

1. *Think about what could be going wrong with our calculation. Think about a better way to evaluate this data.*

Answer:

The data has a few very large transactions that either could be fraudulent or contains errors. The problematic data can be grouped into two different groups, that is: 1) 17 credit transactions of \$704000 of 2000 items, from shopId 42 and userId 607; and 2) 46 transactions that were greater than \$10000.

A few ways to better evaluate the data include:

- Only including data that is within a valid range, that is eliminating outliers from the dataset
- Taking the median instead of the mean
- Perhaps taking a weighted average

2. *What metric would you report for this dataset?*

Answer:

Rather than taking the mean, the median should be used instead.

3. *What is its value?*

Answer:

284.0