

3CS452 : Data Warehousing and Data Mining Lab

Assignment 2 : To design and implement a data warehouse for a customer order processing system in a company. [Use *MySQL* Database]

Specifications :

The target of data warehouse system is an enterprise that consists of a number of stores located in different cities and states. Each store holds a variety of items in various quantity. In addition, the enterprise keep the information of the customers. There are two kinds of customers: walk-in led by tourism guide and mail-order by post address inclusive. The city location of the customer, together with the data of the customer's first order, is stored by the existing system. Each customer lives in one city only, and the enterprise will try to satisfy the customer's order items by the present stock in the city where the customer lives. Each customer order can be for any quantity of any number of items, and each order is uniquely identified by an order number.

The location of the stores is also recorded. Each store is located in one city, and there can be many stores in the city. Each city has one headquarter for coordinating all of its stores. The enterprise's goal is to meet all of the customer's requirements from stores located in the customer's city. If the requirement cannot be met, the company will turn to the other cities where the item can be found if there is any.

Some processing information is important for the enterprise. For example, the total quantity of item stored in each city. After every time an item is taken, the company needs to know the total quantities of the item in all the stores in a city.

The relational schema of the enterprise's current (operational) databases are:

Headquarter Database:

Relation Customer (Customer_id, Customer_name, City_id, First_order_date)

Relation Walk-in_customers (*Customer_id, tourism_guide, Time)

Relation Mail_order_customers (*Customer_id, post_address, Time)

Sales Databases:

Relation Headquarters (City_id, City_name, Headquarter_addr, State, Time)

Relation Stores (Store_id, *City_id, Phone, Time)

Relation Items (Item_id, Description, Size, Weight, Unit_price, Time)

Relation Stored_items (*Store_id, *Item_id, Quantity_held, Time)

Relation Order (Order_no, Order_date, Customer_id)

Relation Ordered_item (*Order_no, *Item_id, Quantity_ordered, Ordered_price, Time)

Where underlined are primary key and "*" prefixed are foreign keys.

Business requirements:

In order to meet users' demand, the data warehouse system extracts data from the existing two database into a data warehouse, and provides online analytical processing with roll up, drill

down, slice and dice features according to users selections based on dimension tables to meet the user requirements.

Build data warehouse / OLAP which will answer the following queries :

1. Find all the stores along with city, state, phone, description, size, weight and unit price that hold a particular item of stock.
2. Find all the orders along with customer name and order date that can be fulfilled by a given store.
3. Find all stores along with city name and phone that hold items ordered by given customer.
4. Find the headquarter address along with city and state of all stores that hold stocks of an item above a particular level.
5. For each customer order, show the items ordered along with description, store id and city name and the stores that hold the items.
6. Find the city and the state in which a given customer lives.
7. Find the stock level of a particular item in all stores in a particular city.
8. Find the items, quantity ordered, customer, store and city of an order.
9. Find the walk in customers, mail order customers and dual customers (both walk-in and mail order).

Prepare report as follows :

1. Introduction – objective and scope of the project
2. Business requirement – application specification of the data warehousing for the users.
3. Functional specification – input and output specification of the data warehousing
4. Data Warehousing Design – stepwise procedure methodology of designing the data warehousing including star schema.
5. Data cube implementation – computer automation of implementing the data warehousing loading data into data cubes.
6. Observations :
 - a. Online analytical processing reports – invoke commands or panels to generate OLAP reports.
 - b. Data verification – verify the OLAP reports source relational tables' data
7. Conclusion.

Dr. B. F. Momin
Course Teacher