Data Dictionary – Olist DW

OpenAI. (2025). *ChatGPT (GPT-5)* [Large language model]. OpenAI. https://chat.openai.com

Dimension Tables

dim_date

Column Name	Data Type	Description	Key / Notes
date_key	INT	Surrogate key for date	PK
full_date	DATE	Full calendar date	
day_of_week_number	SMALLINT	ISO weekday number (1=Mon)	
day_of_week_name	VARCHAR(10)	Name of the day	
day_of_month	SMALLINT	Day of month (1-31)	
day_of_year	SMALLINT	Day of year (1-365/366)	
week_of_year	SMALLINT	ISO week number	
month_number	SMALLINT	Month number (1-12)	
month_name	VARCHAR(10)	Month name	
quarter_number	SMALLINT	Quarter number (1-4)	
quarter_name	VARCHAR(10)	Quarter name (First, Second, ...)	
year	SMALLINT	Year	
is_weekend	BOOLEAN	TRUE if Saturday/Sunday	
is_holiday	BOOLEAN	TRUE if holiday (placeholder)	

dim_customer

Column Name	Data Type	Description	Key / Notes
customer_key	SERIAL	Surrogate key for customer	PK
customer_unique_id	VARCHAR(50)	Unique customer identifier	Unique
customer_zip_code_prefix	VARCHAR(20)	ZIP code prefix	
customer_city	VARCHAR(50)	City	
customer_state	VARCHAR(50)	State (UF)	

dim_products

Column Name	Data Type	Description	Key / Notes
product_key	SERIAL	Surrogate key for product	PK
product_id	VARCHAR(50)	Unique product identifier	Unique
product_category_name	VARCHAR(100)	Product category	
product_photos_qty	INT	Number of product photos	
product_weight_g	INT	Weight in grams	
product_length_cm	INT	Product length	
product_height_cm	INT	Product height	
product_width_cm	INT	Product width	

dim_seller

Column Name	Data Type	Description	Key / Notes
seller_key	SERIAL	Surrogate key for seller	PK
seller_id	VARCHAR(50)	Unique seller identifier	Unique
seller_city	VARCHAR(100)	Seller city	
seller_state	VARCHAR(50)	Seller state	

dim_payment_type

Column Name	Data Type	Description	Key / Notes
payment_type_key	SERIAL	Surrogate key for payment	PK
payment_type	VARCHAR(50)	Type of payment	Unique

dim_order_status

Column Name	Data Type	Description	Key / Notes
order_status_key	SERIAL	Surrogate key for status	PK
order_status	VARCHAR(50)	Status name	Unique

Fact Tables

fact_order_sales

Column Name	Data Type	Description	Key / Notes
order_sales_key	BIGSERIAL	Surrogate key for sales line	PK
order_id	VARCHAR(50)	Original order ID	
order_item_id	VARCHAR(50)	Original item ID in order	
product_key	INT	FK to dim_products	FK
date_key	INT	FK to dim_date	FK
seller_key	INT	FK to dim_seller	FK
customer_key	INT	FK to dim_customer	FK
shipping_limit_date	DATE	Latest shipping date	
price	DECIMAL(10,2)	Product price	
freight_value	DECIMAL(10,2)	Freight cost	
total_item_value	DECIMAL(10,2)	price + freight_value (calculated)	
quantity	INT	Quantity of product	Default 1
profit_margin	DECIMAL(10,2)	Profit margin	

fact_order_reviews

Column Name	Data Type	Description	Key / Notes
review_key	BIGSERIAL	Surrogate key for review	PK
review_id	VARCHAR(50)	Original review ID	Unique
order_id	VARCHAR(50)	Original order ID	
customer_key	INT	FK to dim_customer	FK
review_date_key	INT	FK to dim_date (review date)	FK
answer_date_key	INT	FK to dim_date (answer date)	FK
review_score	INT	Score 1–5	
has_comment	BOOLEAN	TRUE if comment present	
has_title	BOOLEAN	TRUE if title present	
review_creation_date	TIMESTAMP	Timestamp of review creation	
review_answer_timestamp	TIMESTAMP	Timestamp of answer	

fact_payments

Column Name	Data Type	Description	Key / Notes
payment_key	SERIAL	Surrogate key for payment	PK
order_id	VARCHAR(50)	Original order ID	
customer_key	INT	FK to dim_customer	FK
date_key	INT	FK to dim_date	FK
payment_type_key	INT	FK to dim_payment_type	FK
payment_value	DECIMAL(12,2)	Payment amount	

fact_orders

Column Name	Data Type	Description	Key / Notes
order_key	SERIAL	Surrogate key for order	PK
order_id	VARCHAR(50)	Original order ID	
customer_key	INT	FK to dim_customer	FK
date_key	INT	FK to dim_date	FK
order_status_key	INT	FK to dim_order_status	FK
total_amount	DECIMAL(14,2)	Total order amount	
freight_value	DECIMAL(12,2)	Total freight for order	

Notes on Constraints and Indexes

Primary Keys (PK): All surrogate keys in dimension and fact tables.

Foreign Keys (FK): Link fact tables to dimension tables.

Unique Constraints:

dim_customer.customer_unique_id

dim_products.product_id

fact_order_sales unique(order_id, order_item_id)

fact_order_reviews.review_id

Check Constraints:

fact_order_reviews.review_score between 1 and 5

Indexes for Performance:

Fact tables have indexes on key columns (e.g., date_key, product_key, customer_key) for faster queries.