Database Important Questions & Answers

1. Domain

In a database, a domain is the set of all possible values that can be assigned to an attribute. For example, if we consider an attribute Age in a Student table, its domain may be all integer values ranging from 0 to 120. Domains help maintain data integrity by ensuring only valid values are entered.

2. Data Dependency

Data dependency refers to the relationship between attributes where the value of one attribute depends on another attribute. For example, a student's roll number uniquely determines their name, department, and course. Thus, Roll_No → Name, Department, Course is a data dependency.

3. Armstrong Axioms

Armstrong's axioms are a set of inference rules used to derive all functional dependencies in a relational database. They are:
1. Reflexivity: If Y is a subset of X, then X determines Y (X → Y).
2. Augmentation: If X → Y, then XZ → YZ also holds.
3. Transitivity: If X → Y and Y → Z, then X → Z.

4. Data Independence

Data independence is the ability to modify a schema definition at one level of the database system without affecting the schema at the next higher level. There are two types:
1. Logical Data Independence: Changes in the logical schema do not affect the external schema or applications.
2. Physical Data Independence: Changes in the physical storage do not affect the logical schema.

5. Advantages of Query Decomposition

Query decomposition is the process of breaking down complex queries into smaller parts. Its advantages are:
• It improves query optimization.
• It reduces redundancy and cost of execution.
• It simplifies complex queries for the DBMS.
• It ensures correctness of the results.

6. HAVING vs WHERE Clause

The WHERE clause is used to filter rows before any grouping is performed. It applies to individual rows. The HAVING clause is used to filter groups after the GROUP BY clause

has been applied. It is generally used with aggregate functions such as COUNT, SUM, or AVG.

7. ORDER BY vs GROUP BY

ORDER BY is used to sort the result set of a query in ascending or descending order based on one or more columns. GROUP BY is used to arrange identical data into groups and is often used with aggregate functions. While ORDER BY organizes the output, GROUP BY organizes the data into groups for summary purposes.

8. Relational Model vs Object-Oriented Model

In the Relational Model, data is represented in tables (relations) with rows (tuples) and columns (attributes). It is based on mathematical set theory and predicate logic. In contrast, the Object-Oriented Model stores data as objects, similar to object-oriented programming concepts, and supports encapsulation, inheritance, and polymorphism.

9. Tree and Relational Algebra Operators

A Tree is a hierarchical data structure where elements (nodes) are connected in parent-child relationships. Relational Algebra is a set of operations used to manipulate relations (tables). The basic operators include:
• Selection (σ)
• Projection (π)
• Union (∪)
• Difference (−)
• Cartesian Product (×)
• Join (⋈)

10. Data Abstraction

Data abstraction refers to hiding the complexities of data storage and showing only the essential information to the user. It is organized into three levels:
1. Physical Level: Describes how data is stored physically in storage devices.
2. Logical Level: Describes what data is stored in the database and the relationships among data.
3. View Level: Describes only a part of the database to the user, providing security and simplicity.

11. Projection in Relational Algebra

Projection is a relational algebra operator that extracts specified columns from a relation, eliminating duplicates. For example, π(Name, Age)(Student) will return only the Name and Age columns from the Student table.

12. Relational Algebra Operators

Relational algebra provides a formal foundation for relational databases and SQL. The key operators are:
• Selection

- Projection
- Union
- Set Difference
- Cartesian Product
- Join

These operators are used to query and manipulate relations.

13. Relational Model

The relational model, introduced by E. F. Codd in 1970, represents data in the form of relations (tables). Each table consists of rows (tuples) and columns (attributes). It provides data integrity, simplicity, and flexibility, making it the most widely used data model in database systems.