

CS636 - Data Warehouse

Data Warehouses - Tutorial

Enrico Franconi

`franconi@inf.unibz.it`

`http://www.inf.unibz.it/~franconi`

Faculty of Computer Science, Free University of Bozen-Bolzano

Data Warehouse Tutorial

Consider a data warehouse talking about sales, where the total items sold are stored, organised by customer order and product. Each customer order includes the name of the customer and the date of the order; each product includes a description of the product and its price.

1. Devise the relational schema (specifying the relations, the attributes, the primary keys, and the foreign keys) of the above data warehouse using the star schema.
2. Write a SQL query for the following report: “Select the customer(s) who made an order containing at least five products with different descriptions”.
3. Write a SQL query for the following report: “Select the customer(s) who made the order with the maximum total amount, considering both the price and the quantities of the products included in the order”.
4. Consider to add the new level *product categories* to the product dimension, which assigns a product category to each product. Devise the new relational star schema, and write a SQL query for the following report: “Select the total quantities of products sold aggregated by product category”.

(1)

Consider a data warehouse talking about sales, where the total items sold are stored, organised by customer order and product. Each customer order includes the name of the customer and the date of the order; each product includes a description of the product and its price.

```
product(prodcode, description, price)
```

```
order(ordercode, customer, date)
```

```
sale(ordercode, prodcode, quantity)
```

ordercode FK ref. order

prodcode FK ref. product

(2)

Write a SQL query for the following report: “Select the customer(s) who made an order containing at least five products with different descriptions”.

```
product(prodcode, description, price)
```

```
order(ordercode, customer, date)
```

```
sale(ordercode, prodcode, quantity)
```

```
select customer
```

```
from product, order, sale
```

```
where product.prodcode=sale.prodcode and
```

```
order.ordercode=sale.ordercode
```

```
group by customer, order
```

```
having count(distinct description) >= 5
```

(3)

Write a SQL query for the following report: “Select the customer(s) who made the order with the maximum total amount, considering both the price and the quantities of the products included in the order”.

```
select customer
from product,order,sale
where product.prodcode=sale.prodcode and
       order.ordercode=sale.ordercode
group by customer,order
having sum(price*quantity) >=ALL
      (select sum(price*quantity)
       from product,order,sale
       where product.prodcode=sale.prodcode and
             order.ordercode=sale.ordercode
       group by order)
```

(4)

Consider to add the new level *product categories* to the product dimension, which assigns a product category to each product. Devise the new relational star schema, and write a SQL query for the following report: “Select the total quantities of products sold aggregated by product category”.

```
product(prodcode, description, price, category)
```

```
order(ordercode, customer, date)
```

```
sale(ordercode, prodcode, quantity)
```

ordercode FK ref. order

prodcode FK ref. product

```
select category, sum(quantity)
```

```
from product, sale
```

```
where product.prodcode=sale.prodcode
```

```
group by category
```