# **Task4.**

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

*CREATE DATABASE "Sales Data "*

*WITH*

*OWNER = postgres*

*ENCODING = 'UTF8'*

*LC\_COLLATE = 'English\_United States.1252'*

*LC\_CTYPE = 'English\_United States.1252'*

*TABLESPACE = pg\_default*

*CONNECTION LIMIT = -1*

*IS\_TEMPLATE = False;*

**1) Database creation**

CREATE TABLE Sales\_sample (Product\_Id INT, Region VARCHAR(50), On\_date DATE,

Sales\_Amount NUMERIC);

**2) Data Creation**

|  |  |
| --- | --- |
| INSERT INTO Sales\_sample (Product\_Id, Region, On\_date, Sales\_Amount) VALUES  ('1', 'East', '2023-10-10', '45000'),  ('2', 'West', '2023-09-19', '75000'),  ('2', 'East', '2023-10-21', '65000'),  ('3', 'North', '2023-09-20', '40000'),  ('4', 'North', '2023-08-06', '70000'),  ('2', 'South', '2023-08-25', '76000'),  ('5', 'North', '2023-11-23', '48000'),  ('5', 'West', '2023-11-11', '58000'),  ('3', 'East', '2023-09-19', '72000'),  ('1', 'West', '2023-09-29', '63000');  Select Select \* from Sales\_Sample; |  |

**3) OLAP operations**

a) Drill down - Write a query to perform drill down from region to product level to understand sales performance.

|  |  |
| --- | --- |
| SELECT Region, Product\_Id, Sum(Sales\_Amount) AS Sales\_Amount  FROM Sales\_Sample  GROUP BY 1,2  ORDER BY Region, Product\_Id, Sales\_Amount; |  |

b) Roll Up - 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 Sales\_Amount*  *FROM Sales\_Sample*  *GROUP BY ROLLUP (1,2)*  *ORDER BY Region;* |  |

c) Cube - Write a query to explore sales data from different perspectives, such as product, region, and date

|  |  |
| --- | --- |
| *SELECT Region, Product\_Id, On\_Date, SUM(Sales\_Amount) AS Sales\_Amount*  *FROM Sales\_Sample*  *GROUP BY Cube (1,2,3)*  *ORDER BY Region, Product\_Id, On\_Date, Sales\_Amount;* |  |
|  |  |

Note: Due to the length of the Cube output. Reduced output length till North and not included South and West.

d) Slice - Write a query to slice the data to view sales for a particular region or date range

|  |  |
| --- | --- |
| *SELECT Region, Product\_Id, On\_Date, SUM(Sales\_Amount) AS Sales\_Amount*  *FROM Sales\_Sample*  *WHERE Region in('North', 'South') OR On\_Date BETWEEN To\_date('2023-08-20','YYYY-MM-DD') AND To\_Date('2023-10-20','YYYY-MM-DD')*  *GROUP BY 1,2,3*  *ORDER BY Region, Product\_Id, On\_Date, Sales\_Amount;* |  |

e) Dice - Write a query to view sales for specific combinations of product, region, and date

|  |  |
| --- | --- |
| *SELECT Region, Product\_Id, On\_Date, SUM(Sales\_Amount) AS Sales\_Amount*  *FROM Sales\_Sample*  *WHERE Region in ('North', 'South') AND Product\_Id IN (1,2) AND On\_Date*  *BETWEEN To\_date('2023-08-20','YYYY-MM-DD') And To\_Date('2023-10-20','YYYY-MM-DD')*  *GROUP BY 1,2,3*  *ORDER BY Region, Product\_Id, On\_Date, Sales\_Amount;* |  |