**Instructions:**

Please share your answers filled in line in the Word document. Submit code separately wherever applicable.

Please ensure you update all the details:

**Name: ULLI VENKATA SAI KUMAR Batch ID:** 04072024HYD10AM

**Topic: Introduction to Database**

1. Create a Supermart\_DB with the tables created from the datasets shared (Customer.csv, Sales.csv and Product.csv files)
   1. Create a new database in your database management system, and name it Supermart\_DB.

CREATE DATABASE SUPERMART\_DB;

USE SUPERMART\_DB;

* 1. Create a new table called "customers" in the Supermart\_DB database

CREATE TABLE CUSTOMERS (

CUSTOMER\_ID VARCHAR(100) PRIMARY KEY,

CUSTOMER\_NAME VARCHAR(25),

segment VARCHAR(30),

Age INT,

Country VARCHAR(30),

region VARCHAR(30),

city VARCHAR(30),

state VARCHAR(30),

postal\_code VARCHAR(30)

);

* 1. Load the data from the Customer.csv file into the customers table
  2. Create a new table called "products" in the Supermart\_DB database

CREATE TABLE PRODUCTS (

PRODUCT\_ID VARCHAR(30) PRIMARY KEY,

Category VARCHAR(20),

Sub\_category VARCHAR(30),

product\_name VARCHAR(400)

);

* 1. Load the data from the Product.csv file into the products table
  2. Create a new table called "sales" in the Supermart\_DB database

CREATE TABLE SALES (

order\_line INT,

order\_id VARCHAR(20),

order\_date VARCHAR(30),

ship\_date VARCHAR(30),

ship\_mode VARCHAR(40),

customer\_id VARCHAR(30),

product\_id VARCHAR(30),

sales INT,

quantity INT,

discount INT,

profit INT,

PRIMARY KEY (order\_id, order\_line),

FOREIGN KEY (customer\_id) REFERENCES PRODUCTS(PRODUCT\_ID)

FOREIGN KEY (product\_id) REFERENCES PRODUCTS(PRODUCT\_ID)

);

* 1. Load the data from the Sales.csv file into the sales table

**SELECTION OPERATORS:- (FILTERING):- in, like, between**

**Note:** use products, customers and sales table

1. Define the relationship between the tables using constraints/keys.

Step-by-Step Relationship Definition

Define Primary Keys (PK) for each table.

In CUSTOMERS, the primary key is likely CUSTOMER\_ID.

In PRODUCTS, the primary key is PRODUCT\_ID.

In SALES, there’s no clear single field that acts as a unique identifier for each row. Therefore, a composite primary key of order\_id and order\_line might be appropriate.

Define Foreign Keys (FK) to establish relationships between the tables.

The CUSTOMER\_ID in SALES should reference the CUSTOMER\_ID in CUSTOMERS.

The PRODUCT\_ID in SALES should reference the PRODUCT\_ID in PRODUCTS.

Explanation of the Relationships

**CUSTOMERS → SALES:**

The SALES table has a foreign key CUSTOMER\_ID referencing the CUSTOMER\_ID in the CUSTOMERS table. This creates a one-to-many relationship: one customer can have many sales.

**PRODUCTS → SALES:**

The SALES table also has a foreign key PRODUCT\_ID referencing the PRODUCT\_ID in the PRODUCTS table. This creates another one-to-many relationship: one product can be associated with many sales.

**Primary Keys:**

CUSTOMERS: The primary key is CUSTOMER\_ID, which ensures each customer is unique.

PRODUCTS: The primary key is PRODUCT\_ID, ensuring each product is unique.

SALES: The composite primary key order\_id + order\_line ensures that each combination of order and line is unique within the SALES table.

**Constraints Overview:**

Primary Key on CUSTOMER\_ID and PRODUCT\_ID ensures the uniqueness of each customer and product.

Foreign Key constraints in the SALES table ensure that customer\_id and product\_id values in SALES must exist in the CUSTOMERS and PRODUCTS tables respectively. This ensures referential integrity across the tables.

1. In the database Supermart \_DB, find the following:
2. Get the list of all the cities where the region is north or east without any duplicates using the IN statement.
3. Get the list of all orders where the ‘sales’ value is between 100 and 500 using the BETWEEN operator.
4. Get the list of customers whose last name contains only 4 characters using LIKE.

**SELECTION OPERATORS:- ordering**

1. Retrieve all orders where the ‘discount’ value is greater than zero ordered in descending order basis ‘discount’ value
2. Limit the number of results in the above query to the top 10.

**Aggregate operators:-**

1. Find the sum of all ‘sales’ values.
2. Find count of the number of customers in the north region with ages between 20 and 30
3. Find the average age of east region customers
4. Find the minimum and maximum aged customers from Philadelphia

**GROUP BY OPERATORS:-**

1. Create a display with the information below for each product ID.
2. Total sales (in $) order by this column in descending
3. Total sales quantity
4. The number of orders
5. Max Sales value
6. Min Sales value
7. Average sales value
8. Get the list of product ID’s where the quantity of product sold is greater than 10