

Problem 5 (15 points)

Imagine you are designing a table to store recent transactions for an online shopping platform and there are 1 trillion transactions. You want to record the following information:

- user id
- user name
- item id
- item name
- transaction id
- amount of money (\$) for the transaction (e.g. \$7.81, \$470.80, etc)

- a.* What data type should you use for each column? You need to fill one of the following data types: byte, short, int, long, float, double, boolean, char.

Column	Data Type
User ID	int
User Name	char
Item ID	int
Item Name	char
Transaction ID	long
Amount	double

- b.* What is the size of each row in bytes? Think about the size of each column by selecting proper data types. You need to select the most suitable data type for each column by considering efficiency.

Assuming User Name has 20 characters, and Item Name has 30 characters. Therefore User Name will be char[20], taking 20 Bytes, and Item Name will be char[30], taking 30 Bytes.

Column	Data Type	Size (Bytes)
User ID	int	4
User Name	char[20]	20
Item ID	int	4
Item Name	char[30]	30
Transaction ID	long	8
Amount	double	8
Total Row Size		74 Bytes

Each Row will consume 74 Bytes

- c.* What is the size of the table in TB?

Total size = Row size * Number of Transactions = $74 * 10^{12}$ bytes

From Bytes to Terabytes (1TB = 10^{12} Bytes): $(74 * 10^{12}) / 10^{12} = 74\text{TB}$

Answer: 74TB