

Bài tập DW: Nhóm 1

Hãng dược phẩm ABC là một hãng sản xuất và phân phối dược phẩm. Hãng sản xuất các loại thuốc từ dược liệu và phân phối qua nhiều kênh khác nhau, bao gồm các nhà thuốc, bệnh viện, kênh phân phối trực tuyến. Hãng muốn xây dựng một Data Warehouse để phân tích hiệu quả kinh doanh, từ doanh thu bán hàng đến chi phí sản xuất, nhằm tối ưu hóa các quyết định chiến lược và tăng trưởng.

Biết rằng DW gồm các bảng dim và bảng fact như sau:

1. Các bảng dim

1.1. Bảng dim_product: Thông tin dược phẩm, gồm các thuộc tính:

- product_id: Mã sản phẩm
- product_name: Tên sản phẩm
- product_type: Loại thuốc (kê đơn, không kê đơn)
- manufacturer: Công ty sản xuất
- form: Dạng thuốc (viên, lỏng, mỡ)
- dosage: Hàm lượng

1.2. Bảng dim_customer: Thông tin khách, gồm các thuộc tính:

- customer_id: Mã khách hàng
- customer_name: Tên khách hàng
- customer_type: Loại khách hàng (Nhà thuốc, Bệnh viện, ...)
- location: Vị trí địa lý (thành phố, khu vực)
- contact_info: Thông tin liên hệ

1.3. Bảng dim_channel: Thông tin kênh phân phối, gồm các thuộc tính:

- channel_id: Mã kênh
- channel_name: Tên kênh (Nhà thuốc, Bệnh viện, Phân phối trực tuyến)
- channel_type: Loại kênh (Bán lẻ, Bán buôn)

1.4. Bảng dim_date: Thông tin về ngày, tháng, quý, gồm các thuộc tính:

- date_id: Mã ngày
- date: Ngày cụ thể
- month: Tháng
- quarter: Quý
- year: Năm
- week: Tuần

1.5. Bảng dim_supplier: Thông tin nhà cung cấp, gồm các thuộc tính:

- supplier_id: Mã nhà cung cấp

- supplier_name: Tên nhà cung cấp
- location: Địa điểm
- contact_info: Thông tin liên hệ

2. Các bảng fact

2.1. Bảng fact_sales: Thông tin giao dịch bán, gồm các thuộc tính:

- sale_id: Mã giao dịch
- product_id: Mã sản phẩm (Khóa ngoại từ dim_product)
- customer_id: Mã khách hàng (Khóa ngoại từ dim_customer)
- channel_id: Mã kênh phân phối (Khóa ngoại từ dim_channel)
- date_id: Mã ngày (Khóa ngoại từ dim_date)
- quantity_sold: Số lượng bán
- unit_price: Giá bán mỗi đơn vị
- total_sales: Tổng doanh thu ($\text{quantity_sold} * \text{unit_price}$)
- discount: Chiết khấu
- net_sales: Doanh thu sau chiết khấu

2.2. Bảng fact_production_costs: Thông tin chi phí sản xuất, gồm các thuộc tính:

- production_id: Mã sản xuất
- product_id: Mã sản phẩm (Khóa ngoại từ dim_product)
- supplier_id: Mã nhà cung cấp (Khóa ngoại từ dim_supplier)
- date_id: Mã ngày (Khóa ngoại từ dim_date)
- raw_material_cost: Chi phí nguyên liệu
- labor_cost: Chi phí lao động
- overhead_cost: Chi phí quản lý
- total_production_cost: Tổng chi phí sản xuất ($\text{raw_material_cost} + \text{labor_cost} + \text{overhead_cost}$)

3. Yêu cầu:

- 3.1. Hãy xây dựng và giả lập dữ liệu trong DW;
- 3.2. Phân tích doanh thu theo kênh phân phối: Viết câu truy vấn SQL để tính tổng doanh thu theo từng kênh phân phối tháng vừa qua, phân tích kênh nào mang lại doanh thu cao nhất?
- 3.3. Phân tích chi phí sản xuất theo từng loại dược phẩm: Viết câu truy vấn SQL để tính tổng chi phí sản xuất cho từng loại dược phẩm và phân tích dược phẩm nào có chi phí cao nhất trong năm nay?
- 3.4. Phân tích xu hướng doanh thu theo mùa: Viết câu truy vấn SQL để tính tổng doanh thu của từng loại thuốc theo mùa (mùa xuân, hạ, thu, đông)?

- 3.5. So sánh chi phí và doanh thu giữa các khu vực: Viết câu truy vấn SQL để phân tích tổng doanh thu và tổng chi phí sản xuất theo khu vực trong năm hiện tại?
 - 3.6. Phân tích lợi nhuận theo nhà cung cấp: Viết câu truy vấn SQL để tính lợi nhuận từ mỗi nhà cung cấp, bao gồm doanh thu bán hàng và chi phí sản xuất từ nhà cung cấp đó?
-