

Data Catalog

The Gold layer is the business level data representation, structured to support analytical & reporting use cases. It consists of dimension tables and fact tables for specific business metrics.

1. gold.dim_customers

- Purpose: Stores customer details enriched with demography and geographic data.
- Columns:

Column Name	Data Type	Description
customer_key	INT	Surrogate key uniquely identifying each customer record in the customer dimension table
customer_id	INT	Unique numerical identifier assigned to each customer
customer_number	NVARCHAR(50)	Alphanumeric identifier representing the customer, used for tracking & referencing
first_name	NVARCHAR(50)	The customer's first name
last_name	NVARCHAR(50)	The customer's last name
country	NVARCHAR(50)	The country of residence of the customer (eg: 'Australia')
marital_status	NVARCHAR(50)	The marital status of the customer (eg: 'Married', 'Single')
gender	NVARCHAR(50)	The gender of the customer (eg: 'Male', 'Female', 'n/a')
birthdate	DATE	The date of birth of the customer, formatted as YYYY-MM-DD (eg: 1971-10-06)
create_date	DATE	The date and time when the customer record was created in the system

2. gold.dim_products

- Purpose: Provides information about the products and their attributes.
- Columns:

Column Name	Data Type	Description
product_key	INT	Surrogate key uniquely identifying each product record in the product dimension table
product_id	INT	Unique numerical identifier assigned to each product
product_number	NVARCHAR(50)	Alphanumeric identifier representing the product, used for tracking & referencing
Category_id	NVARCHAR(50)	Descriptive name of the product, including key details such as type, color & size
category	NVARCHAR(50)	The broader classification of the product (eg: Bikes, Components) to group related items
subcategory	NVARCHAR(50)	More detailed classification of the products within category

maintenance	NVARCHAR(50)	Indicates whether the product requires maintenance (eg: 'Yes', 'No')
cost	INT	The cost of the product, measured in the monetary units
Product_line	NVARCHAR(50)	The specific product line or series to which the product belongs (eg: Roads)
Start_date	DATE	The date when the product became available for sale or use, stored in

3. gold.fact_sales

- Purpose: Stores transactional sales data for analytical purposes
- Columns:

Column Name	Data Type	Description
order_number	NVARCHAR(50)	A unique alphanumeric identifier for each sales order
product_key	INT	Surrogate key linking the order to the product dimension table
customer_key	INT	Surrogate key linking the order to the customer dimension table
order_date	DATE	The date when the order was placed
shipping_date	DATE	The date when the order was shipped to the customer
due_date	DATE	The date when the order payment was due
sales_amount	INT	The total monetary value of the sale for the line item, in the whole currency units
quantity	INT	The number of units the products ordered for the line item
price	INT	The price per unit of the product for the line item, in the whole currency