

Data Catalog

Overview

The Gold layer contains business-level representations of data, designed to support business-level analytics. It aggregates and reconciles data from the Silver layer, providing unified customer and product profiles along with sales fact data. These views are optimized for reporting, dashboards, and analytical queries, leveraging surrogate keys for dimensional modeling and ensuring data consistency through predefined business rules.

1. gold.dim_customers

Column Name	Data Type	Description	Example Value
customer_number	INTEGER	A unique surrogate key assigned sequentially based on cst_id, may shift on data changes.	1
customer_id	VARCHAR	Unique identifier for the customer from CRM data.	11000
customer_key	VARCHAR	Alternate key for the customer, used for cross-system matching.	AW00011000
first_name	VARCHAR	Customer's first name from CRM.	John
last_name	VARCHAR	Customer's last name from CRM.	Doe
country	VARCHAR	Customer's country from ERP location data, defaults to 'Unknown' if missing.	Australia
gender	VARCHAR	Customer's gender, prefers CRM value unless 'Unknown', then uses ERP gen if available, defaults to 'Unknown'.	Male
marital_status	VARCHAR	Customer's marital status from CRM.	Married / Single
birth_date	DATE	Customer's birth date from ERP, falls back to create_date if missing.	05-15-1990
create_date	DATE	Date when the customer record was created in CRM.	01-10-2023

2. gold.dim_products

Column Name	Data Type	Description	Example Value
product_number	INTEGER	A unique surrogate key assigned sequentially based on prd_start_dt and prd_id, may shift on data changes.	1
product_id	VARCHAR	Unique identifier for the product from CRM data.	210
product_key	VARCHAR	Alternate key for the product, used for cross-system matching.	FR-R92B-58
product_name	VARCHAR	Name of the product from CRM.	HL ROAD FRAME - BLACK- 58
category_id	VARCHAR	Identifier for the product category from CRM.	CO_RF
category	VARCHAR	Main category name from ERP, defaults to 'Unknown' if missing.	Components
sub_category	VARCHAR	Sub-category name from ERP, defaults to 'Unknown' if missing.	Road Frames
maintenance	VARCHAR	Maintenance status from ERP, defaults to 'N/A' if missing.	Yes
cost	DECIMAL	Cost of the product from CRM.	12
product_line	VARCHAR	Product line classification from CRM.	Road
start_date	DATE	Date when the product became available.	01-10-2023
end_date	DATE	Date when the product was discontinued, NULL for active products.	01-10-2023

3. gold.fact_sales

Column Name	Data Type	Description	Example Value
order_number	VARCHAR	Unique identifier for the sales order from CRM.	SO43697
product_number	INTEGER	Surrogate key linking to gold.dim_products, NULL if no match.	1
customer_number	INTEGER	Surrogate key linking to gold.dim_customers, NULL if no match.	1
order_date	DATE	Date when the sales order was placed.	01-10-2023
shipping_date	DATE	Date when the sales order was shipped.	01-10-2023
due_date	DATE	Date when the sales order is due.	01-10-2023
sales_amount	DECIMAL	Total sales amount for the order from CRM.	1299.99
quantity	INTEGER	Number of units sold in the order.	2
price	DECIMAL	Unit price of the product in the order.	649.99