Sales Data Task

The sales dataset contains information on sales transactions for XYZ company over a year. The dataset includes attributes such as Customer Name, product ID, Product category, quantity, and Sales. The company needs to analyze this dataset to identify the factors that are impacting its sales performance.

Note: The granularity of the data is by Order, so the one customer could be repeated for different orders.

Before working on the task:

**Step-1:** Create a Database called ‘SalesDB’.

CREATE DATABASE SalesDB;

**Step-2:** Query on the ‘SalesDB’ database and run the attached script to create the dataset.

Done

Use Window Functions to answer the following questions:

**Question-1:** Find the top 10 customers based on total sales.

-- Using the Window

SELECT DISTINCT TOP 10

CustomerName,

SUM(Sales) OVER (PARTITION BY CustomerName) AS TotalSales

FROM Sales

ORDER BY TotalSales DESC;

-- Using the Group by

SELECT DISTINCT TOP 10

CustomerName,

SUM(Sales) as TotalSales

FROM Sales

GROUP BY CustomerName

ORDER BY TotalSales DESC;

**Question-2:** How much is the total sales of the highest 20% of customers in our company?

SELECT SUM(TotalSales) as SumSales

FROM

(

-- the window here will only select the top 20 percent, and sum of sales

-- for each custName and then order it by sales

SELECT DISTINCT TOP 20 PERCENT CustomerName,

SUM(Sales) OVER (PARTITION BY CustomerName ORDER BY Sales) AS TotalSales

FROM Sales

) AS TOP20;

**Question-3:** We need to categorize our customers based on how much sales they make, divide the customers into 3 equal groups, and give them labels from lowest to highest sales as follows

● Customers Needing Attention

● Promising Customers

● High-Value Customers

WITH CTE AS (

SELECT

CustomerName,

SUM(Sales) AS TotalSales,

NTILE(3) OVER (ORDER BY SUM(Sales) DESC) AS SalesGroup

FROM Sales

GROUP BY CustomerName

)

SELECT

CustomerName,

TotalSales,

SalesGroup,

CASE

WHEN SalesGroup = 1 THEN 'High-Value Customers'

WHEN SalesGroup = 2 THEN 'Promising Customers'

WHEN SalesGroup = 3 THEN 'Customers Needing Attention'

END AS CustomerCategory

FROM CTE;

**Question-4:** The company wants to analyze the purchase behavior of its customers over time to identify trends and patterns.

The company wants to answer one question: "For each customer, what was their first and last purchase date?"

WITH CTE AS (

SELECT

CustomerName,

MIN(OrderDate) OVER (PARTITION BY CustomerName) AS FirstPurchaseDate,

MAX(OrderDate) OVER (PARTITION BY CustomerName) AS LastPurchaseDate

FROM Sales

)

SELECT DISTINCT

CustomerName,

FirstPurchaseDate,

LastPurchaseDate

FROM CTE;

Answer the following questions using simple queries:

**Question-1:** Find customers whose names start with 'A' and handle NULLs in Sales.

SELECT

CustomerName,

COALESCE(Sales, 0) AS Sales

FROM Sales

WHERE CustomerName LIKE 'A%';

**Question-2:** Categorize orders based on the total sales into “small” and “large” transactions.

SELECT

Sales,

Quantity,

CASE

WHEN Sales < 1000 THEN 'small'

ELSE 'large'

END AS TransactionSize

FROM Sales;

**Question-3:** Find unique categories of products bought by customers named 'Rick Hansen' and not by 'Mick Brown'.

SELECT DISTINCT

Category, CustomerName

FROM Sales

WHERE CustomerName = 'Rick Hansen'