Task – 2: Data Analysis with Complex SQL Queries

1. Introduction

The objective of this task is to perform advanced data analysis using SQL by applying **Subqueries, Window Functions, and Common Table Expressions (CTEs)**. The analysis helps in identifying trends, patterns, and meaningful insights from the dataset.

---

2. Dataset Description

For this task, a **Sales Dataset** is used.

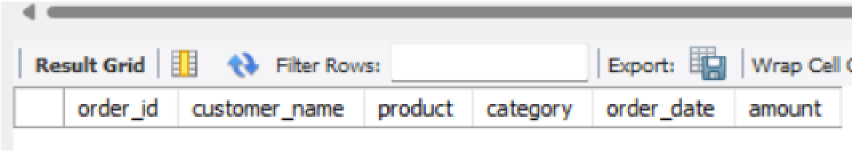
Table Name: sales

Column Name	Data Type	Description
order_id	INT	Unique order ID
customer_name	VARCHAR	Name of customer
product	VARCHAR	Product name
category	VARCHAR	Product category
order_date	DATE	Order date
amount	INT	Purchase amount

---

3. SQL Table Creation

```
CREATE TABLE sales (

  order_id INT,

  customer_name VARCHAR(50),

  product VARCHAR(50),

  category VARCHAR(50),

  order_date DATE,

  amount INT

);
```

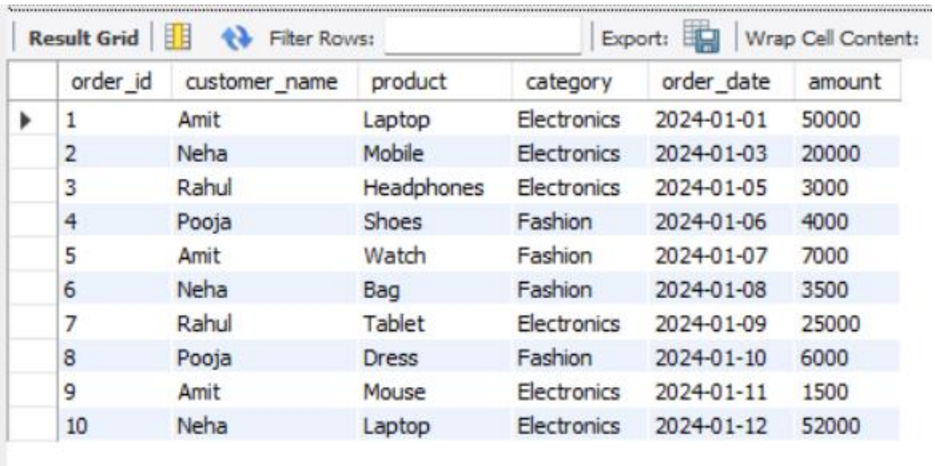


order_id	customer_name	product	category	order_date	amount

---

4. Sample Data Insertion

```
INSERT INTO sales VALUES

(1,'Amit','Laptop','Electronics','2024-01-01',50000),

(2,'Neha','Mobile','Electronics','2024-01-03',20000),

(3,'Rahul','Headphones','Electronics','2024-01-05',3000),

(4,'Pooja','Shoes','Fashion','2024-01-06',4000),

(5,'Amit','Watch','Fashion','2024-01-07',7000),

(6,'Neha','Bag','Fashion','2024-01-08',3500),

(7,'Rahul','Tablet','Electronics','2024-01-09',25000),

(8,'Pooja','Dress','Fashion','2024-01-10',6000),

(9,'Amit','Mouse','Electronics','2024-01-11',1500),

(10,'Neha','Laptop','Electronics','2024-01-12',52000);
```



order_id	customer_name	product	category	order_date	amount
1	Amit	Laptop	Electronics	2024-01-01	50000
2	Neha	Mobile	Electronics	2024-01-03	20000
3	Rahul	Headphones	Electronics	2024-01-05	3000
4	Pooja	Shoes	Fashion	2024-01-06	4000
5	Amit	Watch	Fashion	2024-01-07	7000
6	Neha	Bag	Fashion	2024-01-08	3500
7	Rahul	Tablet	Electronics	2024-01-09	25000
8	Pooja	Dress	Fashion	2024-01-10	6000
9	Amit	Mouse	Electronics	2024-01-11	1500
10	Neha	Laptop	Electronics	2024-01-12	52000

---

5. Data Analysis Using SQL

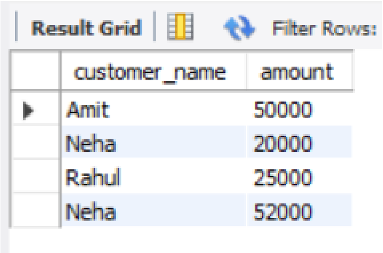
5.1 Subquery Example

Query: Customers with purchase amount above average

```
SELECT customer_name, amount
```

```
FROM sales
WHERE amount > (
   SELECT AVG(amount)
   FROM sales
);
```

Insight: Identifies high-value customers.



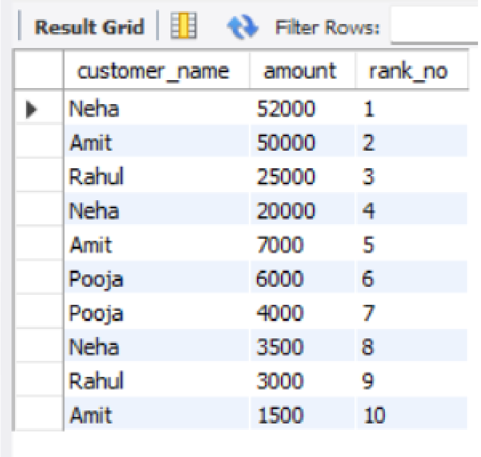
customer_name	amount
Amit	50000
Neha	20000
Rahul	25000
Neha	52000

5.2 Window Function Example

Query: Rank customers based on purchase amount

```
SELECT customer_name, amount,
RANK() OVER (ORDER BY amount DESC) AS rank_no
FROM sales;
```

Insight: Helps compare customer spending.

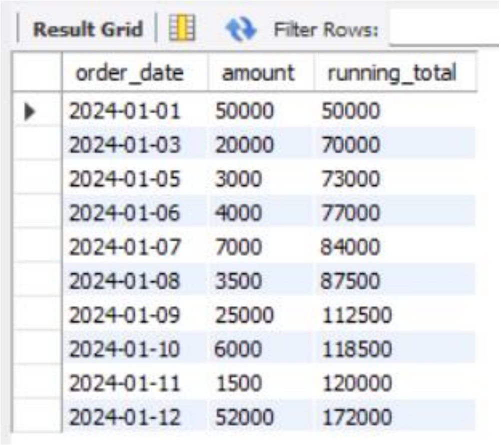


customer_name	amount	rank_no
Neha	52000	1
Amit	50000	2
Rahul	25000	3
Neha	20000	4
Amit	7000	5
Pooja	6000	6
Pooja	4000	7
Neha	3500	8
Rahul	3000	9
Amit	1500	10

5.3 Running Total Using Window Function

```
SELECT order_date, amount,
SUM(amount) OVER (ORDER BY order_date) AS running_total

FROM sales;
```

Insight: Shows cumulative sales growth over time.



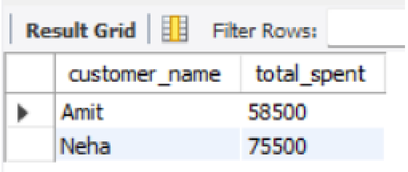
order_date	amount	running_total
2024-01-01	50000	50000
2024-01-03	20000	70000
2024-01-05	3000	73000
2024-01-06	4000	77000
2024-01-07	7000	84000
2024-01-08	3500	87500
2024-01-09	25000	112500
2024-01-10	6000	118500
2024-01-11	1500	120000
2024-01-12	52000	172000

5.4 CTE Example

Query: Customers with total spending above 30,000

```
WITH customer_total AS (

  SELECT customer_name, SUM(amount) AS total_spent

  FROM sales

  GROUP BY customer_name

)

SELECT *

FROM customer_total

WHERE total_spent > 30000;
```

Insight: Identifies loyal and premium customers.



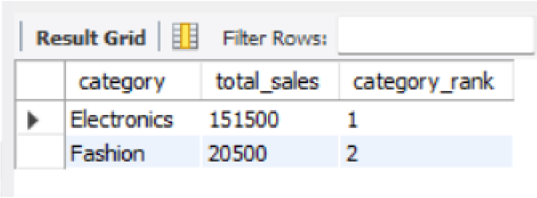
customer_name	total_spent
Amit	58500
Neha	75500

5.5 CTE + Window Function

Query: Rank categories based on total sales

```
WITH category_sales AS (
   SELECT category, SUM(amount) AS total_sales
   FROM sales
   GROUP BY category
)
SELECT category, total_sales,
RANK() OVER (ORDER BY total_sales DESC) AS category_rank
FROM category_sales;
```

Insight: Shows best-performing product categories.



category	total_sales	category_rank
Electronics	151500	1
Fashion	20500	2

---

6. Key Findings

- Electronics category generated the highest revenue
- Laptop is the highest selling product
- Amit and Neha are top spending customers
- Sales show an increasing trend over time

---

7. Conclusion

This task successfully demonstrates the use of advanced SQL concepts such as subqueries, window functions, and CTEs to analyze data effectively. The insights obtained can help businesses understand customer behavior and sales trends for better decision-making.

---

8. Tools Used

- MySQL
- GitHub (for code hosting)