

DICIONÁRIO DE DADOS

1. ETAPA DE OLTP

Tabela: oltp_customers

Descrição: Armazena informações dos clientes

Nome da Coluna	Tipo de Dado	Descrição
customer_id	VARCHAR	Identificador único do cliente no sistema Olist
customer_unique_id	VARCHAR	Identificador único do cliente (mesmo cliente em múltiplos pedidos)
zip_code_prefix	VARCHAR	Prefixo do CEP do cliente (primeiros 5 dígitos)
customer_city	VARCHAR	Cidade de residência do cliente
customer_state	VARCHAR	Estado de residência do cliente (sigla)

Tabela: oltp_sellers

Descrição: Armazena informações dos vendedores

Nome da Coluna	Tipo de Dado	Descrição
seller_id	VARCHAR	Identificador único do vendedor
zip_code_prefix	VARCHAR	Prefixo do CEP do vendedor
seller_city	VARCHAR	Cidade do vendedor
seller_state	VARCHAR	Estado do vendedor (sigla)

Tabela: oltp_products

Descrição: Armazena informações dos produtos

Nome da Coluna	Tipo de Dado	Descrição
product_id	VARCHAR	Identificador único do produto
product_category_name	VARCHAR	Nome da categoria do produto em português
product_name_length	INTEGER	Quantidade de caracteres no nome do produto
product_description_length	INTEGER	Quantidade de caracteres na descrição do produto
product_photos_qty	INTEGER	Número de fotos do produto disponíveis
product_weight_g	DOUBLE	Peso do produto em gramas
product_length_cm	DOUBLE	Comprimento do produto em centímetros
product_height_cm	DOUBLE	Altura do produto em centímetros
product_width_cm	DOUBLE	Largura do produto em centímetros

Tabela: oltp_order_payments

Descrição: Armazena informações de pagamento dos pedidos

Nome da Coluna	Tipo de Dado	Descrição
order_id	VARCHAR	Identificador do pedido
payment_sequential	INTEGER	Número sequencial para pagamentos múltiplos
payment_type	VARCHAR	Método de pagamento (credit_card, boleto, voucher, debit_card)
payment_installments	INTEGER	Número de parcelas do pagamento
payment_value	DOUBLE	Valor do pagamento

Tabela: oltp_orders

Descrição: Armazena informações dos pedidos

Nome da Coluna	Tipo de Dado	Descrição
order_id	VARCHAR	Identificador único do pedido
customer_id	VARCHAR	Identificador do cliente que fez o pedido
order_status	VARCHAR	Status atual do pedido (ex: delivered, shipped, canceled)
order_purchase_timestamp	TIMESTAMP	Data e hora do momento da compra
order_approved_at	TIMESTAMP	Data e hora da aprovação do pagamento
order_delivered_carrier_date	TIMESTAMP	Data e hora da entrega à transportadora
order_delivered_customer_date	TIMESTAMP	Data e hora da entrega final ao cliente
order_estimated_delivery_date	TIMESTAMP	Data e hora estimada para entrega

Tabela: oltp_order_items

Descrição: Armazena os itens de cada pedido

Nome da Coluna	Tipo de Dado	Descrição
order_id	VARCHAR	Identificador do pedido
order_item_id	INTEGER	Número sequencial identificando a ordem dos itens no pedido
product_id	VARCHAR	Identificador do produto
seller_id	VARCHAR	Identificador do vendedor
shipping_limit_date	TIMESTAMP	Data limite para envio pelo vendedor
price	DOUBLE	Preço do produto
freight_value	DOUBLE	Valor do frete do item

Tabela: oltp_order_reviews

Descrição: Armazena avaliações dos pedidos

Nome da Coluna	Tipo de Dado	Descrição
review_id	VARCHAR	Identificador único da avaliação
order_id	VARCHAR	Identificador do pedido avaliado
review_score	INTEGER	Nota da avaliação (1 a 5)
review_comment_title	VARCHAR	Título opcional do comentário
review_comment_message	VARCHAR	Mensagem opcional do comentário
review_creation_date	TIMESTAMP	Data de criação da avaliação
review_answer_timestamp	TIMESTAMP	Data da resposta do vendedor à avaliação

Tabela: oltp_geolocation

Descrição: Armazena informações geográficas

Nome da Coluna	Tipo de Dado	Descrição
zip_code_prefix	VARCHAR	Prefixo do CEP (primeiros 5 dígitos)
geolocation_lat	DOUBLE	Latitude da localização
geolocation_lng	DOUBLE	Longitude da localização
geolocation_city	VARCHAR	Cidade da localização
geolocation_state	VARCHAR	Estado da localização (sigla)

Tabela: oltp_category_translation

Descrição: Traduções das categorias de produtos

Nome da Coluna	Tipo de Dado	Descrição
product_category_name	VARCHAR	Nome da categoria em português
product_category_name_english	VARCHAR	Nome da categoria em inglês

2.FATO E DIMENSÕES DO DATA WAREHOUSE

Dimensão: dim_date

Descrição: Dimensão de data para análises temporais

Nome da Coluna	Tipo de Dado	Descrição
date_key	INTEGER	Chave substituta no formato YYYYMMDD
full_date	DATE	Data completa
year	INTEGER	Ano (ex: 2024)
quarter	INTEGER	Trimestre (1-4)
month	INTEGER	Mês numérico (1-12)
month_name	VARCHAR	Nome do mês (ex: Janeiro)
week	INTEGER	Número da semana no ano (1-52)
day	INTEGER	Dia do mês (1-31)
weekday	VARCHAR	Dia da semana (ex: Segunda-feira)

Dimensão: dim_customer

Descrição: Dimensão de clientes com SCD Type 2

Nome da Coluna	Tipo de Dado	Descrição
customer_key	INTEGER	Chave substituta do cliente (auto-incremento)
customer_id	VARCHAR	ID original do cliente (chave natural)
customer_unique_id	VARCHAR	ID único do cliente entre pedidos
customer_city	VARCHAR	Cidade do cliente
customer_state	VARCHAR	Estado do cliente
zip_code_prefix	VARCHAR	Prefixo do CEP do cliente
start_date	DATE	Data de início da validade do registro
end_date	DATE	Data de fim da validade do registro (NULL se atual)
is_current	BOOLEAN	Indicador se é o registro atual (TRUE/FALSE)

Dimensão: dim_payment_type

Descrição: Dimensão de tipos de pagamento

Nome da Coluna	Tipo de Dado	Descrição
payment_type_key	INTEGER	Chave substituta do tipo de pagamento (auto-incremento)
payment_type	VARCHAR	Descrição do tipo de pagamento (credit_card, boleto, etc)

Dimensão: dim_product

Descrição: Dimensão de produtos com SCD Type 2

Nome da Coluna	Tipo de Dado	Descrição
product_key	INTEGER	Chave substituta do produto (auto-incremento)
product_id	VARCHAR	ID original do produto (chave natural)
product_category_name	VARCHAR	Nome da categoria em português
product_category_english	VARCHAR	Nome da categoria em inglês
product_weight_g	DOUBLE	Peso do produto em gramas
product_length_cm	DOUBLE	Comprimento do produto em cm
product_height_cm	DOUBLE	Altura do produto em cm
product_width_cm	DOUBLE	Largura do produto em cm
start_date	DATE	Data de início da validade do registro
end_date	DATE	Data de fim da validade do registro (NULL se atual)
is_current	BOOLEAN	Indicador se é o registro atual (TRUE/FALSE)

Dimensão: dim_seller

Descrição: Dimensão de vendedores com SCD Type 2

Nome da Coluna	Tipo de Dado	Descrição
seller_key	INTEGER	Chave substituta do vendedor (auto-incremento)
seller_id	VARCHAR	ID original do vendedor (chave natural)
seller_city	VARCHAR	Cidade do vendedor
seller_state	VARCHAR	Estado do vendedor
zip_code_prefix	VARCHAR	Prefixo do CEP do vendedor
start_date	DATE	Data de início da validade do registro
end_date	DATE	Data de fim da validade do registro (NULL se atual)
is_current	BOOLEAN	Indicador se é o registro atual (TRUE/FALSE)

Fato: fact_sales

Descrição: Tabela fato de vendas com granularidade por item de pedido

Nome da Coluna	Tipo de Dado	Descrição
sales_key	BIGINT	Chave primária da tabela fato
date_key	INTEGER	Chave da data da compra (referência dim_date)
customer_key	INTEGER	Chave do cliente (referência dim_customer)
product_key	INTEGER	Chave do produto (referência dim_product)
seller_key	INTEGER	Chave do vendedor (referência dim_seller)
payment_type_key	INTEGER	Chave do tipo de pagamento (referência dim_payment_type)
price	DOUBLE	Preço unitário do produto
freight_value	DOUBLE	Valor do frete do item
payment_value	DOUBLE	Valor total do pagamento
review_score	INTEGER	Nota da avaliação do pedido (1-5)
quantity	INTEGER	Quantidade de itens (normalmente 1)
order_id	VARCHAR	ID original do pedido (para rastreabilidade)
order_item_id	INTEGER	ID original do item do pedido (para rastreabilidade)