

Cho lược đồ CSDL toàn cục sau:

KHACHHANG(MaKH(10), hoten(20), tinh/TP(20), tel(10),phai(1))

SANPHAM(MaSP(10), tenSP(30), mota(50), loai(15))

HOADON(SHD(12), MaHK(10), ngay(8))

CTHD(SHD(12), MaSP(10), SL(4), gia(4))

Câu hỏi

1. Viết câu lệnh SQL trả lời cho câu Tìm họ tên và tỉnh thành của các khách hàng nữ có mua sản phẩm thuộc loại '*chăm sóc sắc đẹp*'.

2. Viết biểu thức ĐSQL tối ưu tương đương câu lệnh SQL trên

3. Thay biểu thức ĐSQLH toàn cục tối ưu bởi các đoạn và đơn giản biểu thức thu được. Biết rằng CSDL này được phân thành các đoạn như sau:

Site 1 : $KH_1 = \sigma_{\text{phai} = 'F'}(KHACHHANG)$

Site 2 : $KH_2 = \sigma_{\text{phai} = 'M'}(KHACHHANG)$

Site 3 : $SP_1 = \sigma_{\text{loai} = 'thực phẩm'}(SANPHAM)$

Site 4 : $SP_2 = \sigma_{\text{loai} \neq 'thực phẩm'}(SANPHAM)$

Site 5 : $HD_1 = HOADON \bowtie KH_1$

Site 6 : $HD_2 = HOADON \bowtie KH_2$

Site 7 : $CTHD_1 = CTHD \bowtie SP_1$

Site 8 : $CTHD_2 = CTHD \bowtie SP_2$

4. Với các giả thuyết sau:

- Biết rằng chi phí di chuyển 1 bộ từ nút này qua nút khác bằng kích thước của 1 bộ
- Chỉ có 1% các sản phẩm khác thực phẩm là sản phẩm chăm sóc sắc đẹp.
- Mỗi sản phẩm thuộc loại chăm sóc sắc đẹp được mua ít nhất 1 lần.
- Có 1000 khách hàng nữ và 500 khách hàng nam
- Có 1000 sản phẩm loại thực phẩm và 5000 sản phẩm loại khác
- Trung bình mỗi hoá đơn có 5 mặt hàng, trong đó có 2 mặt hàng thực phẩm và 5 hoá đơn thì có 1 sản phẩm làm đẹp.
- Mỗi khách hàng nữ trung bình có 10 hoá đơn và nam 2 hoá đơn.
- Có 11 % khách hàng nữ có mua sản phẩm làm đẹp và mỗi khách hàng trung bình có mua 20 mặt hàng sản phẩm làm đẹp trong các hoá đơn (nam không mua các sản phẩm CSSĐ).
- Kết quả đặt tại nút 1

Để thực hiện câu truy vấn trên, hãy cho biết phép JOIN hay SEMI-JOIN sẽ cho chi phí truyền dữ liệu thấp hơn? Hãy chứng minh câu trả lời.