Data Catalogue – GOLD Layer

This document describes the GOLD layer of the Data Warehouse, which is designed for analytical consumption and reporting purposes. It contains dimension and fact views built on top of the cleansed Silver layer data.

The structure of the GOLD layer follows a star schema. At the center of this schema is the `fact_sales` view, which captures transactional sales data. It is connected to surrounding dimension views: `dim_products` and `dim_customers`, enabling efficient slicing and dicing of sales metrics across various customer and product attributes.

Data Catalogue – GOLD Layer Entities

gold.dim_customers

Column Name	Data Type	Description
customer_key	INT	Surrogate key for each customer (generated using ROW_NUMBER)
customer_id	INT	Unique customer ID from CRM
customer_number	NVARCHAR(50)	Customer key (textual ID) from CRM
first_name	NVARCHAR(50)	Customer's first name
last_name	NVARCHAR(50)	Customer's last name
country	NVARCHAR(50)	Country from ERP location table
marital_status	NVARCHAR(50)	Marital status, standardized ('Single', 'Married', etc.)
gender	NVARCHAR(50)	Gender resolved from CRM or ERP ('Male', 'Female', or 'n/a')
birthdate	DATE	Date of birth from ERP
create_date	DATE	Date the customer record was created

gold.dim_products

Column Name	Data Type	Description
product_key	INT	Surrogate key for each product (generated using ROW_NUMBER)
product_id	INT	Product ID from CRM
product_number	NVARCHAR(50)	Extracted product key
product_name	NVARCHAR(50)	Name of the product
category_id	NVARCHAR(50)	Category ID derived from product key
category	NVARCHAR(50)	Product category from ERP
subcategory	NVARCHAR(50)	Product subcategory from ERP
maintenance	NVARCHAR(50)	Maintenance type/category
cost	INT	Product cost
product_line	NVARCHAR(50)	Product line (e.g., "Mountain", "Road")
start_date	DATE	Start date of product availability

gold.fact_sales

Column Name	Data Type	Description
order_number	NVARCHAR(50)	Unique sales order number
product_key	INT	Foreign key from dim_products
customer_key	INT	Foreign key from dim_customers
order_date	DATE	Date of order placement
shipping_date	DATE	Date of product shipment

due_date	DATE	Due date for the order
sales_amount	INT	Total sales amount (validated and derived if needed)
quantity	INT	Number of units sold
price	INT	Price per unit