Respected Sir,

Introduction

We have been given data of transactions approximation of a week.With the feature columns transaction\_id, timestamp, product\_id, category, customer\_type, unit\_price, quantity, total, payment\_type

Summarize your findings

* We have payment\_type classified into 4 groups cash, credit card, e-wallet, debit card. And the majority of them using is cash , these classes are uniformly distributed
* The data types of columns transaction\_id is object, timestamp is object, product\_id is object, category is object, customer\_type is object, unit\_price is float, quantity is int, total is float, payment\_type is object
* Range of unit\_price is 23.8 and quantity is 3
* When we analyse the scatter plot we get to see that points are aligned in horizontal
* And once we plot the hist graph the frequency of quantity is high

Recommendations

* It would be better if we get to have more features for better analysis , as more the number of features better the model because too less features may not give the exact prediction
* It would be better if the data is a bit large for making train and test sets and training the model accordingly , which in turn helps in feature engineering
* To better stock the items that they sell , we need classification model , for the classification model we need uniformly distributed data hence when we increase the size of dataset we must ensure the uniformity in classes