Q4 Part a, b, c

Monday, 29 April 2024 6:41 PM

Question 4 Relational algebra and normalisation

Consider the following schema for product sales for an online shop.

product(pid, cat, price)
customer(cid, cname)
sale(pid, cid, date, qty)

("cat" is short for category and "qty" for quantity). For estimating efficiency, assume around 100 products, 10,000 customers and several million sale records.

(a) State in English what the following relational algebra expression computes:

$$\pi_{\text{cname}}(\sigma_{\text{qty}*\text{price}\geq 100}(\sigma_{\text{cat}=\text{'clothes'}}((\text{product} \bowtie \text{customer}) \bowtie \text{sale})))$$

[2 marks]

(b) The expression given above is considered inefficient. State all the ways in which it is inefficient. **[4 marks]**

(c) Give a more efficient version of the expression. Make it as efficient as you can, and explain how it achieves efficiency. **[6 marks]**