**REQUIREMENTS & SOLUTIONS:**

**Generic Question**

1. How many unique cities does the data have?

SELECT distinct(City) FROM walmart\_table;

1. In which city is each branch?

SELECT distinct(branch), city FROM walmart\_table;

**Product**

1. How many unique product lines does the data have?

SELECT count(product\_line) FROM walmart\_table;

1. What is the most common payment method?

SELECT payment, count(payment) FROM walmart\_table

group by payment

order by count(payment) desc

limit 1;

1. What is the most selling product line?

SELECT product\_line, count(product\_line) FROM walmart\_table

group by product\_line

order by count(product\_line) desc

limit 1;

1. What is the total revenue by month?

SELECT month\_name, sum(total) FROM walmart\_table

GROUP BY month\_name;

1. What month had the largest COGS?

SELECT month\_name, sum(cogs) FROM walmart\_table

GROUP BY month\_name

ORDER BY sum(cogs) desc

limit 1;

1. What product line had the largest revenue?

SELECT product\_line, sum(total) FROM walmart\_table

GROUP BY product\_line

ORDER BY sum(total) desc

limit 1;

1. What is the city with the largest revenue?

SELECT city, sum(total) FROM walmart\_table

GROUP BY city

ORDER BY sum(total) desc

limit 1;

1. What product line had the largest VAT?

SELECT product\_line, avg(VAT) FROM walmart\_table

GROUP BY product\_line

ORDER BY avg(VAT) desc

limit 1;

1. Fetch each product line and add a column to those product line showing "Good", "Bad". Good if its greater than average sales
2. Which branch sold more products than average product sold?

SELECT branch, sum(quantity) FROM walmart\_table

GROUP BY branch

HAVING sum(quantity) > (SELECT avg(quantity) FROM walmart\_table)

ORDER BY sum(quantity) desc

limit 1;

1. What is the most common product line by gender?

SELECT gender, product\_line, count(gender) FROM walmart\_table

group by gender, product\_line

order by count(gender) desc;

1. What is the average rating of each product line?

SELECT product\_line, round(avg(rating), 2) FROM walmart\_table

GROUP BY product\_line

ORDER BY avg(rating);

**Sales**

1. Number of sales made in each time of the day per weekday

SELECT time\_of\_day, count(\*) FROM walmart\_table

WHERE day\_name = 'Monday'

GROUP BY time\_of\_day

ORDER BY count(\*) desc;

1. Which of the customer types brings the most revenue?

SELECT customer\_type, sum(total) FROM walmart\_table

GROUP BY customer\_type

ORDER BY sum(total) desc;

1. Which city has the largest tax percent/ VAT (**Value Added Tax**)?

SELECT city, round(avg(VAT), 2) as VAT\_perc FROM walmart\_table

GROUP BY city

ORDER BY avg(VAT) desc;

1. Which customer type pays the most in VAT?

SELECT customer\_type, round(AVG(VAT), 2) FROM walmart\_table

GROUP BY customer\_type

ORDER BY avg(VAT) desc;

**Customer**

1. How many unique customer types does the data have?

SELECT count(DISTINCT(customer\_type)) FROM walmart\_table;

1. How many unique payment methods does the data have?

SELECT count(DISTINCT(payment)) FROM walmart\_table;

1. What is the most common customer type?

SELECT customer\_type, count(customer\_type) FROM walmart\_table

GROUP BY customer\_type

ORDER BY count(customer\_type) desc;

1. Which customer type buys the most?

SELECT customer\_type, count(quantity) FROM walmart\_table

GROUP BY customer\_type

ORDER BY count(quantity) desc

1. What is the gender of most of the customers?

SELECT gender, count(quantity) FROM walmart\_table

GROUP BY gender

ORDER BY count(quantity) desc;

1. What is the gender distribution per branch?

SELECT gender, branch, count(gender) FROM walmart\_table

GROUP BY gender, branch

ORDER BY count(gender);

1. Which time of the day do customers give most ratings?

SELECT time\_of\_day, avg(rating) FROM walmart\_table

GROUP BY time\_of\_day

ORDER BY avg(rating) desc;

1. Which time of the day do customers give most ratings per branch?

SELECT time\_of\_day, branch, avg(rating) FROM walmart\_table

GROUP BY time\_of\_day, branch

ORDER BY avg(rating) desc;

1. Which day of the week has the best avg ratings?

SELECT day\_name, avg(rating) FROM walmart\_table

GROUP BY day\_name

ORDER BY avg(rating) desc;

1. Which day of the week has the best average ratings per branch?

SELECT day\_name, branch, avg(rating) FROM walmart\_table

GROUP BY day\_name, branch

ORDER BY avg(rating) desc;

**TO CREATE TABLE**

create table if not exists walmart\_table(

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

branch VARCHAR(30) 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,

VAT FLOAT(6, 4) NOT NULL,

total DECIMAL(10, 2) NOT NULL,

date DATETIME NOT NULL,

time TIME NOT NULL,

payment VARCHAR(50) NOT NULL,

cogs DECIMAL(10, 2) NOT NULL,

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

gross\_income DECIMAL(10, 2) NOT NULL,

rating FLOAT(2,1)

)

**TO CREATE time\_of\_day COLUMN**

SELECT

time,

(CASE

WHEN time BETWEEN '00:00:00' AND '12:00:00' THEN 'Morning'

WHEN time BETWEEN '12:01:00' AND '16:00:00' THEN 'Afternoon'

ELSE 'Evening'

END) AS time\_of\_day

FROM

walmart\_table;

ALTER TABLE walmart\_table ADD time\_of\_day VARCHAR(20);

UPDATE walmart\_table

SET time\_of\_day = (

CASE

WHEN time BETWEEN '00:00:00' AND '12:00:00' THEN 'Morning'

WHEN time BETWEEN '12:01:00' AND '16:00:00' THEN 'Afternoon'

ELSE 'Evening'

END

);

**TO CREATE day\_name COLUMN**

SELECT date, dayname(date) as day\_name FROM walmart\_table;

ALTER TABLE walmart\_table ADD day\_name VARCHAR(10);

UPDATE walmart\_table

SET day\_name = dayname(date);

**TO CREATE month\_name COLUMN**

SELECT date, monthname(date) FROM walmart\_table;

ALTER TABLE walmart\_table ADD month\_name VARCHAR(10);

UPDATE walmart\_table

SET month\_name = monthname(date);