Project 4

Title: OLAP Operations (SQL)

1. Database Creation

Question:

Create a database to store the sales data (Redshift or PostgreSQL).

Create a table named "sales_sample" with the specified columns:

Product_Id (Integer)

Region (varchar(50))-like East ,West etc

Date (Date)

Sales_Amount (int/numeric)

Query:

```
-- Creating Database
CREATE DATABASE SalesData;
USE SalesData;
-- Creating Table
CREATE TABLE sales_sample (
    Product_Id INT,
    Region VARCHAR(50),
    Date DATE,
    Sales_Amount INT
);
```

2. Data Creation

Question:

Insert 10 sample records into the "sales_sample" table, representing sales data.

Query:

```
INSERT INTO sales_sample (Product_Id, Region, Date, Sales_Amount) VALUES
(1, 'East', '2024-11-01', 100),
(2, 'West', '2024-11-02', 200),
```

(3, 'North', '2024-11-03', 150),

(4, 'South', '2024-11-04', 120),

(5, 'East', '2024-11-05', 180),

(6, 'West', '2024-11-06', 210),

(7, 'North', '2024-11-07', 130),

(8, 'South', '2024-11-08', 170),

(9, 'East', '2024-11-09', 90),

(10, 'West', '2024-11-10', 220);

3. Perform OLAP operations

3.1 Question:

Drill Down (From Region to Product Level)

Query:

```
SELECT Region, Product_Id, SUM(Sales_Amount) AS Total_Sales
FROM sales_sample
GROUP BY 1, 2
ORDER BY 1, 2;
```

Output:

Region	Product_Id	Total_Sales
East	1	100
East	5	180
East	9	90
North	3	150
North	7	130
South	4	120
South	8	170
West	2	200
West	6	210
West	10	220

3.2 Question:

Roll Up (From Product to Region Level)

Query:

```
SELECT Region, SUM(Sales_Amount) AS Total_Sales
FROM sales_sample
GROUP BY 1
ORDER BY 1;
```

Output:

Region	Total_Sales
East	370
North	280
South	290
West	630

3.3 Question:

Cube (Simulate using UNION for Product, Region, and Date)

Query:

```
-- Aggregate by Product
SELECT Product_Id AS Dimension, 'Product' AS Category, SUM(Sales_Amount) AS Total_Sales
FROM sales_sample
GROUP BY Product_Id
UNION ALL

-- Aggregate by Region
SELECT Region AS Dimension, 'Region' AS Category, SUM(Sales_Amount) AS Total_Sales
FROM sales_sample
GROUP BY Region
UNION ALL
```

-- Aggregate by Date

SELECT Date AS Dimension, 'Date' AS Category, SUM(Sales_Amount) AS Total_Sales

FROM sales_sample

GROUP BY Date;

Output:

Dimension	Category	Total_Sales
1	Product	100
2	Product	200
3	Product	150
4	Product	120
5	Product	180
6	Product	210
7	Product	130
8	Product	170
9	Product	90
10	Product	220
East	Region	370
West	Region	630
North	Region	280
South	Region	290
2024-11-01	Date	100
2024-11-02	Date	200
2024-11-03	Date	150
2024-11-04	Date	120
2024-11-05	Date	180
2024-11-06	Date	210
2024-11-07	Date	130
2024-11-08	Date	170
2024-11-09	Date	90
2024-11-10	Date	220

3.4 Question:

Slice (Filter for a specific region or date range)

Query:

-- Sales for a specific region (eg - East)

SELECT *

FROM sales_sample

WHERE Region = 'East';

Output:

Product_Id	Region	Date	Sales_Amount
1	East	2024-11-01	100
5	East	2024-11-05	180
9	East	2024-11-09	90

3.5 Question:

Dice (Filter based on multiple criteria)

Query:

```
SELECT *
FROM sales_sample
WHERE Product_Id IN (1, 2, 3)
AND Region IN ('East', 'West')
AND Date BETWEEN '2024-11-01' AND '2024-11-07';
```

Output:

Product_Id	Region	Date	Sales_Amount
1	East	2024-11-01	100
2	West	2024-11-02	200