Ibrahim Khalid Ibrahim

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;

USE SalesDB;

CREATE TABLE Sales (

OrderID INT PRIMARY KEY IDENTITY(1,1),

CustomerName VARCHAR(100),

ProductID INT,

ProductCategory VARCHAR(50),

Quantity INT,

Sales DECIMAL(10, 2),

OrderDate DATE

);

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

INSERT INTO Sales (CustomerName, ProductID, ProductCategory, Quantity, Sales, OrderDate) VALUES

('Ibrahim Khalid', 10, 'Electronics', 2, 300.00, '2023-01-15'),

('Amr Ali', 12, 'Furniture', 1, 450.00, '2023-01-20'),

('Ibrahim Khalid', 13, 'Books', 3, 45.00, '2023-02-10'),

('Nada Salah', 10, 'Electronics', 1, 150.00, '2023-03-05'),

('Ahmed Mohamed', 14, 'Clothing', 4, 200.00, '2023-04-12'),

('Amr Ali', 15, 'Kitchen', 2, 80.00, '2023-05-15'),

('Ibrahim Khalid', 16, 'Books', 5, 75.00, '2023-06-20'),

('Nada Salah', 17, 'Toys', 1, 20.00, '2023-07-25'),

('Ahmed Mohamed', 18, 'Gardening', 3, 90.00, '2023-08-30'),

('Rick Hansen', 12 ,'Furniture', 2, 1000.00, '2023-09-21'),

('Rick Hansen', 17 ,'Toys', 2, 500.00, '2023-11-2'),

('Mick Brown', 12 ,'Furniture', 1, 450.00, '2023-2-10'),

('Mick Brown', 18,'Gardening', 1, 720.00, '2023-12-10');

Use Window Functions to answer the following questions:

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

WITH CustomerSales AS (

SELECT

CustomerName,

SUM(Sales) AS TotalSales

FROM Sales

GROUP BY CustomerName

),

RankedCustomerSales AS (

SELECT

CustomerName,

TotalSales,

ROW\_NUMBER() OVER (ORDER BY TotalSales DESC) AS SalesRank

FROM CustomerSales

)

SELECT CustomerName, TotalSales

FROM RankedCustomerSales

WHERE SalesRank <= 10

ORDER BY TotalSales DESC;

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

WITH CustomerSales AS (

SELECT

CustomerName,

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

FROM Sales

),

RankedCustomerSales AS (

SELECT

CustomerName,

TotalSales,

PERCENT\_RANK() OVER (ORDER BY TotalSales DESC) AS SalesPercentRank

FROM CustomerSales

)

SELECT SUM(TotalSales) AS Top20PercentSales

FROM RankedCustomerSales

WHERE SalesPercentRank <= 0.20;

**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 CustomerSales AS (

SELECT

CustomerName,

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

FROM Sales

),

RankedCustomerSales AS (

SELECT

CustomerName,

TotalSales,

NTILE(3) OVER (ORDER BY TotalSales DESC) AS SalesGroup

FROM CustomerSales

)

SELECT

CustomerName,

TotalSales,

CASE SalesGroup

WHEN 1 THEN 'Customers Needing Attention'

WHEN 2 THEN 'Promising Customers'

WHEN 3 THEN 'High-Value Customers'

END AS CustomerCategory

FROM RankedCustomerSales

ORDER BY TotalSales DESC;

**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?"

SELECT DISTINCT

CustomerName,

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

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

FROM Sales

ORDER BY CustomerName;

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

OrderID,

CustomerName,

Sales,

CASE

WHEN Sales < 100 THEN 'small'

ELSE 'large'

END AS TransactionCategory

FROM Sales;

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

SELECT DISTINCT ProductCategory

FROM Sales

WHERE CustomerName = 'Rick Hansen'

AND ProductCategory NOT IN (

SELECT DISTINCT ProductCategory

FROM Sales

WHERE CustomerName = 'Mick Brown'

);