

## Data Dictionary

**Customer Dataset:** Use it to locate the orders' delivery location and identify distinct customers in the orders dataset.

**customer\_id:** Key to the orders dataset. Each order has a unique customer\_id.

**customer\_unique\_id:** Unique identifier for each customer

**customer\_zip\_code\_prefix:** first five digits of customer zip code

**customer\_city:** Customer's city name

**customer\_state:** Customer's State name

**Geolocation Dataset:** Use it to measure the distances between vendors and customers and plot maps.

**geolocation\_zip\_code\_prefix:** first five digits of zip code

**geolocation\_lat:** Latitude

**geolocation\_lng:** Longitude

**geolocation\_city:** City Name

**geolocation\_state:** State Name

**Order Items Dataset:** includes information on the things that were bought in each order.

**order\_id:** order unique identifier

**order\_item\_id:** sequential number identifying number of items included in the same order.

**product\_id:** product unique identifier

**seller\_id:** seller unique identifier

**shipping\_limit\_date:** Shows the seller shipping limit date for handling the order over to the logistic partner.

**Price:** item price

**freight\_value:** item freight value item (if an order has more than one item the freight value is splitted between items)

**Payments Dataset:** contains information on the orders' payment choices.

**order\_id:** unique identifier of an order.

**payment\_sequential:** customer may pay an order with more than one payment method. If he does so, a sequence will be created to accommodate all payments.

**payment\_type:** method of payment chosen by the customer.

**payment\_installments:** number of installments chosen by the customer.

**payment\_value:** transaction value.

**Order Reviews Dataset:** contains information about the consumer reviews.

**review\_id:** unique review identifier

**order\_id:** unique order identifier

**review\_score:** Note ranging from 1 to 5 given by the customer on a satisfaction survey.

**review\_comment\_title:** Comment title from the review left by the customer, in Portuguese.

**review\_comment\_message:** Comment message from the review left by the customer, in Portuguese.

**review\_creation\_date:** Shows the date in which the satisfaction survey was sent to the

**customer\_review\_answer\_timestamp:** Shows satisfaction survey answer timestamp.

## Order Dataset

order\_id: unique identifier of the order.

customer\_id: key to the customer dataset. Each order has a unique customer\_id.

order\_status: Reference to the order status (delivered, shipped, etc).

order\_purchase\_timestamp: Shows the purchase timestamp.

order\_approved\_at: Shows the payment approval timestamp.

order\_delivered\_carrier\_date: Shows the order posting timestamp. When it was handled to the logistic partner.

order\_delivered\_customer\_date: Shows the actual order delivery date to the customer.

order\_estimated\_delivery\_date: Shows the estimated delivery date that was informed to customer at the purchase moment.

Products Dataset: includes information on the merchandise sold

product\_id: unique product identifier

product\_category\_name: root category of product, in Portuguese.

product\_name\_length: number of characters extracted from the product name.

product\_description\_length: number of characters extracted from the product description.

product\_photos\_qty: number of product published photos

product\_weight\_g: product weight measured in grams.

product\_length\_cm: product length measured in centimeters.

product\_height\_cm: product height measured in centimeters.

product\_width\_cm: product width measured in centimeters.

Sellers Dataset: includes information on the vendors who filled orders placed. Use it to locate the seller's business and determine which seller shipped each item.

seller\_id: seller unique identifier

seller\_zip\_code\_prefix: first 5 digits of seller zip code

seller\_city: seller city name

seller\_state: seller state

Category Name Translation: English translation of the product category name.

product\_category\_name: category name in Portuguese

product\_category\_name\_english: category name in English



