

# ■ DBMS Practice Questions (5 Marks Each)

## Table of Contents

1. Query Processing and Optimization
2. Steps of Query Processing
3. Measures of Query Cost
4. Evaluation of Expressions
5. Basics and Goals of Query Optimization
6. Transformations of Relational Expression
7. Estimating Statistics and Evaluation Plans
8. Data Warehousing
9. Data Warehouse Architecture & Design
10. Dimensional Modeling (Star & Snowflake Schema)
11. Factless Fact Tables & Dimension Updates
12. OLAP and Operations
13. OLAP Architectures
14. ETL Process
15. Techniques of Data Loading

# Query Processing and Optimization

**Scenario-based Question:** How does query optimization improve efficiency in a banking system when retrieving account balances?

**Answer:** Query optimization reduces execution cost by using indexes and efficient join methods, ensuring fast access to account details without scanning entire tables.

**Theory-based Question:** What are the main steps of query processing in DBMS?

**Answer:** Parsing – syntax check and internal query representation.  
Optimization – selecting the best execution plan.  
Evaluation – executing the plan to produce results.

**Query-based Question:** Write an SQL query to find employees who earn more than the average salary of their department.

**Answer:** SELECT name, department, salary  
FROM Employees e  
WHERE salary > (  
    SELECT AVG(salary)  
    FROM Employees  
    WHERE department = e.department  
);

## Steps of Query Processing

**Scenario-based Question:** Why are steps of query processing important in an airline reservation system?

**Answer:** They ensure correctness and efficiency while retrieving flight schedules and booking data, reducing delays in seat availability checks.

**Theory-based Question:** Define the steps of query processing.

**Answer:** The steps include parsing, translation into relational algebra, optimization, and execution of the chosen plan.

**Query-based Question:** Write an SQL query to list the top 3 highest paid employees in each department.

**Answer:** SELECT name, department, salary  
FROM Employees e1  
WHERE 3 > (  
    SELECT COUNT(\*)  
    FROM Employees e2  
    WHERE e2.department = e1.department  
    AND e2.salary > e1.salary  
);

## Measures of Query Cost

**Scenario-based Question:** How do query cost measures affect performance in an e-commerce system?

**Answer:** By estimating CPU, I/O, and memory cost, the system chooses execution plans that minimize query response time when retrieving product details.

**Theory-based Question:** What are the main measures of query cost in DBMS?

**Answer:** Disk I/O, CPU processing, memory usage, and communication overhead.

**Query-based Question:** Write an SQL query to display customers who placed more than 5 orders.

**Answer:** SELECT customer\_id, COUNT(order\_id) AS total\_orders  
FROM Orders  
GROUP BY customer\_id  
HAVING COUNT(order\_id) > 5;

# Evaluation of Expressions

**Scenario-based Question:** How does evaluation of expressions help in a payroll system?

**Answer:** It ensures salary computations (like gross and net pay) are optimized using efficient arithmetic and relational operations.

**Theory-based Question:** Explain evaluation of expressions in DBMS.

**Answer:** It refers to how relational algebra expressions are computed efficiently, often using temporary relations and optimized join orders.

**Query-based Question:** Write an SQL query to list products that have a price higher than all products in the "Stationery" category.

**Answer:** SELECT product\_name, price  
FROM Products  
WHERE price > ALL (  
    SELECT price  
    FROM Products  
    WHERE category = 'Stationery'  
);

# Basics and Goals of Query Optimization

**Scenario-based Question:** Why is query optimization necessary in a hospital management system?

**Answer:** It ensures fast retrieval of patient history and reduces delays by selecting low-cost query execution strategies.

**Theory-based Question:** State two goals of query optimization.

**Answer:** Minimize response time.  
Reduce overall system resource usage.

**Query-based Question:** Write an SQL query to find the second highest salary from the Employees table.

**Answer:** SELECT MAX(salary) AS second\_highest  
FROM Employees  
WHERE salary < (SELECT MAX(salary) FROM Employees);

# Transformations of Relational Expression

**Scenario-based Question:** How do relational expression transformations help in retail databases?

**Answer:** They allow equivalent queries to be executed with lower cost, improving performance for inventory and sales reports.

**Theory-based Question:** What is the purpose of relational expression transformation?

**Answer:** It rewrites queries into equivalent forms to find more efficient execution plans.

**Query-based Question:** Write an SQL query to find departments that have at least one employee earning above 1,00,000.

**Answer:** SELECT DISTINCT department  
FROM Employees  
WHERE salary > 100000;

# Estimating Statistics and Evaluation Plans

**Scenario-based Question:** Why is statistics estimation important in telecom databases?

**Answer:** It helps the optimizer choose efficient join and selection methods when querying large customer call records.

**Theory-based Question:** What is the role of statistics in query optimization?

**Answer:** Statistics (like table size, distinct values, histograms) guide the optimizer in selecting efficient plans.

**Query-based Question:** Write an SQL query to find customers who spent more than the average total spending.

```
Answer: SELECT customer_id, SUM(amount) AS total_spent
FROM Transactions
GROUP BY customer_id
HAVING SUM(amount) > (
    SELECT AVG(total)
    FROM (SELECT SUM(amount) AS total
    FROM Transactions
    GROUP BY customer_id) t
);
```

## Data Warehousing

**Scenario-based Question:** How does a data warehouse help a retail chain?

**Answer:** It consolidates sales data from all branches, enabling trend analysis and better decision-making.

**Theory-based Question:** Define a data warehouse and give one feature.

**Answer:** A data warehouse is subject-oriented, integrated, and time-variant.  
Feature: Provides historical data for analysis.

**Query-based Question:** Write an SQL query to calculate monthly total sales from a Sales table.

```
Answer: SELECT EXTRACT(MONTH FROM sale_date) AS month, SUM(amount) AS monthly_sales
FROM Sales
GROUP BY EXTRACT(MONTH FROM sale_date);
```

## Data Warehouse Architecture & Design

**Scenario-based Question:** Why is architecture important in data warehouses?

**Answer:** It determines how data is stored, accessed, and analyzed efficiently, ensuring scalability and performance.

**Theory-based Question:** What are the layers of a data warehouse architecture?

**Answer:** Data sources, ETL layer, warehouse storage, and OLAP/analysis tools.

**Query-based Question:** Write an SQL query to show top 5 products by sales value.

```
Answer: SELECT product_id, SUM(amount) AS total_sales
FROM Sales
GROUP BY product_id
ORDER BY total_sales DESC
LIMIT 5;
```

## Dimensional Modeling (Star & Snowflake Schema)

**Scenario-based Question:** Why is dimensional modeling important in sales analysis?

**Answer:** It simplifies queries by organizing data into fact and dimension tables for efficient reporting.

**Theory-based Question:** Define star schema.

**Answer:** A star schema has a central fact table linked to dimension tables like product, time, and location.

**Query-based Question:** Write an SQL query to get total sales per year and product category.

**Answer:** SELECT category, EXTRACT(YEAR FROM sale\_date) AS year, SUM(amount) AS total\_sales  
FROM Sales s  
JOIN Products p ON s.product\_id = p.product\_id  
GROUP BY category, EXTRACT(YEAR FROM sale\_date);

## Factless Fact Tables & Dimension Updates

**Scenario-based Question:** How is a factless fact table useful in education analytics?

**Answer:** It captures events like student attendance without numeric facts, enabling useful reports.

**Theory-based Question:** What is a factless fact table?

**Answer:** A table that records relationships between dimensions without measurable facts.

**Query-based Question:** Write an SQL query to find students who attended all classes in a semester.

**Answer:** SELECT student\_id  
FROM Attendance  
GROUP BY student\_id  
HAVING COUNT(DISTINCT class\_id) = (SELECT COUNT(DISTINCT class\_id) FROM Classes);

## OLAP and Operations

**Scenario-based Question:** How does OLAP help in analyzing sales?

**Answer:** OLAP provides roll-up, drill-down, and slice-dice operations for multidimensional analysis.

**Theory-based Question:** List two OLAP operations with examples.

**Answer:** Roll-up (monthly to yearly sales), Drill-down (yearly to monthly sales).

**Query-based Question:** Write an SQL query to show total sales by region and product.

**Answer:** SELECT region, product\_id, SUM(amount) AS total\_sales  
FROM Sales  
GROUP BY region, product\_id;

## OLAP Architectures

**Scenario-based Question:** Why are OLAP architectures important?

**Answer:** They define how OLAP systems are implemented (ROLAP, MOLAP, HOLAP) for performance and scalability.

**Theory-based Question:** Differentiate between ROLAP and MOLAP.

**Answer:** ROLAP uses relational databases; MOLAP uses multidimensional cubes for faster queries.

**Query-based Question:** Write an SQL query to calculate quarterly sales per region.

**Answer:** SELECT region, EXTRACT(QUARTER FROM sale\_date) AS quarter, SUM(amount) AS total\_sales  
FROM Sales  
GROUP BY region, EXTRACT(QUARTER FROM sale\_date);

## ETL Process

**Scenario-based Question:** Why is ETL important in a banking system?

**Answer:** ETL extracts customer and transaction data, transforms it into consistent format, and loads it into a data warehouse.

**Theory-based Question:** Define ETL with its three phases.

**Answer:** Extraction (collecting data), Transformation (cleaning, formatting), Loading (inserting into warehouse).

**Query-based Question:** Write an SQL query to remove duplicate rows from a Customers table.

**Answer:** DELETE FROM Customers c1  
WHERE EXISTS (  
SELECT 1  
FROM Customers c2  
WHERE c1.customer\_id > c2.customer\_id  
AND c1.email = c2.email  
);

## Techniques of Data Loading

**Scenario-based Question:** How do efficient loading techniques help a telecom data warehouse?

**Answer:** They allow bulk loading of millions of records quickly, reducing downtime.

**Theory-based Question:** Name two techniques of data loading.

**Answer:** Incremental loading and bulk loading.

**Query-based Question:** Write an SQL query to insert only new records from a staging table into the main Customers table.

**Answer:** INSERT INTO Customers (customer\_id, name, email)  
SELECT s.customer\_id, s.name, s.email  
FROM Staging\_Customers s  
WHERE NOT EXISTS (  
SELECT 1 FROM Customers c WHERE c.customer\_id = s.customer\_id  
);