**Data dictionary and data wrangling/manipulations done.**

**olist\_order\_payments\_dataset.csv**

order\_id- unique identifier of an order.

payment\_sequentials- a 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\_types- method of payment chosen by the customer.

payment\_installments-number of installments chosen by the customer.

payment\_values- transaction value.

#### olist\_order\_items\_dataset.csv

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)

**Data cleaning/checks/manipulation**

Duplicates were dropped and null values were checked for.

There are 775 missing entries in the olist order\_items (take note the one missing id in olist order\_payment is present in here)

There is one missing id/entry in the olist order\_payment - bfbd0f9bdef84302105ad712db648a6c

There are 775 missing entries in the olist order\_items (take note the one missing id in olist order\_payment is present in here) There is one missing id/entry in the olist order\_payment - bfbd0f9bdef84302105ad712db648a6c

An order might have multiple items.