Assignment Title:

Lab sheet 2

Course:

CCS3307 Data Warehousing

Submission Date:

2025/07/01

Index Number:

CIT-23-02-0094

Name :

M.B.Thrithwaka Preethi Shakya

Task 1: Identify Key Components

1. What are 3 possible data sources for this organization?

- **Sales System** – Contains daily sales transactions, invoices, and revenue data.
- **Inventory System** – Tracks stock levels, product movements, and restocking information.
- **Customer Support Database** – Includes customer complaints, service tickets, and feedback logs.

2. What tools or processes are needed for ETL (Extract, Transform, Load)?

- **ETL Tools**:
  - **Apache Nifi** (open-source) or **Talend** – for managing the ETL pipeline.
  - **SQL scripts** – for data cleaning and transformation.
- **ETL Process**:
  - **Extract**: Pull data from source systems (Sales, Inventory, Support).
  - **Transform**: Clean, filter, and normalize data (format dates, remove duplicates).
  - **Load**: Store the cleaned data into the staging area, then into the data warehouse.

3. What type of data storage will be used in the warehouse?

- A **centralized relational database** such as **Amazon Redshift**, **Google BigQuery**, or **Microsoft SQL Server Data Warehouse**.
- Designed using **star or snowflake schema** for performance and simplicity.

4. Who are the potential users of this data warehouse?

- **Business Analysts** – analyze trends and performance metrics.
- **Marketing Team** – identify customer preferences and campaign performance.
- **Inventory Managers** – optimize stock levels and order forecasting.
- **Executives** – monitor overall business performance and KPIs.

5. What reporting or BI tools could be useful?

- **Power BI** – for interactive dashboards and reports.
- **Tableau** – for advanced data visualization.
- **Google Data Studio** – for quick and easy reporting (free option).
- These tools help convert data into charts, graphs, and KPI dashboards.

Task 2: Draw the Architecture Diagram