Handling duplicates:

There are several duplicated Rows in the order\_items table: see the 1.Duplicated\_IN\_orderitems.csv - 2.duplicated\_order\_id\_counts.csv

I removed them using the code block 2, in this code block I removed rows with row item\_id bigger than 1 (bigger than one is only happens when row is duplicated, all other values of rows are the same)

There is a complex condition on order\_payment Table see what you should do for handling duplicated rows

<https://www.kaggle.com/datasets/olistbr/brazilian-ecommerce/data>

read this link too understand the data.