

Dataset Name	Primary Key	Description
olist_customers_dataset.csv	customer_id	Contains unique customers along with their ZIP code, city, and state. customer_unique_id can help identify repeat customers across multiple customer_ids (used when a customer makes multiple accounts or purchases).
olist_order_items_dataset.csv	order_id, product_id, seller_id	Line-item level detail for each product in an order, including product ID, seller ID, pricing, and freight/shipping costs.
olist_order_payments_dataset.csv	order_id	Payment-related information per order, including payment type, installment details, and payment value.
olist_order_reviews_dataset.csv	review_id, order_id	Customer reviews for orders, including score, comments, creation date, and response timestamp. order_id links back to the Orders dataset.
olist_orders_dataset.csv	order_id	Contains order-level information, such as status, purchase timestamp, approval, shipment, delivery dates, and customer ID linking back to the Customers table.
olist_products_dataset.csv	product_id	Metadata about products such as category, name length, description length, number of photos, weight, and dimensions.
olist_sellers_dataset.csv	seller_id	Information about individual sellers including their ZIP code prefix, city, and state.
olist_geolocation_dataset.csv	geolocation_zip_code_prefix	Geolocation coordinates associated with ZIP code prefixes, cities, and states. It does not have a strict unique row identifier, but the combination can serve as a surrogate primary key.
product_category_name_translation.csv	product_category_name	Mapping between original Portuguese product category names and their English translations.