Step 1: Create Table

CREATE TABLE walmart\_sales (

User\_ID INT,

Product\_ID VARCHAR(20),

Gender CHAR(1),

Age VARCHAR(10),

Occupation INT,

City\_Category CHAR(1),

Stay\_In\_Current\_City\_Years VARCHAR(5),

Marital\_Status TINYINT,

Product\_Category INT,

Purchase INT

);

INSERT INTO walmart\_sales (User\_ID, Product\_ID, Gender, Age, Occupation, City\_Category, Stay\_In\_Current\_City\_Years, Marital\_Status, Product\_Category, Purchase) VALUES

(1000001, 'P00069042', 'F', '0-17', 10, 'A', '2', 0, 3, 8370),

(1000001, 'P00248942', 'F', '0-17', 10, 'A', '2', 0, 1, 15200),

(1000001, 'P00087842', 'F', '0-17', 10, 'A', '2', 0, 12, 1422),

(1000001, 'P00085442', 'F', '0-17', 10, 'A', '2', 0, 12, 1057),

(1000002, 'P00285442', 'M', '55+', 16, 'C', '4+', 0, 8, 7969),

(1000003, 'P00193542', 'M', '26-35', 15, 'A', '3', 0, 1, 15227),

(1000004, 'P00184942', 'M', '46-50', 7, 'B', '2', 1, 1, 19215),

(1000004, 'P00346142', 'M', '46-50', 7, 'B', '2', 1, 1, 15854),

(1000004, 'P0097242', 'M', '46-50', 7, 'B', '2', 1, 1, 15686),

(1000005, 'P00274942', 'M', '26-35', 20, 'A', '1', 1, 8, 7871),

(1000005, 'P00251242', 'M', '26-35', 20, 'A', '1', 1, 5, 5254),

(1000005, 'P00014542', 'M', '26-35', 20, 'A', '1', 1, 8, 3957),

(1000005, 'P00031342', 'M', '26-35', 20, 'A', '1', 1, 8, 6073),

(1000005, 'P00145042', 'M', '26-35', 20, 'A', '1', 1, 1, 15665),

(1000006, 'P00231342', 'F', '51-55', 9, 'A', '1', 0, 5, 5378),

(1000006, 'P00190242', 'F', '51-55', 9, 'A', '1', 0, 4, 2079),

(1000006, 'P0096642', 'F', '51-55', 9, 'A', '1', 0, 2, 13055),

(1000006, 'P00058442', 'F', '51-55', 9, 'A', '1', 0, 5, 8851),

(1000007, 'P00036842', 'M', '36-45', 1, 'B', '1', 1, 1, 11788),

(1000008, 'P00249542', 'M', '26-35', 12, 'C', '4+', 1, 1, 19614);

## Step 2: Load Data

Assuming you have the CSV file walmart.csv imported into the table walmart\_sales.

## Step 3: SQL Queries for Insights

1. Total Sales by Gender

SELECT Gender, SUM(Purchase) AS Total\_Sales

FROM walmart\_sales

GROUP BY Gender;

## 2. Average Purchase Amount by Age Group

SELECT Age, AVG(Purchase) AS Avg\_Purchase

FROM walmart\_sales

GROUP BY Age

ORDER BY FIELD(Age, '0-17', '18-25', '26-35', '36-45', '46-50', '51-55', '55+');

3. Top 5 Product Categories by Total Sales

SELECT Product\_Category, SUM(Purchase) AS Total\_Sales

FROM walmart\_sales

GROUP BY Product\_Category

ORDER BY Total\_Sales DESC

LIMIT 5;

## 4. Sales Trend by City Category

SELECT City\_Category, SUM(Purchase) AS Total\_Sales

FROM walmart\_sales

GROUP BY City\_Category;

## 5. Effect of Marital Status on Purchase Amount

SELECT Marital\_Status, AVG(Purchase) AS Avg\_Purchase

FROM walmart\_sales

GROUP BY Marital\_Status;

## 6. Purchase Behavior by Length of Stay in Current City

SELECT Stay\_In\_Current\_City\_Years, AVG(Purchase) AS Avg\_Purchase, COUNT(\*) AS Number\_of\_Purchases

FROM walmart\_sales

GROUP BY Stay\_In\_Current\_City\_Years

ORDER BY

CASE

WHEN Stay\_In\_Current\_City\_Years = '0' THEN 0

WHEN Stay\_In\_Current\_City\_Years = '1' THEN 1

WHEN Stay\_In\_Current\_City\_Years = '2' THEN 2

WHEN Stay\_In\_Current\_City\_Years = '3' THEN 3

WHEN Stay\_In\_Current\_City\_Years = '4+' THEN 4

ELSE 5

END;

## 7. Top Occupations by Total Sales

SELECT Occupation, SUM(Purchase) AS Total\_Sales

FROM walmart\_sales

GROUP BY Occupation

ORDER BY Total\_Sales DESC

LIMIT 5;

## 8. Number of Purchases per User and their Total Spending

SELECT User\_ID, COUNT(\*) AS Number\_of\_Purchases, SUM(Purchase) AS Total\_Spent

FROM walmart\_sales

GROUP BY User\_ID

ORDER BY Total\_Spent DESC

LIMIT 10;

## 9. Distribution of Purchases by Product Category and Gender

SELECT Product\_Category, Gender, COUNT(\*) AS Number\_of\_Purchases, SUM(Purchase) AS Total\_Sales

FROM walmart\_sales

GROUP BY Product\_Category, Gender

ORDER BY Product\_Category, Gender;

## 10. Average Purchase by Age and Gender

SELECT Age, Gender, AVG(Purchase) AS Avg\_Purchase

FROM walmart\_sales

GROUP BY Age, Gender

ORDER BY Age, Gender;