USE WALLMART;

CREATE TABLE WALLMART\_ANALYSIS

    (

        invoice\_id VARCHAR(30) NOT NULL PRIMARY KEY,

        branch VARCHAR(5) NOT NULL,

        city VARCHAR(30) NOT NULL,

        customer\_type VARCHAR(30) NOT NULL,

        gender VARCHAR(30) NOT NULL,

        product\_line VARCHAR(100) NOT NULL,

        unit\_price DECIMAL(10, 2) NOT NULL,

        quantity INT NOT NULL,

        tax\_pct FLOAT (6, 4) NOT NULL,

        total DECIMAL(12, 4) NOT NULL,

        date DATETIME NOT NULL,

        time TIME NOT NULL,

        payment VARCHAR(15) NOT NULL,

        cogs DECIMAL(10, 2) NOT NULL,

        gross\_margin\_pct FLOAT (11, 9),

        gross\_income DECIMAL(12, 4),

        rating FLOAT (2, 1));

-- displaying the column names

SELECT \* from wallmart\_analysis;

           -- -- -- -- --  Feature Engineering -- -- -- -- --

-- Time of Day

SELECT

    time,

    CASE

        WHEN `time` >= "00:00:00" AND `time` < "12:00:00" THEN "Morning"

        WHEN `time` >= "12:00:00" AND `time` < "16:00:00" THEN "Afternoon"

        ELSE "Evening"

    END AS time\_of\_day

FROM wallmart\_analysis;--

-- Adding column day name

ALTER TABLE wallmart\_analysis

ADD COLUMN day\_name VARCHAR(10);

UPDATE wallmart\_analysis

SET day\_name = DAYNAME(date);

--  Adding column month name

ALTER TABLE wallmart\_analysis ADD column month\_name VARCHAR(10);

UPDATe wallmart\_analysis SET month\_name =MONTHNAME(DATE);

--  Q1How many unique cities does the data have?

SELECT DISTINCT city FROM wallmart\_analysis;

--  Q2 In which city is each branch?

SELECT DISTINCT city,branch FROM wallmart\_analysis;

-- ---------------------------------------------------------------------

-- ---------------------------- Product --------------------------------

-- ---------------------------------------------------------------------

-- Q3 How many unique product lines does the data have?

SELECT DISTINCT PRODUCT\_LINE FROM WALLMART\_ANALYSIS ;

-- Q4 What is the most common payment method?

SELECT PAYMENT,COUNT(PAYMENT)AS PAYMENT\_COUNT

FROM WALLMART\_ANALYSIS

GROUP BY PAYMENT

ORDER BY PAYMENT\_COUNT DESC;

-- Q5 Which selling product line is the most ?

SELECT PRODUCT\_LINE,COUNT(PRODUCT\_LINE)AS PRODUCT\_LINE\_COUNT

FROM WALLMART\_ANALYSIS

GROUP BY PRODUCT\_LINE

ORDER BY PRODUCT\_LINE\_COUNT DESC;

-- Q6 What is the most selling product line

SELECT

    SUM(quantity) as qty,

    product\_line

FROM WALLMART\_ANALYSIS

GROUP BY product\_line

ORDER BY qty DESC;

-- Q7 What is the total revenue by month?

SELECT month\_name as month , sum(total) as total\_revenue

from wallmart\_analysis group by month order by total\_revenue desc;

-- Q8 What month had the largest COGS?

select month\_name,sum(cogs) as largest\_cogs from wallmart\_analysis

group by month\_name

order by largest\_cogs desc;

-- Q9 What product line had the largest revenue?

select product\_line, sum(total) as total\_Revenue

from wallmart.wallmart\_analysis

group by product\_line

order by  total\_Revenue desc;

-- Q10 What is the city with the largest revenue?

select city , branch , sum(total)as total\_Revenue

from wallmart.wallmart\_analysis

group by city, branch

order by  total\_Revenue desc;

-- Q11 What product line had the largest VAT?

SELECT

    product\_line,

    SUM(0.05 \* cogs) AS total\_vat

FROM wallmart.wallmart\_analysis

GROUP BY product\_line

ORDER BY total\_vat DESC

LIMIT 1;

-- Q12 Fetch each product line and add a column to those product line

-- showing "Good", "Bad". Good if its greater than average sales

select avg(quantity) from wallmart\_analysis;

SELECT

    product\_line,

    CASE

        WHEN AVG(quantity) > 5.49 THEN "Good"

        ELSE "Bad"

    END AS remark

    FROM wallmart.wallmart\_analysis

GROUP BY product\_line;

-- Q13 Which branch sold more products than average product sold

Select branch, sum(quantity) as quantity\_sold from wallmart.wallmart\_analysis

GROUP BY branch

HAVING SUM(quantity) > (SELECT AVG(quantity) FROM wallmart.wallmart\_analysis);

-- Q14 What is the most common product line by gender?

SELECT

    gender,

    product\_line,

    COUNT(gender) AS total\_cnt

FROM wallmart.wallmart\_analysis

GROUP BY gender, product\_line

ORDER BY total\_cnt DESC;

-- Q15 What is the average rating of each product line?

select product\_line , round(avg(rating),2) as avg\_rating

from wallmart.wallmart\_analysis

group by product\_line

order by  avg\_rating;

-- --------------------------------------------------------------------

-- -------------------------- Sales -----------------------------------

-- --------------------------------------------------------------------

-- Q16 Number of sales made in each time of the day per weekday

SELECT

    time\_of\_day,

    COUNT(\*) AS total\_sales

FROM WALLMART.WALLMART\_ANALYSIS

WHERE day\_name = "Sunday"

GROUP BY time\_of\_day

ORDER BY total\_sales DESC;

-- Q17 Which of the customer types brings the most revenue?

SELECT

    customer\_type,

    SUM(TOTAL) AS total\_revenue

FROM

    wallmart.wallmart\_analysis

GROUP BY

    customer\_type

ORDER BY

    total\_revenue DESC;

-- Q18 Which city has the largest tax percent/ VAT (Value Added Tax)?

select

    city,

    ROUND(AVG(tax\_pct), 2) AS avg\_tax\_pct

from wallmart.wallmart\_analysis

group by city order by avg\_tax\_pct desc;

-- Q19 Which customer type pays the most in VAT?

    SELECT

        customer\_type,

        round(Max(tax\_pct),2) AS total\_vat

    FROM

        wallmart.wallmart\_analysis

    GROUP BY

        customer\_type

    ORDER BY

        total\_vat DESC;

-- --------------------------------------------------------------------

-- -------------------------- Customers -------------------------------

-- --------------------------------------------------------------------

-- Q20 How many unique customer types does the data have?

select distinct customer\_type from wallmart.wallmart\_analysis;

-- Q21 How many unique payment methods does the data have?

 select distinct payment as payment\_method from wallmart.wallmart\_analysis;

-- Q22 What is the most common customer type?

select

    customer\_type,

    count(\*) as count

FROM wallmart.wallmart\_analysis

GROUP BY customer\_type

ORDER BY count DESC;

-- Q23 Which customer type buys the most?

SELECT

    customer\_type,

    round(SUM(total),2) AS total\_purchases

FROM

    wallmart.wallmart\_analysis

GROUP BY

    customer\_type

ORDER BY

    total\_purchases DESC;

-- Q24 What is the gender of most of the customers?

SELECT

    gender,

    COUNT(\*) AS gender\_cnt

FROM

    WALLMART.WALLMART\_ANALYSIS

GROUP BY

    gender

ORDER BY

    gender\_cnt DESC;

-- Q25 What is the gender distribution per branch?

select branch ,count(gender) as gendercount

from wallmart.wallmart\_analysis

group by branch

order by gendercount desc;

-- Q26 Which time of the day do customers give most ratings?

SELECT

    time\_of\_day,

    avg(rating) as avg\_rating

FROM wallmart.wallmart\_analysis

GROUP BY time\_of\_day

ORDER BY avg\_rating DESC ;

-- Q27 Which time of the day do customers give most ratings per branch?

SELECT

    time\_of\_day,branch,

    avg(rating) as avg\_rating

FROM wallmart.wallmart\_analysis

GROUP BY time\_of\_day,branch

ORDER BY avg\_rating DESC ;

-- Q28 Which day of the week has the best avg ratings?

SELECT

day\_name,

avg(rating) as avg\_rating

FROM wallmart.wallmart\_analysis

GROUP BY day\_name

ORDER BY avg\_rating DESC ;

-- Q29 Which day of the week has the best average ratings per branch?

select  day\_name,

    COUNT(day\_name) total\_sales

FROM wallmart.wallmart\_analysis

WHERE branch = "C"

GROUP BY day\_name

ORDER BY total\_sales DESC;