

Data Dictionary – Gold Schema

Overview

The Gold Layer consists of **dimension tables** and **fact tables** for specific business metrics. It integrates and enriches data from the Silver Layer, transforming it into **business-ready datasets** optimized for analytics and reporting.

1.gold.dim_customers

- **Purpose:** Stores enriched customer details with demographic and segmentation attributes.
- **Column:**

Column Name	Data Type	Description
customer_key	INT	Surrogate key uniquely identifying each customer record in the dimension table.
customer_id	VARCHAR(10)	Unique identifier of the customer from the source system.
first_name	VARCHAR(50)	Customer's first name (as recorded in the system).
last_name	VARCHAR(50)	Customer's last name or family name.
age_category	VARCHAR(10)	Age group classification of the customer (e.g., '18-25', '26-35').
gender	VARCHAR(10)	Gender of the customer (e.g., 'Male', 'Female').
country	VARCHAR(50)	Customer's country of residence.
city	VARCHAR(50)	Customer's city of residence.
customer_segment	VARCHAR(50)	Segment classification (e.g., 'Premium', 'Regular').

2.gold.dim_products

- **Purpose:** Provides information about products and their attributes.
- **Column:**

Column Name	Data Type	Description
product_key	INT	Surrogate key uniquely identifying each product record.
product_id	VARCHAR(10)	Unique identifier of the product from the source system.
product_name	VARCHAR(100)	Descriptive name of the product.
category	VARCHAR(10)	High-level classification of the product (e.g., 'Electronics').
Brand	VARCHAR(10)	Brand or manufacturer of the product.
Price	DECIMAL(10,2)	Selling price of the product.

Cost	DECIMAL(10,2)	Base cost of the product.
launch_year	INT	Year when the product was launched.
launch_month	VARCHAR(10)	Month when the product was launched.
current_inventory	INT	Current stock quantity available.
reorder_level	INT	Threshold quantity at which the product should be reordered.

3.gold.fact_orders

- **Purpose:** Stores transactional order data, linking customers and products for analysis.
- **Column:**

Column Name	Data Type	Description
order_id	VARCHAR(20)	Unique identifier of the order.
customer_key	INT	Foreign key referencing gold.dim_customers.
product_key	INT	Foreign key referencing gold.dim_products.
order_year	INT	Year when the order was placed.
order_month	VARCHAR(10)	Month when the order was placed.
order_day	INT	Day when the order was placed.
order_status	VARCHAR(20)	Status of the order (e.g., 'Completed', 'Pending').
quantity	INT	Quantity of the product ordered.
unit_price	DECIMAL(10,2)	Price per unit of the product.
line_total	INT	Total value for the line item (quantity × unit_price).
subtotal	INT	Order subtotal before discounts and taxes.
discount_amount	FLOAT	Discount applied to the order.
shipping_cost	INT	Shipping charges applied.
tax_amount	FLOAT	Tax applied to the order.
payment_method	VARCHAR(20)	Payment method used (e.g., 'Credit Card', 'PayPal').
order_source	VARCHAR(20)	Source of the order (e.g., 'Online', 'In-Store').
billing_country	VARCHAR(20)	Country used for billing address.