# Assignment No 06

*Advanced Database System Lab(5CS372)*

**Name: Anushka Ajit Jadhav  PRN: 21520003  Batch:T1**

## Title:

To design and implement a data warehouse for a customer order processing system in a company.

## Theory:

**Data Warehouse:**

- A data warehouse is a large, centralized repository of data that is used to support business decision-making activities.
- It is designed to store data from multiple sources, such as transactional databases, applications, and external sources, in a format that is optimized for querying and analysis.
- Data warehouses typically use a process called ETL (extract, transform, load) to integrate data from different sources and transform it into a format that is consistent and usable for reporting and analysis.
- This process involves extracting data from the source systems, transforming it to meet the needs of the data warehouse, and loading it into the data warehouse.

**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 Headqarters (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, Qantity_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.

**SQL Queries:**

```
CREATE schema Headquarter;
CREATE schema Sales;

CREATE TABLE Customer (
    Customer_id INT PRIMARY KEY,
    Customer_name VARCHAR(50),
    City_id INT,
    First_order_date DATE,
    FOREIGN KEY (City_id) REFERENCES sales.Headqarters(City_id)
);

CREATE TABLE Walk_in_customers (
    Customer_id INT PRIMARY KEY,
    tourism_guide VARCHAR(50),
    Time DATETIME,
    FOREIGN KEY (Customer_id) REFERENCES Customer(Customer_id)
);

CREATE TABLE Mail_order_customers (
    Customer_id INT PRIMARY KEY,
    post_address VARCHAR(100),
    Time DATETIME,
    FOREIGN KEY (Customer_id) REFERENCES Customer(Customer_id)
);

CREATE TABLE Headqarters (
    City_id INT PRIMARY KEY,
    City_name VARCHAR(50),
    Headquarter_addr VARCHAR(100),
    State VARCHAR(50),
    Time DATETIME
);

CREATE TABLE Stores (
    Store_id INT PRIMARY KEY,
    City_id INT,
    Phone VARCHAR(20),
```

```
    Time DATETIME,
    FOREIGN KEY (City_id) REFERENCES Headqarters(City_id)
);

CREATE TABLE Items (
    Item_id INT PRIMARY KEY,
    Description VARCHAR(100),
    Size VARCHAR(20),
    Weight DECIMAL(10,2),
    Unit_price DECIMAL(10,2),
    Time DATETIME
);

CREATE TABLE Stored_items (
    Store_id INT,
    Item_id INT,
    Quantity_held INT,
    Time DATETIME,
    PRIMARY KEY (Store_id, Item_id),
    FOREIGN KEY (Store_id) REFERENCES Stores(Store_id),
    FOREIGN KEY (Item_id) REFERENCES Items(Item_id)
);

CREATE TABLE OrderT (
    Order_no INT PRIMARY KEY,
    Order_date DATE,
    Customer_id INT,
    FOREIGN KEY (Customer_id) REFERENCES headquarter.Customer(Customer_id)
);

CREATE TABLE Ordered_item (
    Order_no INT,
    Item_id INT,
    Quantity_ordered INT,
    Ordered_price DECIMAL(10,2),
    Time DATETIME,
    PRIMARY KEY (Order_no, Item_id),
    FOREIGN KEY (Order_no) REFERENCES OrderT(Order_no),
    FOREIGN KEY (Item_id) REFERENCES Items(Item_id)
);

CREATE schema Datawarehouse;

CREATE TABLE Sales_Fact (
    Store_id INT,
    Item_id INT,
    Quantity_held INT,
    Unit_price DECIMAL(10,2),
```

```
    PRIMARY KEY (Store_id, Item_id),
    FOREIGN KEY (Store_id) REFERENCES sales.Stores(Store_id),
    FOREIGN KEY (Item_id) REFERENCES sales.Items(Item_id)
);



CREATE TABLE Store_Dim (
    Store_id INT PRIMARY KEY,
    City_id INT,
    Phone VARCHAR(20),
    FOREIGN KEY (City_id) REFERENCES sales.Headqarters(City_id)
);

CREATE TABLE Item_Dim (
    Item_id INT PRIMARY KEY,
    Description VARCHAR(100),
    Size VARCHAR(20),
    Weight DECIMAL(10,2)
);

CREATE TABLE City_Dim (
    City_id INT PRIMARY KEY,
    City_name VARCHAR(50),
    State VARCHAR(50)
);

INSERT INTO Headqarters (City_id, City_name, Headquarter_addr, State, Time)
VALUES
(1, 'New York City', '123 Broadway, Suite 500', 'NY', '2022-01-01
00:00:00'),
(2, 'Los Angeles', '456 Main St, 12th Floor', 'CA', '2022-02-01 00:00:00'),
(3, 'Chicago', '789 Elm St, Suite 200', 'IL', '2022-03-01 00:00:00');

INSERT INTO Customer (Customer_id, Customer_name, City_id, First_order_date)
VALUES
(1, 'John Smith', 1, '2022-01-01'),
(2, 'Mary Johnson', 2, '2022-02-15'),
(3, 'David Lee', 3, '2022-03-20');

INSERT INTO Walk_in_customers (Customer_id, tourism_guide, Time) VALUES
(1, 'Tourist Guide A', '2022-01-01 10:00:00'),
(2, 'Tourist Guide B', '2022-02-15 15:30:00');

INSERT INTO Mail_order_customers (Customer_id, post_address, Time) VALUES
(3, '123 Main St, Anytown USA', '2022-03-20 09:00:00');

INSERT INTO Stores (Store_id, City_id, Phone, Time) VALUES
```

```
(1, 1, '555-1234', '2022-01-02 09:00:00'),
(2, 1, '555-5678', '2022-01-02 09:00:00'),
(3, 2, '555-9876', '2022-02-15 10:30:00'),
(4, 3, '555-4321', '2022-03-20 11:45:00');

INSERT INTO Items (Item_id, Description, Size, Weight, Unit_price, Time)
VALUES
(1, 'Widget', 'Small', 1.0, 10.00, '2022-01-01 00:00:00'),
(2, 'Gizmo', 'Large', 2.5, 25.00, '2022-02-01 00:00:00'),
(3, 'Thingamajig', 'Medium', 0.5, 5.00, '2022-03-01 00:00:00');

INSERT INTO Stored_items (Store_id, Item_id, Quantity_held, Time) VALUES
(1, 1, 100, '2022-01-02 09:00:00'),
(1, 2, 50, '2022-01-02 09:00:00'),
(2, 1, 75, '2022-01-02 09:00:00'),
(3, 2, 100, '2022-02-15 10:30:00'),
(4, 3, 200, '2022-03-20 11:45:00');

INSERT INTO OrderT (Order_no, Order_date, Customer_id) VALUES
(1001, '2022-02-15', 1),
(1000,'2022-01-01',2),
(1002,'2022-03-20',3);

INSERT INTO Ordered_item (Order_no, Item_id, Quantity_ordered,
Ordered_price, Time) VALUES
(1000,1,2,20.00,'2022-01-01 10:00:00');

INSERT INTO Ordered_item (Order_no, Item_id, Quantity_ordered,
Ordered_price, Time) VALUES
(1001,2,3,75.00,'2022-02-15 15:30:00');

INSERT INTO Ordered_item (Order_no, Item_id, Quantity_ordered,
Ordered_price,Time) VALUES
(1002,3,1,5.00,'2022-03-20 09:00:00');

/******************* inserting data into dimension
table**************************/

INSERT INTO Store_Dim (Store_id, City_id, Phone)
SELECT DISTINCT Store_id, City_id, Phone
FROM Sales.Stores;

INSERT INTO City_Dim (City_id, City_name, State)
SELECT DISTINCT City_id, City_name, State
FROM Sales.Headqarters;

INSERT INTO Sales_Fact (Store_id, Item_id, Quantity_held, Unit_price)
SELECT si.Store_id, si.Item_id, si.Quantity_held, i.Unit_price
```

```
FROM Sales.Stored_items si
JOIN Sales.Items i ON si.Item_id = i.Item_id;
```

**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.**

→

```
SELECT s.Store_id, c.City_name, c.State, s.Phone, i.Description,
i.Size, i.Weight, f.Unit_price
FROM Sales_Fact f
JOIN Store_Dim s ON f.Store_id = s.Store_id
JOIN Item_Dim i ON f.Item_id = i.Item_id
JOIN City_Dim c ON s.City_id = c.City_id
WHERE i.Description='Widget';
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| Store_id | City_name | State | Phone | Description | Size | Weight | Unit_price |
|---|---|---|---|---|---|---|---|
| 1 | New York City | NY | 555-1234 | Widget | Small | 1.00 | 10.00 |
| 2 | New York City | NY | 555-5678 | Widget | Small | 1.00 | 10.00 |

2. **Find all the orders along with customer name and order date that can be fulfilled by a given store.**

→

```
SELECT o.Order_no, c.Customer_name, o.Order_date
FROM Sales.OrderT o
JOIN Headquarter.Customer c ON o.Customer_id = c.Customer_id
JOIN Sales.Ordered_item oi ON oi.Order_no = o.Order_no
JOIN Sales.Stored_items si ON si.Item_id=oi.Item_id and si.Store_id=3;
```

| Order_no | Customer_name | Order_date |
|---|---|---|
| 1001 | John Smith | 2022-02-15 |

```
SELECT OrderT.Order_no, Customer.Customer_name, OrderT.Order_date
FROM Sales_Fact
JOIN Store_Dim ON Sales_Fact.Store_id = Store_Dim.Store_id
JOIN Item_Dim ON Sales_Fact.Item_id = Item_Dim.Item_id
JOIN Sales.Ordered_item ON Sales_Fact.Store_id = Ordered_item.Item_id
JOIN Sales.OrderT ON Ordered_item.Order_no = OrderT.Order_no
JOIN Headquarter.Customer ON OrderT.Customer_id =
Customer.Customer_id
WHERE Store_Dim.Store_id = 3;
```

Result Grid | Filter Rows:

| Order_no | Customer_name | Order_date |
|---|---|---|
| 1002 | David Lee | 2022-03-20 |

3. **Find all stores along with city name and phone that hold items ordered by given customer.**

→

```
SELECT DISTINCT Stores.Store_id, Headqarters.City_name, Stores.Phone
FROM Sales.Stores
INNER JOIN Sales.Headqarters ON Stores.City_id = Headqarters.City_id
INNER JOIN Sales.Stored_items ON Stores.Store_id =
Stored_items.Store_id
INNER JOIN Sales.Ordered_item ON Stored_items.Item_id =
Ordered_item.Item_id
INNER JOIN Sales.OrderT ON Ordered_item.Order_no = OrderT.Order_no
INNER JOIN Headquarter.Customer ON OrderT.Customer_id =
Customer.Customer_id
WHERE Customer.Customer_id = 2;
```

Result Grid | Filter Rows: | Export: | Wrap Cel

| Store_id | City_name | Phone |
|---|---|---|
| 1 | New York City | 555-1234 |
| 2 | New York City | 555-5678 |

4. **Find the headquarter address along with city and state of all stores that hold stocks of an item above a particular level.**

→

```
SELECT DISTINCT h.Headquarter_addr, h.City_name, h.State
FROM Sales.Stored_items si
JOIN Sales.Stores s ON si.Store_id = s.Store_id
JOIN Sales.Headqarters h ON s.City_id = h.City_id
JOIN Sales.Items i ON si.Item_id = i.Item_id
WHERE si.Quantity_held > 75
AND i.Description = 'Widget';
```

Result Grid | Filter Rows: | Export: | Wrap Cell

| Headquarter_addr | City_name | State |
|---|---|---|
| 123 Broadway, Suite 500 | New York City | NY |

5. **For each customer order, show the items ordered along with description, store id and city name and the stores that hold the items.**

→

```
SELECT o.Order_no, i.Description, si.Store_id, h.City_name AS
Store_City
FROM Sales.OrderT o
JOIN Sales.Ordered_item oi ON o.Order_no = oi.Order_no
JOIN Sales.Items i ON oi.Item_id = i.Item_id
JOIN Sales.Stored_items si ON oi.Item_id = si.Item_id
JOIN Sales.Stores s ON si.Store_id = s.Store_id
JOIN Sales.Headqarters h ON s.City_id = h.City_id
WHERE o.Customer_id = 3;
```

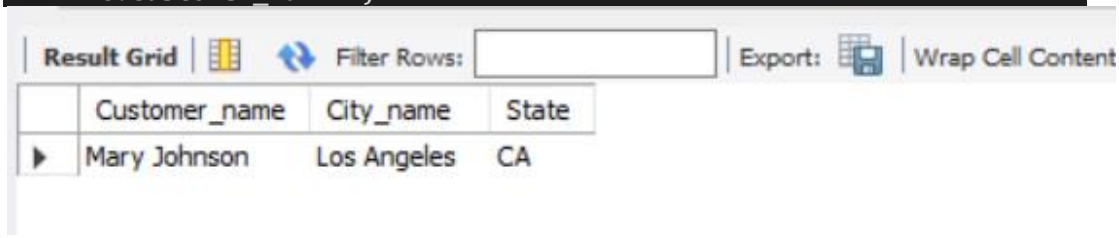
Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| Order_no | Description | Store_id | Store_City |
|---|---|---|---|
| 1002 | Thingamajig | 4 | Chicago |

**6. Find the city and the state in which a given customer lives.**

→

```
SELECT c.Customer_name, h.City_name, h.State
FROM Headquarter.Customer c
JOIN Sales.Headqarters h
ON c.City_id = h.City_id
WHERE c.Customer_id = 2;
```




**Conclusion:**

From this assignment I able to learn concept of data Warehouse and how we can implement data warehouse.