Data dictionary:

1. olist_customers_dataset:
Customer_id: Unique identifier for a customer entry in the dataset; may appear multiple times for the same person across different orders.
Customer_unique_id: Unique identifier for each customer
Customer_zip_code_prefix: Postal code prefix of the customer's address
customer_city: City where the customer is located
Customer_state: state where the customer is located

2. Olist_orders_dataset:
order_id: Unique identifier for each order
order_status: Current status of the order (e.g., delivered, shipped, canceled…)
Order_purchase_timestamp: Date and time when the order was placed
order_approved_at: Date and time when the order was approved
order_delivered_carrier_date: Date and time when the order was handed over to the shipping carrier
order_delivered_customer_date: Date and time when the order was delivered to the customer
order_estimated_delivery_date: Estimated delivery date provided at the time of purchase

3.olist_order_items_dataset:
order_id: unique identifier for the order
order_item_id: unique identifier for each item within an order
product_id: Unique identifier for the product
seller_id: Unique identifier for the seller of the product
shipping_limit_date: Maximum date for shipping the item
price: Price of the product
freight_value: shipping fee for the product

4.olist_order_reviews_dataset:
review_id: Unique identifier for each review
order_id: Identifier linking the review to a specific order
review_score: Rating given by the customer (from 1 to 5)
review_comment_title: Optional title summarizing the customer's feedback
review_comment_message: Customer's written review or detailed feedback
review_creation_date: Date when the review was submitted by the customer
review_answer_timestamp: Date/time when the seller or platform responded to the review

5.olist_products_dataset:
product_id: Unique identifier for each product
product_category_name: Category to which the product belongs

product_name_lenght: Length of the product name (number of characters)
product_description_lenght: Length of the product description (number of characters)
product_photos_qty: Number of photos available for the product
product_weight_g: Weight of the product in grams
product_length_cm: Product length in centimeters
product_height_cm: Product height in centimeters
product_width_cm: Product width in centimeters

6.olist_sellers_dataset:
seller_id: unique identifier for each seller
seller_zip_code_prefix: zip code prefix for the seller's location
seller_city: city where the seller is located
seller_state: state where the seller is located

7.olist_geolocation_dataset:
geolocation_zip_code_prefix: zip code prefix for the location
geolocation_lat: latitude of the location
geolocation_lng: longitude of the location
geolocation_city: city of the location
geolocation_state: state of the location

8.olist_order_payments_dataset:
order_id: Unique identifier linking the payment to a specific order
Payment_sequential: Indicates the sequence number of the payment attempt for the order
order payment_type: Type of payment used (e.g., credit card, boleto, voucher)
payment_installments: Number of installments chosen by the customer
payment_value: Total amount paid for the order

9. Product_category_name_translation:
 product_category_name: name of the product category in Portuguese
product_category_name_english: name of the product category in English