SQL Retail Sales Analysis

**--Create database**

CREATE DATABASE project;

Use project;

**-- Create TABLE**

DROP TABLE IF EXISTS retail\_sale\_analysis;

CREATE TABLE retail\_sale\_analysis

(

transactions\_id INT PRIMARY KEY,

sale\_date DATE,

sale\_time TIME,

customer\_id INT,

gender VARCHAR(15),

age INT,

category VARCHAR(15),

quantity INT,

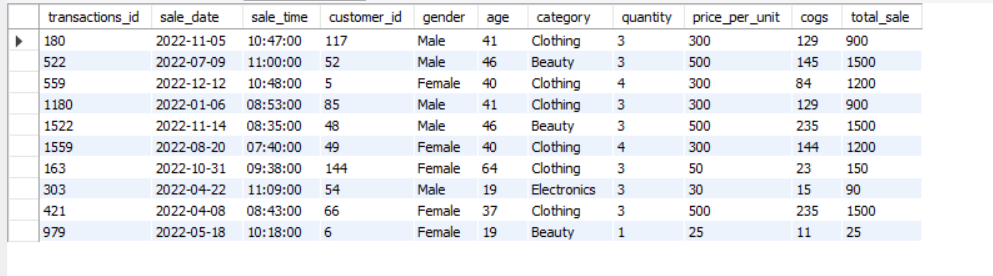
price\_per\_unit FLOAT,

cogs FLOAT,

total\_sale FLOAT

);

SELECT \* FROM retail\_sale\_analysis LIMIT 10;



SELECT COUNT(\*) FROM retail\_sale\_analysis;



**-- Data Cleaning : Finding Null values**

SELECT \* FROM retail\_sales\_analysis

WHERE

transactions\_id IS NULL

OR

sale\_date IS NULL

OR

sale\_time IS NULL

OR

gender IS NULL

OR

category IS NULL

OR

quantity IS NULL

OR

cogs IS NULL

OR

total\_sale IS NULL;

**Deleting Null values:**

DELETE FROM retail\_sale\_analysis

WHERE

transactions\_id IS NULL

OR

sale\_date IS NULL

OR

sale\_time IS NULL

OR

gender IS NULL

OR

category IS NULL

OR

quantity IS NULL

OR

cogs IS NULL

OR

total\_sale IS NULL;

**-- Data Exploration**

**-- How many sales we have?**

SELECT COUNT(\*) as total\_sale FROM retail\_sale\_analysis;

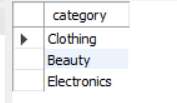


**-- How many unique customers we have ?**

SELECT COUNT(DISTINCT customer\_id) as total\_sale FROM retail\_sale\_analaysis;



SELECT DISTINCT category FROM retail\_sale\_analysis;



**-- Data Analysis & Business Key Problems & Answers**

**-- Analysis & Findings**

-- Q.1 Write a SQL query to retrieve all columns for sales made on '2023-10-07

-- Q.2 Write a SQL query to calculate the total sales (total\_sale) for each category.

-- Q.3 Write a SQL query to find the average age of customers who purchased items from the 'Beauty' category.

-- Q.4 Write a SQL query to find all transactions where the total\_sale is greater than 1000.

-- Q.5 Write a SQL query to find the total number of transactions (transaction\_id) made by each gender in each category.

-- Q.6 Write a SQL query to calculate the average sale for each month. Find out best selling month in each year

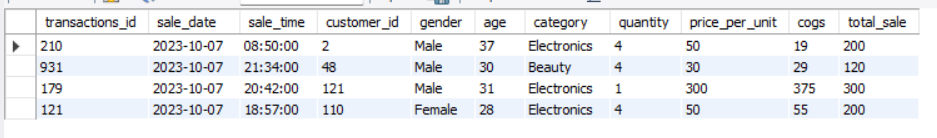
-- Q.7 Write a SQL query to find the top 5 customers based on the highest total sales

-- Q.8 Write a SQL query to find the number of unique customers who purchased items from each category.

-- Q.9 Write a SQL query to create each shift and number of orders (Example Morning <=12, Afternoon Between 12 & 17, Evening >17)

-- Q.1 Write a SQL query to retrieve all columns for sales made on '2023-10-07

**SELECT \* FROM retail\_sale\_analysis WHERE sale\_date = '2023-10-07';**



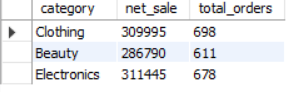
-- Q.2 Write a SQL query to calculate the total sales (total\_sale) for each category.

**SELECT category,**

**SUM(total\_sale) as net\_sale,**

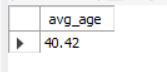
**COUNT(\*) as total\_orders**

**FROM retail\_sale\_analysis GROUP BY 1;**



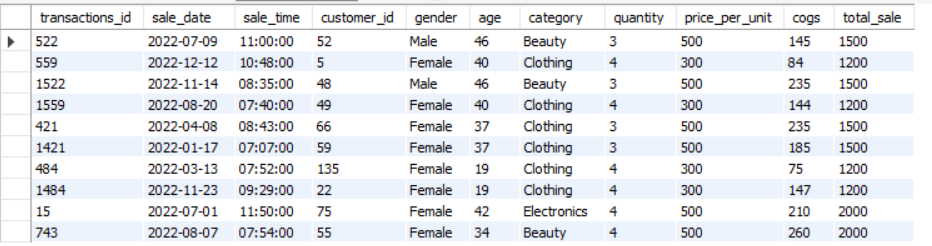
-- Q.3 Write a SQL query to find the average age of customers who purchased items from the 'Beauty' category.

**SELECT ROUND(AVG(age), 2) as avg\_age FROM retail\_sale\_analysis WHERE category = 'Beauty';**



-- Q.4 Write a SQL query to find all transactions where the total\_sale is greater than 1000.

**SELECT \* FROM retail\_sale\_analysis WHERE total\_sale > 1000 LIMIT 10;**



-- Q.5 Write a SQL query to find the total number of transactions (transaction\_id) made by each gender in each category.

**SELECT category, gender, COUNT(\*) as total\_trans**

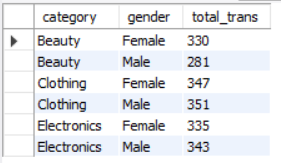
**FROM retail\_sales**

**GROUP BY**

**category,**

**gender**

**ORDER BY 1**



-- Q.6 Write a SQL query to calculate the average sale for each month. Find out best selling month in each year

**SELECT year, month,avg\_sale FROM (**

**SELECT**

**EXTRACT(YEAR FROM sale\_date) as year,**

**EXTRACT(MONTH FROM sale\_date) as month,**

**AVG(total\_sale) as avg\_sale,**

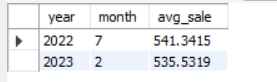
**RANK() OVER(PARTITION BY EXTRACT(YEAR FROM sale\_date) ORDER BY AVG(total\_sale) DESC) as \_rank**

**FROM retail\_sale\_analysis**

**GROUP BY 1, 2**

**) as t1**

**WHERE \_rank = 1;**



-- Q.7 Write a SQL query to find the top 5 customers based on the highest total sales

**SELECT**

**customer\_id,**

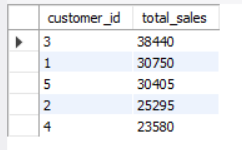
**SUM(total\_sale) as total\_sale\_analysis**

**FROM retail\_sale\_aanlysis**

**GROUP BY 1**

**ORDER BY 2 DESC**

**LIMIT 5;**

****

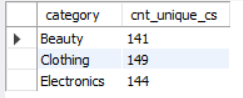
-- Q.8 Write a SQL query to find the number of unique customers who purchased items from each category.

**SELECT category,**

**COUNT(DISTINCT customer\_id) as cnt\_unique\_cs**

**FROM retail\_sale\_analaysis**

**GROUP BY category;**

****

-- Q.9 Write a SQL query to create each shift and number of orders (Example Morning <12, Afternoon Between 12 & 17, Evening >17)

**WITH hourly\_sale**

**AS**

**(**

**SELECT \*,**

**CASE**

**WHEN EXTRACT(HOUR FROM sale\_time) < 12 THEN 'Morning'**

**WHEN EXTRACT(HOUR FROM sale\_time) BETWEEN 12 AND 17 THEN 'Afternoon'**

**ELSE 'Evening'**

**END as shift**

**FROM retail\_sales**

**)**

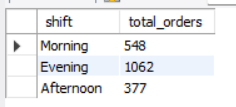
**SELECT**

**shift,**

**COUNT(\*) as total\_orders**

**FROM hourly\_sale**

**GROUP BY shift;**

****

-- End of project