# KJSCE/IT/SYBTech/SEM IV/AD/2022-23



**Experiment No. 9**

**Title: Execution of OLAP Operations**

**Batch:A2 Roll No.:16010421073 Experiment No.:9**

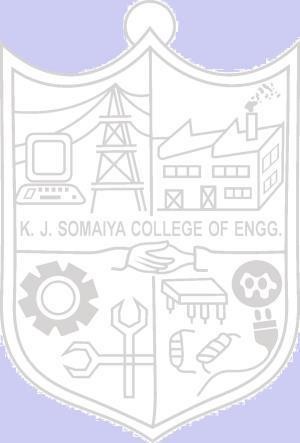
**Aim:** Execution of OLAP operations

**Resources needed: MySQL, Postgres**

**Theory**

## OLAP:

In computing, online analytical processing, or OLAP is an approach to answering multi- dimensional analytical (MDA) queries. OLAP is part of the broader category of business intelligence, which also encompasses relational database report writing and data mining. Typical applications of OLAP include business reporting for sales, marketing, management reporting, business process management (BPM), budgeting and forecasting, financial reporting and similar areas, with new applications coming up, such as agriculture. The term OLAP was created as a slight modification of the traditional database term OLTP (Online Transaction Processing).

OLAP tools enable users to analyze multidimensional data interactively from multiple perspectives. OLAP consists of three basic analytical operations: consolidation (roll-up), drill- down, and slicing and dicing. Consolidation involves the aggregation of data that can be accumulated and computed in one or more dimensions. For example, all sales offices are rolled up to the sales department or sales division to anticipate sales trends. By contrast, the drill-down is a technique that allows users to navigate through the details. For instance, users can view the sales by individual products that make up a region’s sales. Slicing and dicing is a feature whereby users can take out (slicing) a specific set of data of the OLAP cube and view (dicing) the slices from different viewpoints.

OLAP queries can be implemented by using analytical SQL functions

Oracle has extensions to ANSI SQL to allow to quickly computing aggregations and rollups. These new statements include:

* rollup
* cube
* grouping

These simple SQL operators allow creating easy aggregations directly inside the SQL.

## Creating tabular aggregates with ROLLUP:

ROLLUP enables an SQL statement to calculate multiple levels of subtotals across a specified group of dimensions. It also calculates a grand total. ROLLUP is a simple extension to the GROUP BY clause, so its syntax is extremely easy to use. Create cross-tabular reports with CUBE:

In multidimensional jargon, a “cube” is a cross-tabulated summary of detail rows. CUBE enables a SELECT statement to calculate subtotals for all possible combinations of a group of dimensions. It also calculates a grand total.

This is the set of information typically needed for all cross-tabular reports, so CUBE can calculate a cross-tabular report with a single select statement

**Activities:**

# Create a dataset in PostgreSQL and MySQL

1. Apply rollup and cube operations to the same

**PostgresSQL**

**CREATE TABLE emp**

**(**

**deptno INT,**

**job VARCHAR(20),**

**sal INT**

**);**

**insert into emp(deptno,job,sal) values(10,'Sales',30000);**

**insert into emp(deptno,job,sal) values(10,'Marketing',40000);**

**insert into emp(deptno,job,sal) values(10,'Marketing',45000);**

**insert into emp(deptno,job,sal) values(36,'HR',20000);**

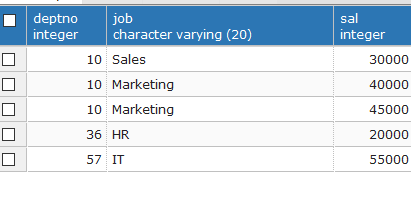
**insert into emp(deptno,job,sal) values(57,'IT',55000);**

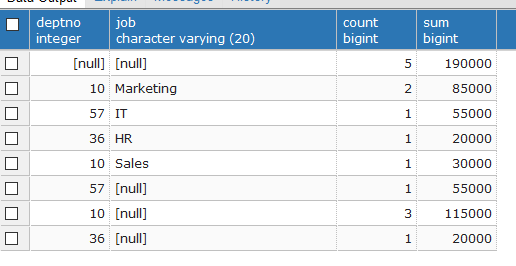
**SELECT \* FROM emp;**

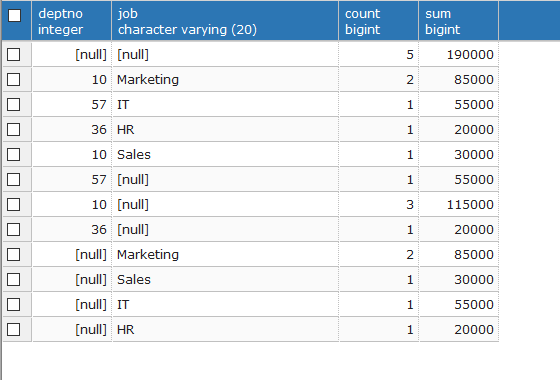
**SELECT deptno,job,count(\*),sum(sal)FROM emp GROUP BY ROLLUP(deptno,job);**

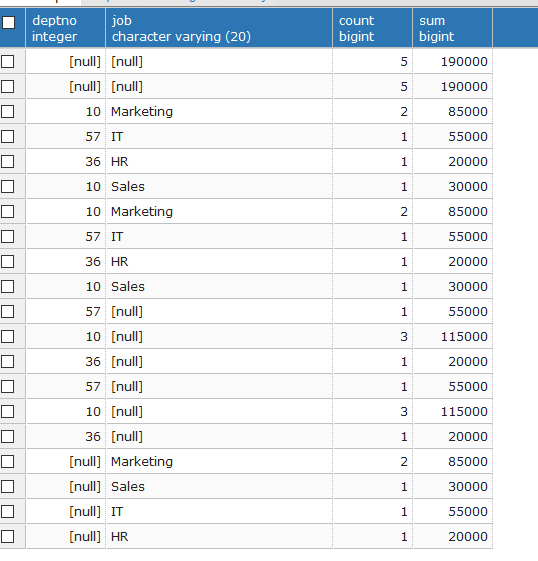
**SELECT deptno, job, count(\*), sum(sal) FROM Emp GROUP BY CUBE(deptno,job);**

**SELECT deptno, job, count(\*), sum(sal)FROM emp GROUP BY GROUPING SETS ( ROLLUP (deptno, job), CUBE (deptno, job) );**

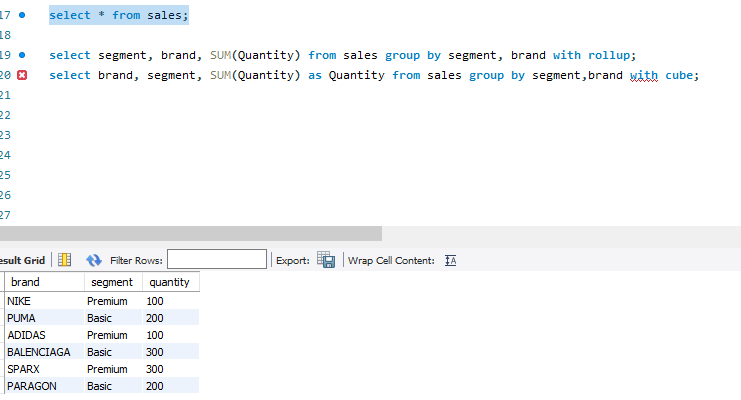


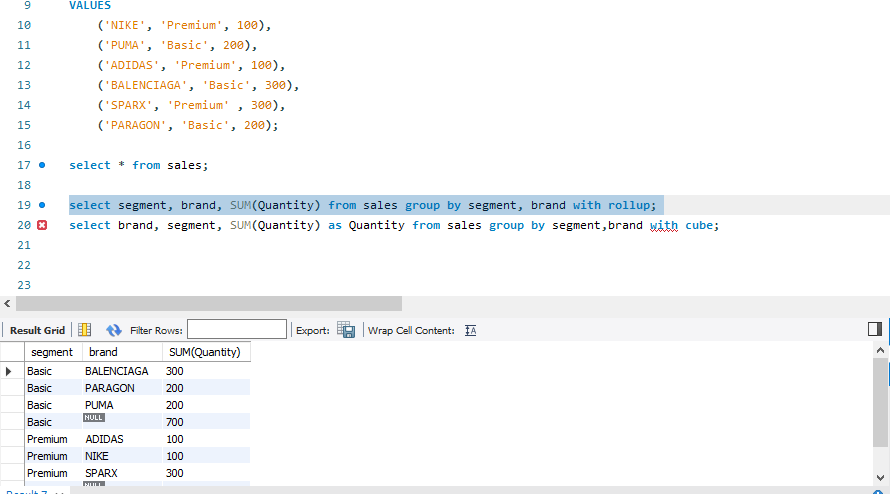






**Mysql**





**Questions:**

1. **Elaborate on the operations applied and results generated to your dataset**

**Ans:** Here we applied Rollup and cube operation in our dataset.

* **Rollup** is an operation that allows us to view data at different levels of detail, by aggregating or summarizing data up a hierarchy.

For example, if we have data on sales revenue at the individual product level, we could use rollup to summarize the data at the product category level, and then at the overall company level. This allows us to see a summary of the data at different levels of detail, and to analyze trends and patterns across different levels.

* **Cube**, on the other hand, is an operation that allows us to analyze data across multiple dimensions, by creating a multi-dimensional view of the data.

For example, if we have data on sales revenue by product, by region, and by time period, we could use cube to analyze the data by any combination of these dimensions. This allows us to see how the data varies across different dimensions, and to identify patterns and relationships that might not be apparent when viewing the data in a traditional tabular format.

1. **Explain if Drill-down, Drill-across can be applied in relational database, Justify with a query implementation.**

**Ans:**

Drill-down and drill-across are two commonly used techniques in data analysis and reporting that can be applied in a relational database environment.

* **Drill-down** refers to the process of navigating from a higher-level summary of data to a more detailed level of information. For example, if we have a sales report that summarizes the total sales revenue by region, we could drill down into a specific region to see the sales revenue by each individual store in that region.
* **Drill-across**, on the other hand, refers to the process of navigating from one dimension of data to another. For example, we might have a sales report that summarizes the total sales revenue by product category, and we could drill across to see the sales revenue by region for each product category.

Here's an example query that demonstrates how drill-down and drill-across can be applied in a relational database environment using SQL:

**SELECT**

**region,**

**product\_category,**

**SUM(sales\_revenue) AS total\_sales**

**FROM**

**sales**

**GROUP BY**

**region,**

**product\_category**

This query would generate a report that summarizes the total sales revenue by region and product category. If we wanted to drill-down to see the sales revenue by individual store for a specific region, we could modify the query like this:

**SELECT**

**store,**

**product\_category,**

**sales\_revenue**

**FROM**

**sales**

**WHERE**

**region = 'North'**

**AND product\_category = 'Electronics'**

This modified query would return the sales revenue for each individual store in the North region that sells electronics products.

If we wanted to drill-across to see the sales revenue by region for a specific product category, we could modify the query like this:

**SELECT**

**region,**

**product\_category,**

**SUM(sales\_revenue) AS total\_sales**

**FROM**

**sales**

**WHERE**

**product\_category = 'Electronics'**

**GROUP BY**

**region,**

**product\_category**

This modified query would return the total sales revenue for the electronics product category in each region.

**Outcomes:**

**CO4:** Apply ETL processing and Online Analytical Processing on the warehouse data.

**Conclusion: (Conclusion to be based on the outcomes achieved)**

**Thus we successfully executed OLAP operation like rollup and cube operation datatset in postgresql and mysql workbench.**

**Grade: AA / AB / BB / BC / CC / CD /DD**

**Signature of faculty in-charge with date**

**References:**

# Paulraj Ponniah, “Data Warehousing: Fundamentals for IT Professionals”, Wiley India