Data Catalog: Gold Layer

The Gold Layer represents the business-ready data in the Data Warehouse. It is the final curated layer designed for analytical and reporting use cases. This layer contains dimension tables for descriptive attributes and fact tables for measurable events.

Table: gold.dim_customers

Purpose: Stores enriched customer data combining demographic and geographic details for reporting and customer segmentation.

Column Name	Data Type	Description
customer_key	INT	Surrogate key uniquely identifying each customer record in the dimension table.
customer_id	INT	Original unique identifier for each customer from the source system.
customer_number	NVARCHAR(50)	Alphanumeric identifier used to reference the customer in transactions.
first_name	NVARCHAR(50)	Customer's given name.
last_name	NVARCHAR(50)	Customer's family or surname.
country	NVARCHAR(50)	Customer's country of residence (e.g., 'Australia').
marital_status	NVARCHAR(50)	Marital status (e.g., 'Single', 'Married').
gender	NVARCHAR(50)	Gender identity of the customer ('Male', 'Female', 'n/a').
birthdate	DATE	Date of birth in format YYYY-MM-DD (e.g., 1971-10-06).
create_date	DATE	Date when the customer record was created in the source system.

Table: gold.dim_products

Purpose: Holds detailed product information including hierarchical categorization, pricing, and availability, enabling product-level analysis.

Column Name	Data Type	Description
product_key	INT	Surrogate key uniquely identifying each product record.
product_id	INT	Unique source identifier for the product.
product_number	NVARCHAR(50)	Structured alphanumeric code representing the product.
product_name	NVARCHAR(50)	Descriptive name of the product (e.g., model, color, size).
category_id	NVARCHAR(50)	Code linking to a high-level product category.
category	NVARCHAR(50)	Broad classification of the product (e.g., 'Bikes', 'Components').
subcategory	NVARCHAR(50)	Specific type of product within the category (e.g., Mountain Bikes).
maintenance_required	NVARCHAR(50)	Whether the product requires maintenance ('Yes' / 'No').
cost	INT	Base cost or price of the product.
product_line	NVARCHAR(50)	The series or line the product belongs to (e.g., 'Road', 'Mountain').
start_date	DATE	Date the product became available for sale.

Table: gold.fact_sales

Purpose: Captures individual sales transactions and revenue metrics, linked to dimension tables for customer and product analytics.

Column Name	Data Type	Description
order_number	NVARCHAR(50)	Unique identifier for the sales order (e.g., "SO54496").
product_key	INT	Foreign key linking to dim_products.product_key.
customer_key	INT	Foreign key linking to dim_customers.customer_key.
order_date	DATE	Date when the order was placed.
shipping_date	DATE	Date when the order was shipped.
due_date	DATE	Date by which the payment was due.
sales_amount	INT	Total revenue for the line item (price × quantity).
quantity	INT	Number of units sold in the transaction.
price	INT	Price per unit of product in the transaction.

Notes

- All foreign key fields (`product_key`, `customer_key`) link to their respective dimension tables to support star schema modeling.
- All dates are expected to be in ISO format (YYYY-MM-DD) for consistency and compatibility.
- Consider adding row-level metadata (e.g., `load_date`, `source_system`) for auditability if needed.