# Capstone Task 4

**Project: OLAP Operations (using Redshift or PostgreSQL)**

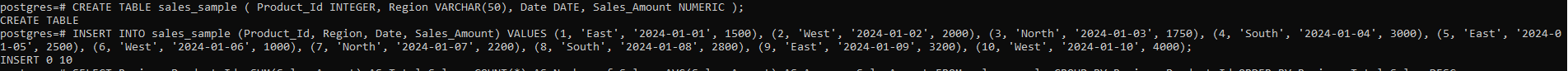
**1. Database Creation**

CREATE DATABASE sales\_data;

CREATE TABLE sales\_sample ( Product\_Id INTEGER, Region VARCHAR(50), Date DATE, Sales\_Amount NUMERIC );

**2. Data Creation**

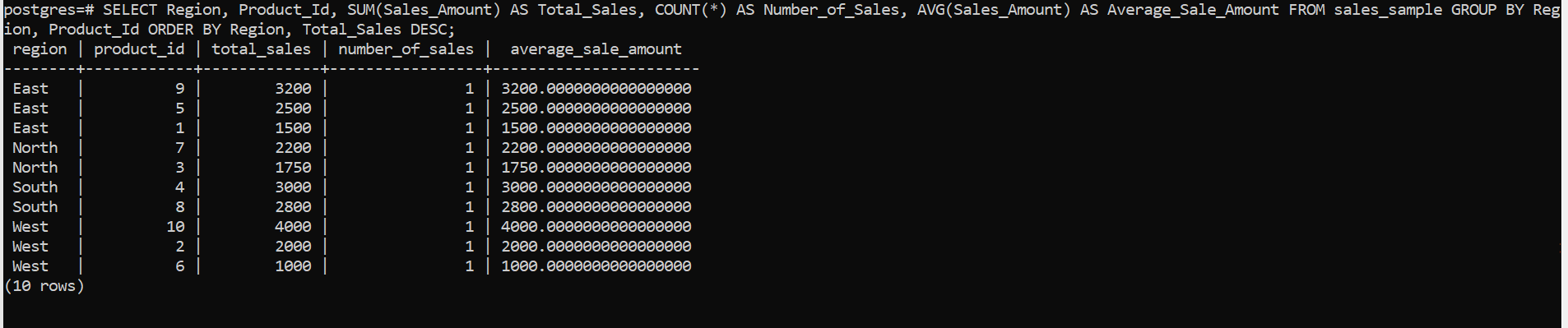
INSERT INTO sales\_sample (Product\_Id, Region, Date, Sales\_Amount) VALUES (1, 'East', '2024-01-01', 1500), (2, 'West', '2024-01-02', 2000), (3, 'North', '2024-01-03', 1750), (4, 'South', '2024-01-04', 3000), (5, 'East', '2024-01-05', 2500), (6, 'West', '2024-01-06', 1000), (7, 'North', '2024-01-07', 2200), (8, 'South', '2024-01-08', 2800), (9, 'East', '2024-01-09', 3200), (10, 'West', '2024-01-10', 4000);



**3. Perform OLAP operations**

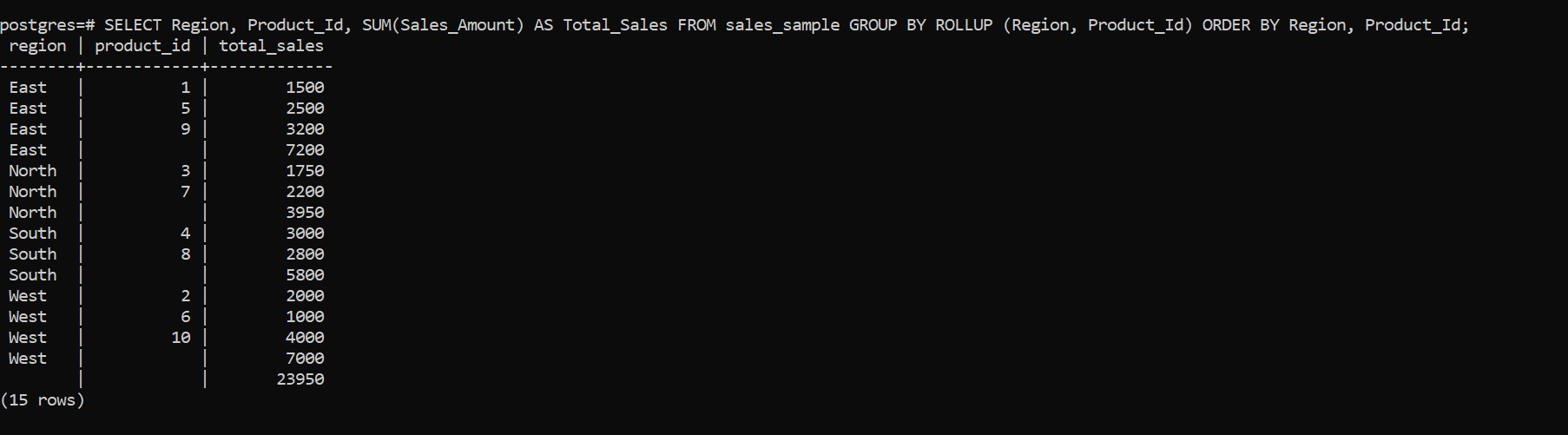
a) Drill Down-Analyze sales data at a more detailed level. Write a query to perform drill down from region to product level to understand sales performance.

SELECT Region, Product\_Id, SUM(Sales\_Amount) AS Total\_Sales, COUNT(\*) AS Number\_of\_Sales, AVG(Sales\_Amount) AS Average\_Sale\_Amount FROM sales\_sample GROUP BY Region, Product\_Id ORDER BY Region, Total\_Sales DESC;



b) Rollup- To summarize sales data at different levels of granularity. Write a query to perform roll up from product to region level to view total sales by region.

SELECT Region, Product\_Id, SUM(Sales\_Amount) AS Total\_Sales FROM sales\_sample GROUP BY ROLLUP (Region, Product\_Id) ORDER BY Region, Product\_Id;



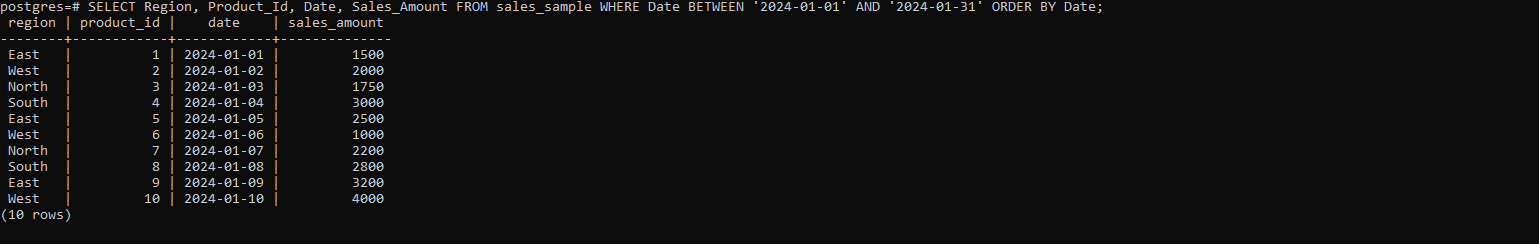
c) Cube - To analyze sales data from multiple dimensions simultaneously. Write a query to Explore sales data from different perspectives, such as product, region, and date.

SELECT Region, Product\_Id, Date, SUM(Sales\_Amount) AS Total\_Sales FROM sales\_sample GROUP BY CUBE (Region, Product\_Id, Date) ORDER BY Region, Product\_Id, Date;



d) Slice- To extract a subset of data based on specific criteria. Write a query to slice the data to view sales for a particular region or date range.

SELECT Region, Product\_Id, Date, Sales\_Amount FROM sales\_sample WHERE Date BETWEEN '2024-01-01' AND '2024-01-31' ORDER BY Date;



e) Dice - To extract data based on multiple criteria. Write a query to view sales for specific combinations of product, region, and date

SELECT Region, Product\_Id, Date, Sales\_Amount FROM sales\_sample WHERE Region = 'East' AND Product\_Id = 1 AND Date BETWEEN '2024-01-01' AND '2024-01-31' ORDER BY Date;

