

Describe Overall Business Domain

Wide World Importers operates as a retailer and wholesaler, dealing with various products supplied by multiple suppliers. The company handles a diverse range of products, each categorized by attributes such as color, packaging, and brand. Effective inventory management, supplier performance monitoring, and transaction tracking are critical for the company's operations and decision-making.

Describe the Particular Business Process

The business process involves managing inventory transactions, including stock receipt from suppliers, stock allocation, and inventory level tracking. The process also includes monitoring supplier performance and ensuring optimal stock levels.

Major Business Entities

1. **StockItem:** Represents individual items in the inventory, with attributes like product name, brand, size, and barcode.
2. **Supplier:** Information about suppliers providing products, including supplier name, contact details, and delivery addresses.
3. **Color:** Different colors associated with products.
4. **PackageType:** Types of packaging used for products.
5. **TransactionType:** Types of inventory transactions, such as purchases, sales, and returns.

Key Performance Indicators (KPIs)

1. **Total Stock Quantity:** Total quantity of stock items available in the inventory.
2. **Reorder Level Compliance:** Measure of how often stock levels are maintained above reorder levels.
3. **Supplier Delivery Time:** Average time taken by suppliers to deliver products.
4. **Stock Turnover Rate:** Frequency at which inventory is sold and replaced over a period.
5. **Inventory Accuracy:** Accuracy of inventory records compared to physical counts.

Potential Analytical Requirements of the Business

1. **Inventory Level Analysis:** Monitoring current stock levels and forecasting future inventory needs.
2. **Supplier Performance Analysis:** Evaluating supplier reliability and delivery performance.
3. **Sales and Transaction Analysis:** Analyzing sales trends and transaction types to identify popular products and seasonal variations.
4. **Cost and Pricing Analysis:** Analyzing cost structures and pricing strategies to optimize profitability.
5. **Product Performance Analysis:** Evaluating product performance based on various attributes such as color, size, and packaging type.

Grain

The grain of the FactInventoryTransaction table is at the level of individual inventory transactions, each uniquely identified by InventoryTransactionID.

Summary of the Inventory Management Database Schema

The Inventory Management database schema for Wide World Importers is designed using a star schema approach to optimize for reporting and analysis. The schema comprises several dimension tables capturing detailed attributes of key business entities and a fact table recording inventory transactions.

Dimensional Tables

1. DimDate

- **Attributes:** DateKey (PK), Date, Year, Quarter, Month, Day
- **Description:** Contains date-related attributes for each day.

2. DimProduct

- **Attributes:** ProductID (PK), ProductName, SupplierID, ColorID, UnitPackageID, OuterPackageID, Brand, Size, LeadTimeDays, IsChillerStock, Barcode, TaxRate, UnitPrice, RetailPrice, WeightPerUnit
- **Description:** Details of individual stock items.

3. DimSupplier

- **Attributes:** SupplierID (PK), SupplierName, SupplierReference, PhoneNumber, Website, DeliveryAddress, PostalCode
- **Description:** Information about suppliers.

4. DimColor

- **Attributes:** ColorID (PK), ColorName
- **Description:** Details of product colors.

5. DimPackageType

- **Attributes:** PackageTypeID (PK), PackageTypeName
- **Description:** Types of product packaging.

6. DimTransactionType

- **Attributes:** TransactionTypeID (PK), TransactionTypeName
- **Description:** Types of inventory transactions.

Fact Table

1. FactInventoryTransaction

- **Attributes:** InventoryTransactionID (PK), DateKey (FK), TransactionTypeID (FK), SupplierID (FK), Quantity, ProductID (FK), PackageTypeID (FK), ColorID (FK)
- **Description:** Records individual inventory transactions.

Data Warehouse Development and ETL Process

Overview

The data warehouse development and ETL (Extract, Transform, Load) process for Wide World Importers was meticulously designed to ensure efficient data management, from extraction to final reporting. This process leverages Microsoft SQL Server Integration Services (SSIS) in Visual Studio to streamline the data pipeline and support robust, up-to-date business intelligence reporting via Power BI.

ETL Process Workflow

1. Data Extraction

- Data is extracted from the Wide World Importers operational database into a staging database. This initial step captures raw transactional data, including details of inventory transactions, product information, supplier details, and other relevant data.

2. Data Loading into Staging Database

- The extracted data is loaded into dedicated tables within the staging database. This step creates a clear separation between raw data and subsequent data processing, facilitating easier data manipulation and error handling.

3. Data Cleaning and Transformation

- Data cleaning and transformation processes are performed using Python scripts. Key tasks include:
 - Handling null values.
 - Correcting data entry errors.
 - Standardizing data formats.
 - Ensuring data consistency and integrity.
- This ensures that the data is accurate and ready for further processing.

4. Star Schema Transformation

- The cleaned data is transformed into a star schema, a data modeling technique optimized for data warehousing and BI reporting. The star schema includes:
 - **Fact Tables:** Central tables containing quantitative data for analysis (e.g., FactInventoryTransaction).
 - **Dimension Tables:** Surrounding tables containing descriptive attributes related to the fact data (e.g., DimProduct, DimSupplier).

5. Loading into Star Schema Database

- The transformed data is loaded into the star schema database. This step ensures that the data is structured according to the star schema design, enhancing query performance and ease of use in reporting.

6. Final Data Cleaning and Loading

- Additional data cleaning and loading processes are performed to ensure that the star schema database is up to date. This includes:
 - Checking for new data updates.
 - Updating previously existing data.
- This step ensures the integrity and timeliness of the data warehouse.

7. Merging and Updating Data Warehouse

- The data from the star schema database is merged and updated into the final data warehouse database. This consolidation step ensures that all historical and current data is accurately reflected.

8. View Snapshot Creation

- A view snapshot is created to provide a consolidated view of the data. The view combines fact and dimension tables to present a comprehensive dataset for analysis. The SQL view includes attributes such as product details, transaction information, supplier data, and more.

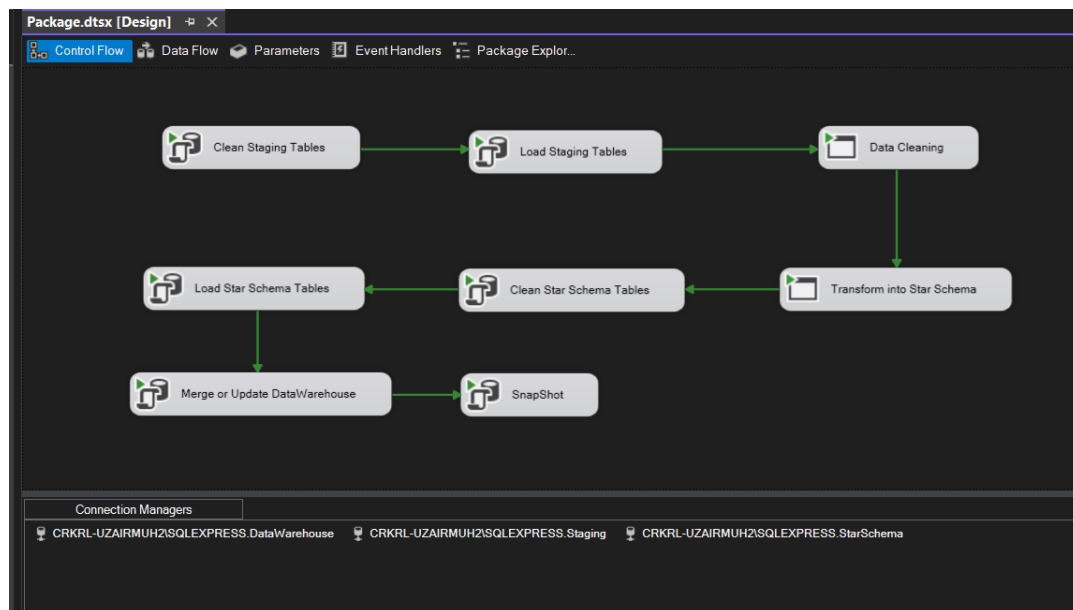
9. Loading into Power BI

- The view snapshot is loaded into Power BI for creating interactive dashboards and visualizations. This enables business users to gain insights into inventory levels, supplier performance, transaction trends, and more.

Final Remarks

The entire ETL process, from data extraction to loading into Power BI, is orchestrated using SSIS in Visual Studio. This pipeline ensures data accuracy, consistency, and timely availability for business intelligence reporting. The comprehensive approach facilitates detailed monitoring and management of inventory, enhancing decision-making processes at Wide World Importers.

This well-structured ETL process not only improves data quality but also supports the scalability and flexibility needed to adapt to future business requirements and data growth.



Thank-you.