Walmart Sales Analysis

About Data

The datasets contains sales transactions from a three different branches of Walmart, respectively located in Mandalay, Yongon and Naypyitaw. The data contains 17 columns and 1000 rows.

|  |  |  |
| --- | --- | --- |
| Columns | Description | Data Types |
| Invoice\_id | Invoice of the sales made | VARCHAR(30) |
| Branch | Branch at which sale were made | VARCHAR(5) |
| City | The location of the branch | VARCHAR(30) |
| Customer\_type | The type of the Customer | VARCHAR(30) |
| Gender | Gender of the customer making purchase | VARCHAR(10) |
| Product\_line | Product line of the Product solf | VARCHAR(100) |
| Unit\_price | The price of each Product | DECIMAL(10,2) |
| Quantity | The amount of product sold | INT |
| VAT | The amount of tax on the purchase | FLOAT(6,4) |
| Total | The Total cost of the purchase | DECIMAL(10,2) |
| Date | The date on which the purchase was made | DATE |
| Time | The time at which the purchase was made | TIMESTAMP |
| Payment\_method | The Payment modes | VARCHAR(30) |
| Cogs | Cost of Goods sold | DECIMAL(10,2) |
| Gross\_margin\_Percentage | Gross margin percentage | FLOAT(11,9) |
| Gross\_income | Gross income | DECIMAL(10,2) |
| Rating | Rating | FLOAT(2,1) |

**Required Analysis**

**Product Analysis:**

Conduct analysis on the data to understand the different product lines, identify the best-performing product lines, and determine which product lines need improvement.

**Sales Analysis:**

This analysis aims to answer questions about sales trends of products. The results can help measure the effectiveness of each sales strategy the business applies and identify necessary modifications to increase sales.

**Customer Analysis:**

This analysis aims to uncover different customer segments, purchase trends, and the profitability of each customer segment.

Adding the columns

1. time\_of\_day (Morning, Afternoon, Evening)

UPDATE sales

SET

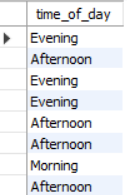
time\_of\_day = (CASE

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);

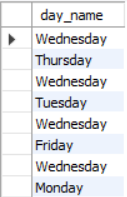


1. day\_name (Sunday, Monday, Tuesday … )

UPDATE sales

SET

day\_name = DAYNAME(`date`);



1. month\_name (January, February, March, April …)

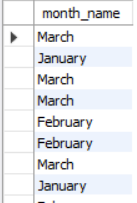
UPDATE sales

SET

`Month` = MONTHNAME(`date`);

ALTER TABLE sales

RENAME COLUMN `Month` to month\_name;



Business Questions to Answer

# Generic question

Q) How many unique cities does the data have ?

SELECT

COUNT(DISTINCT city) AS unique\_city\_count

FROM

sales;



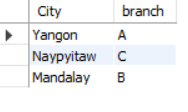
Q) In which city is each branch ?

SELECT

DISTINCT City, branch

FROM

sales;



# Product Questions

Q) How many unique product lines does the data have ?

SELECT

COUNT(DISTINCT product\_line) AS unique\_product\_line\_count

FROM

sales;



Q) What is the Most Common Payment method ?

SELECT

payment\_method, COUNT(payment\_method) AS use\_method

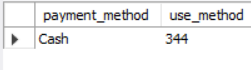
FROM

sales

GROUP BY payment\_method

ORDER BY use\_method DESC

LIMIT 1;



Q) Which is the most selling product line ?

SELECT

product\_line, MAX(total) AS max\_sales

FROM

sales

GROUP BY product\_line

ORDER BY max\_sales DESC

LIMIT 1;



Q) What is the Total Revenue by Month ?

SELECT

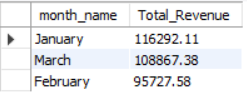
month\_name, SUM(total) AS Total\_Revenue

FROM

sales

GROUP BY month\_name

ORDER BY Total\_Revenue DESC;



Q) Which month had the largest COGS ?

SELECT

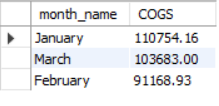
month\_name, SUM(cogs) AS COGS

FROM

sales

GROUP BY month\_name

ORDER BY COGS DESC;



Q) Which product line had the largest revenue ?

SELECT

product\_line, SUM(total) AS Total\_rev

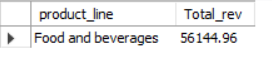
FROM

sales

GROUP BY product\_line

ORDER BY Total\_rev DESC

LIMIT 1;



Q) Which is the city with Maximum Revenue ?

SELECT

city, SUM(total) AS Max\_Rev

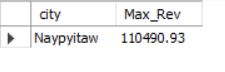
FROM

sales

GROUP BY city

ORDER BY Max\_Rev DESC

LIMIT 1;



Q) Which Product line had largest VAT ?

SELECT

product\_line, SUM(VAT) AS Max\_VAT

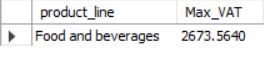
FROM

sales

GROUP BY product\_line

ORDER BY Max\_VAT DESC

LIMIT 1;



Q) Which branch sold more than average product sold ?

SELECT

branch, SUM(quantity) AS Total\_sold\_product

FROM

sales

GROUP BY branch

HAVING SUM(quantity) > (SELECT

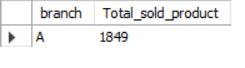
AVG(quantity)

FROM

sales)

ORDER BY Total\_sold\_product DESC

LIMIT 1;



Q) What is the most common product line by gender ?

SELECT

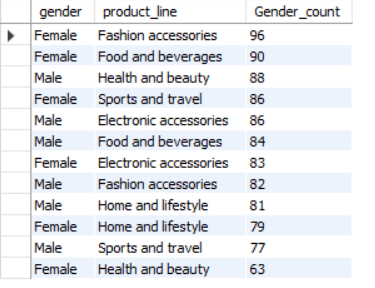
gender, product\_line, COUNT(gender) AS Gender\_count

FROM

sales

GROUP BY gender , product\_line

ORDER BY Gender\_count DESC;



Q) What is the Average rating of each product line ?

SELECT

product\_line, AVG(rating) AS Avg\_rating

FROM

sales

GROUP BY product\_line;



# Sales Question

Q) Number of sales made in each time of the day per weekday ?

SELECT

time\_of\_day, COUNT(invoice\_id) AS sales\_count

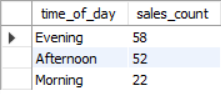
FROM

sales

WHERE day\_name = 'Sunday'

GROUP BY time\_of\_day

ORDER BY sales\_count DESC;



Q) Which of the Customer types bring the most revenue ?

SELECT

customer\_type, SUM(total) AS revenue

FROM

sales

GROUP BY customer\_type

ORDER BY Revenue DESC

LIMIT 1;



Q) Which city has the largest tax percentage / VAT (Value Added Tax) ?

SELECT

city, AVG(VAT) AS tax\_percentage

FROM

sales

GROUP BY city

ORDER BY tax\_percentage DESC

LIMIT 1;



Q) Which customer type pays the most in VAT ?

SELECT

customer\_type, SUM(VAT) AS total\_vat

FROM

sales

GROUP BY customer\_type

ORDER BY total\_vat DESC

LIMIT 1;



# Customer Questions

Q) How many the unique customer types does the data have ?

SELECT

COUNT(DISTINCT customer\_type) AS cust\_type\_cnt

FROM

sales;



Q) How Many unique payment methods does the data have ?

SELECT

COUNT(DISTINCT payment\_method) cnt\_of\_unique\_method

FROM

sales;



Q) Which Customer types buys the most ?

SELECT

customer\_type, SUM(total) AS total\_sales

FROM

sales

GROUP BY customer\_type

ORDER BY total\_sales DESC

LIMIT 1;



Q) What is the gender of most of the customer ?

SELECT

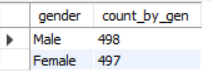
gender, COUNT(\*) AS count\_by\_gen

FROM

sales

GROUP BY gender

ORDER BY count\_by\_gen DESC;



Q) What is the gender distribution per branch ?

SELECT

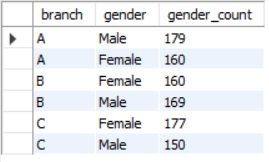
branch, gender, COUNT(\*) AS gender\_count

FROM

sales

GROUP BY gender , branch

ORDER BY branch;



Q) Which time of the day do customer give most rating ?

SELECT

time\_of\_day, COUNT(rating) AS rating\_count

FROM

sales

GROUP BY time\_of\_day

ORDER BY rating\_count DESC

LIMIT 1;



Q) Which time of the day do customer give most rating per branch ?

SELECT

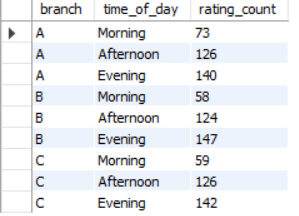
branch, time\_of\_day, COUNT(rating) AS rating\_count

FROM

sales

GROUP BY branch , time\_of\_day

ORDER BY branch , rating\_count;



Q) Which day of the week has the best avg rating ?

SELECT

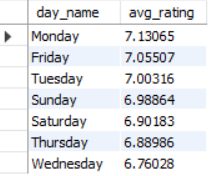
day\_name, AVG(rating) AS avg\_rating

FROM

sales

GROUP BY day\_name

ORDER BY avg\_rating DESC;



Q) Which day of the week has the best average rating per branch ?

SELECT

branch,

DAYNAME(`date`),

AVG(rating) AS avg\_rating,

WEEKDAY(`date`) AS Day\_num

FROM

sales

GROUP BY branch , DAYNAME(`date`) , WEEKDAY(`date`)

ORDER BY branch , WEEKDAY(`date`);

