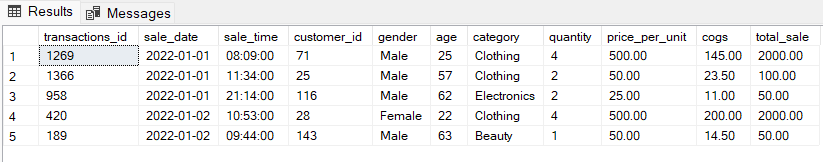
**DATA ANALYSIS PROJECT – SQL**

**Retrieve all columns from the Sales\_Analysis dataset in ascending order by sale\_date.**

SELECT TOP(5) \* FROM Sales\_Analysis ORDER BY sale\_date ASC;

****

**------------------------DATA** **CLEANING----------------------------**

**Modify the cogs column in the dataset to display two decimal places.**

ALTER TABLE Sales\_Analysis ALTER COLUMN cogs DECIMAL(10, 2);

**Modify the price\_per\_unit column in the dataset to display two decimal places.**

ALTER TABLE Sales\_Analysis ALTER COLUMN price\_per\_unit DECIMAL(10, 2);

**Modify the total\_sale column in the dataset to display two decimal places.**

ALTER TABLE Sales\_Analysis ALTER COLUMN total\_sale DECIMAL(10, 2);

**Modify the sale\_time column in the dataset to display hh:mm:ss format.**

ALTER TABLE Sales\_Analysis ALTER COLUMN sale\_time TIME(0);

**Check if there are any NULL values in the dataset.**

SELECT \* FROM Sales\_Analysis

WHERE transactions\_id IS NULL

OR sale\_date IS NULL

OR sale\_time IS NULL

OR customer\_id IS NULL

OR gender IS NULL

OR age IS NULL

OR category IS NULL

OR quantity IS NULL

OR price\_per\_unit IS NULL

OR cogs IS NULL

OR total\_sale IS NULL;

**Delete rows with NULL values from the dataset.**

DELETE FROM Sales\_Analysis

WHERE transactions\_id IS NULL

OR sale\_date IS NULL

OR sale\_time IS NULL

OR customer\_id IS NULL

OR gender IS NULL

OR age IS NULL

OR category IS NULL

OR quantity IS NULL

OR price\_per\_unit IS NULL

OR cogs IS NULL

OR total\_sale IS NULL;

**Rename the column ‘quantiy’ to ‘quantity’.**

EXEC sp\_rename 'Sales\_Analysis.quantiy', 'quantity', 'COLUMN';

**--------------------------DATA** **ANALYSIS---------------------------**

**Display the total number of transactions.**

SELECT COUNT(\*) AS 'Total Transactions' FROM Sales\_Analysis;

A screenshot of a message

Description automatically generated

**Retrieve unique values from the category column.**

SELECT DISTINCT category AS 'Category' FROM Sales\_Analysis;

A screenshot of a computer

Description automatically generated

**Retrieve the ‘Total quantity’ and ‘Total sale’ from the dataset.**

SELECT COUNT(quantity) AS 'Total Quantity', SUM(total\_sale) AS 'Total Sale' FROM Sales\_Analysis;

A screenshot of a message

Description automatically generated

**Retrieve ‘Category’, ‘Gender’ and ‘the number of transactions’, grouped by ‘Category’ and ‘Gender’**

SELECT category AS 'Category', gender AS 'Gender', COUNT(transactions\_id) AS 'Transactions' FROM Sales\_Analysis

GROUP BY category, gender

ORDER BY COUNT(transactions\_id) DESC;

A screenshot of a computer

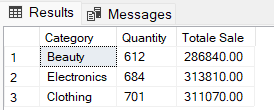
Description automatically generated

**Retrieve ‘Category’, ‘Quantity’ and ‘the sum of total sales’, grouped by ‘Category’.**

SELECT category AS 'Category' ,COUNT(quantity) AS 'Quantity', SUM(total\_sale) AS 'Totale Sale' FROM Sales\_Analysis

GROUP BY category

ORDER BY COUNT(quantity);



**Retrieve the ‘Year’ and the count of ‘Transactions’ from the dataset.**

SELECT YEAR(sale\_date) AS 'Year', COUNT(transactions\_id) AS 'Transactions' FROM Sales\_Analysis

GROUP BY YEAR(sale\_date)

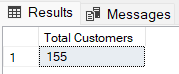
ORDER BY YEAR(sale\_date);

A screenshot of a computer

Description automatically generated

**Retrieve the total number of distinct customers.**

SELECT COUNT(DISTINCT customer\_id) AS 'Total Customers' FROM Sales\_Analysis;

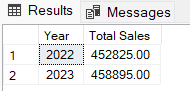


**Retrieve ‘Year’ and the sum of ‘Total sales’, grouped by ‘Year’.**

SELECT YEAR(sale\_date) AS 'Year', SUM(total\_sale) AS 'Total Sales' FROM Sales\_Analysis

GROUP BY YEAR(sale\_date)

ORDER BY YEAR(sale\_date) ASC;



**Compare the total sales between 2022 and 2023 for each month.**

SELECT MONTH(sale\_date) AS 'Month',

COALESCE(SUM(CASE WHEN YEAR(sale\_date) = 2022 THEN total\_sale END), 0) AS 'Total Sales 2022',

COALESCE(SUM(CASE WHEN YEAR(sale\_date) = 2023 THEN total\_sale END), 0) AS 'Total Sales 2023'

FROM Sales\_Analysis

WHERE YEAR(sale\_date) IN (2022, 2023)

GROUP BY MONTH(sale\_date)

ORDER BY MONTH(sale\_date);

A screenshot of a data

Description automatically generated

**Compare the average sales between 2022 and 2023 for each month.**

SELECT

MONTH(sale\_date) AS 'Month',

COALESCE(CAST(AVG(CASE WHEN YEAR(sale\_date) = 2022 THEN total\_sale END) AS DECIMAL(10, 2)), 0.00) AS 'Avg Sales 2022',

COALESCE(CAST(AVG(CASE WHEN YEAR(sale\_date) = 2023 THEN total\_sale END) AS DECIMAL(10, 2)), 0.00) AS 'Avg Sales 2023'

FROM Sales\_Analysis

WHERE YEAR(sale\_date) IN (2022, 2023)

GROUP BY MONTH(sale\_date)

ORDER BY MONTH(sale\_date);

A screenshot of a data

Description automatically generated

**Retrieve ‘Age group’, ‘Quantity’ and ‘Total sales’, grouped by ‘Age group’.**

SELECT

CASE

WHEN Age BETWEEN 18 AND 23 THEN '18-23'

WHEN Age BETWEEN 24 AND 29 THEN '24-29'

WHEN Age BETWEEN 30 AND 35 THEN '30-35'

WHEN Age BETWEEN 36 AND 41 THEN '36-41'

WHEN Age BETWEEN 42 AND 47 THEN '42-47'

WHEN Age BETWEEN 48 AND 60 THEN '48-60'

ELSE '60+'

END AS 'Age group', SUM(quantity) AS 'Quantity', SUM(total\_sale) AS 'Total Sales'

FROM Sales\_Analysis

GROUP BY

CASE

WHEN Age BETWEEN 18 AND 23 THEN '18-23'

WHEN Age BETWEEN 24 AND 29 THEN '24-29'

WHEN Age BETWEEN 30 AND 35 THEN '30-35'

WHEN Age BETWEEN 36 AND 41 THEN '36-41'

WHEN Age BETWEEN 42 AND 47 THEN '42-47'

WHEN Age BETWEEN 48 AND 60 THEN '48-60'

ELSE '60+'

END

ORDER BY 'Age group';

A screenshot of a computer screen

Description automatically generated

**Retrieve the ‘Time period’ and the number of ‘Transactions’, grouped by ‘Time period’.**

SELECT

CASE

WHEN DATEPART(HOUR, sale\_time) < 12 THEN 'Morning'

WHEN DATEPART(HOUR, sale\_time) BETWEEN 12 AND 17 THEN 'Afternoon'

WHEN DATEPART(HOUR, sale\_time) BETWEEN 18 AND 21 THEN 'Evening'

ELSE 'Night'

END AS 'Time Period', COUNT(transactions\_id) AS 'Transactions'

FROM Sales\_Analysis

GROUP BY

CASE

WHEN DATEPART(HOUR, sale\_time) < 12 THEN 'Morning'

WHEN DATEPART(HOUR, sale\_time) BETWEEN 12 AND 17 THEN 'Afternoon'

WHEN DATEPART(HOUR, sale\_time) BETWEEN 18 AND 21 THEN 'Evening'

ELSE 'Night'

END

ORDER BY 'Transactions' DESC;

A screenshot of a message

Description automatically generated