

Official Data Dictionary: vw_Master_Modeling_Data_v1

Document Version: 1.0

Date: 2025-07-16

View Name: vw_Master_Modeling_Data_v1

Description: This view is the definitive, denormalized source of truth for all e-commerce transactions. It consolidates data from sales, customer, product, and regional sources into a single, analytics-ready table. Its primary purpose is to support business intelligence, reporting, and modeling activities for government oversight and internal analysis. The lineage of this data is visually represented in the project's official Data Lineage Diagram.

Enhancements for Government-Level Documentation

This data dictionary has been enhanced to meet the rigorous standards of government documentation. In addition to standard column descriptions, the following metadata fields have been included to ensure clarity, data quality assurance, and unambiguous interpretation:

- 1. Data Type: The specific technical data type (e.g., STRING, INTEGER, FLOAT64, DATE).
- 2. Source Table(s): The original source table(s) from the Bronze Layer where the data originates. This provides clear proof.
- 3. Nullable: Indicates if the field can be empty (YES) or not (NO). This is critical for data integrity checks.
- 4. Unique: Indicates if every value in this column is unique across the entire view (YES/NO).

Column Name	Data Type	Source Table	Null -able	Unique	Description
Identifiers					
row_id	Integer	ecom_sales	No	Yes	Transaction-level ID for each row in the fact table of ecom_sales. The primary key of the original ecom_sales table
order_id	String	ecom_sales	Yes	No	ID of each order, including many products. Format: 2 digits (Country code)-XXXX-TEXT
order_date	Date	ecom_sales	Yes	No	Date of the order. Each order matches with only 1 order date Format: YYYY-MM-DD This is the primary partition key
Base financial metrics					
gross_sales	Decimal	ecom_sales	Yes	No	The total revenue for the line item before any discounts. This sales includes all

					quantities of a product in a specific order_id (Original sales column).
discount	Decimal	ecom_sales	Yes	No	The discount percentage applied to the line item Format: e.g. 0.2 for 20% This discount is specific to a particular product in an order_id
quantity	Integer	ecom_sales	Yes	No	The number of units of the product sold in a respective order_id
profit	Decimal	ecom_sales	Yes	No	The total profit generated from a particular product in a order_id
Derived metrics – adding through the Silver layer					
sales after discount	Decimal	CTE table	Yes	No	gross_sales – discount
unit_price_bdiscount	Decimal	CTE table	Yes	No	gross_sales / quantity
profit_margin_bdiscount	Decimal	CTE table	Yes	No	profit / gross_sales
implied_cost_record_level	Decimal	CTE table	Yes	No	sales – discount - profit
Implied_unit_cost	Decimal	CTE table	Yes	No	(sales – discount – profit) / quantity
Customer dimension – The view won't take the personal information of names and emails					
customer_id	String	customer	No	Yes	Format: 2 digit (Country code) – XXXXX The primary code of the original customer table
birth_date	Date	customer	Yes	No	Format: yyyy-mm-dd
marital_status	String	customer	Yes	No	S (Single), M (Married)
gender	String	customer	Yes	No	F (Female), M (Male)
annual_income	Integer	customer	Yes	No	
education_level	String	customer	Yes	No	High School, Partial High School, Partial College, Bachelors, Graduate Degree
occupation	String	customer	Yes	No	Professional, Clerical, Manual, Management, Skilled Manual
home_owner	String	customer	Yes	No	N (No house), Y (Have at least 1 house)
Product dimension – including the product category					
Product_code	String	product	No	Yes	Format: PXXXXXX This is the primary column of the original product table

product_name	String	product	Yes	No	
product_category	String	product	Yes	No	Body care Face care Hair care Home and Accessories Make up
product_subcategory	String	product	Yes	No	<i>Body care ></i> face masks and exfoliators vitamins and supplements bath oils, bubbles and soaks Nail care products Body soaps and washes hand creams body moisturizers <i>Face care ></i> face moisturizing products <i>Hair care ></i> hair colors and toners shampoos and conditioners <i>Home and Accessories ></i> fragrances Accessories brushes and applicators candles, sprays, diffusers <i>Make up ></i> foundations and concealers Lipsticks Eye shadows and pencils
Region dimension – including the respective country and region to which the product serves					
region_code	String	region	No	Yes	Format: RXXXX The primary code of the original region table
market	String	region	Yes	No	Africa Asia Pacific Europe LATAM USCA
region	String	region	Yes	No	<i>Africa ></i> North Africa Western Africa Southern Africa Eastern Africa <i>Asia Pacific></i> Oceania

					Western Asia Southeastern Asia Eastern Asia Southern Asia <i>Europe</i> > Eastern Europe Northern Europe Southern Europe Western Europe <i>LATAM</i> > Central America Caribbean South America <i>USCA</i> > Central US Eastern US Western US Southern US Canada
country	String	region	Yes	No	
state	String	region	Yes	No	
city	String	region	Yes	No	