

Gold Layer – Data Catalog

Overview

This document describes the final **Gold layer** of the data warehouse, modeled using a **star schema** optimized for analytical queries and BI reporting.

The Gold layer exposes business-ready views built on top of the cleansed and standardized Silver layer.

1) gold.fact_sales

Purpose

Stores transactional sales data at the order–product level, serving as the central fact table for analytical reporting.

Column Name	Data Type	Description
order_number	NVARCHAR(50)	Unique identifier of the sales order.
customer_key	INT	Surrogate key referencing the customer in <code>dim_customer</code> .
product_key	INT	Surrogate key referencing the product in <code>dim_product</code> .
customer_id	INT	Business identifier of the customer from the source system.
product_number	NVARCHAR(50)	Business identifier of the product from the source system.
order_date	DATE	Date when the order was placed.
ship_date	DATE	Date when the order was shipped.
due_date	DATE	Date when the order was due.
sales_amount	DECIMAL(10,2)	Total sales value for the line item (price \times quantity).
quantity	INT	Number of units sold.
unit_price	DECIMAL(10,2)	Price per unit at the time of sale.

2) gold.dim_customer

Purpose

Stores consolidated and cleansed customer master data, enriched with demographic and location attributes.

Column Name	Data Type	Description
customer_key	INT	Surrogate key uniquely identifying the customer.
customer_id	INT	Business identifier of the customer from the CRM system.
customer_number	NVARCHAR(50)	Alternative customer identifier used across source systems.
first_name	NVARCHAR(100)	Customer's first name.
last_name	NVARCHAR(100)	Customer's last name.
gender	NVARCHAR(10)	Standardized gender of the customer (e.g., Male, Female, Unknown).
country	NVARCHAR(100)	Country where the customer is located (e.g., United States).
birthdate	DATE	Customer's date of birth.
marital_status	NVARCHAR(10)	Marital status of the customer (e.g., Married, Single).
create_date	DATE	Date when the customer record was created in the source system.

3) gold.dim_product

Purpose

Stores the current version of each product, enriched with category and maintenance attributes, supporting sales analysis by product.

Column Name	Data Type	Description
product_key	INT	Surrogate key uniquely identifying the product.
product_id	NVARCHAR(50)	Business identifier of the product from the CRM system.
product_number	NVARCHAR(50)	Product key used to relate sales transactions.
category_id	NVARCHAR(50)	Identifier of the product category.
product_name	NVARCHAR(250)	Name of the product.
category	NVARCHAR(100)	High-level product category.
subcategory	NVARCHAR(100)	Product subcategory.
product_line	NVARCHAR(50)	Product line classification.
cost	DECIMAL(10,2)	Cost of the product.
maintenance	NVARCHAR(10)	Indicates whether the product requires maintenance (Yes / No).
start_date	DATE	Start date of the current product version.

Notes

- The Gold layer follows a **star schema** design to support efficient analytical queries.
- Surrogate keys are used to decouple analytics from source system identifiers.
- Only **current product versions** are exposed in `dim_product` to ensure clean joins with `fact_sales`.
- All views are built exclusively from the **Silver layer**, which handles cleansing, standardization, normalization, and enrichment.