

DBMS Interview Questions

1. What is database?

A database is a collection of information that is organized. So that it can easily be accessed, managed, and updated.

2. What is DBMS?

DBMS stands for Database Management System. It is a collection of programs that enables user to create and maintain a database.

3. What is a Database system?

The database and DBMS software together is called as Database system.

4. What are the advantages of DBMS?

- I. Redundancy is controlled.
- II. Providing multiple user interfaces.
- III. Providing backup and recovery
- IV. Unauthorized access is restricted.
- V. Enforcing integrity constraints.

5. What is normalization?

It is a process of analysing the given relation schemas based on their Functional Dependencies (FDs) and primary key to achieve the properties
(1).Minimizing redundancy, (2). Minimizing insertion, deletion and update anomalies.

6. What is Data Model?

A collection of conceptual tools for describing data, data relationships data semantics and constraints.

7. What is E-R model?

This data model is based on real world that consists of basic objects called entities and of relationship among these objects. Entities are described in a database by a set of attributes.

8. What is Object Oriented model?

This model is based on collection of objects. An object contains values stored in instance variables with in the object. An object also contains bodies of code that operate on the object. These bodies of code are called methods. Objects that contain same types of values and the same methods are grouped together into classes.

9. What is an Entity?

An entity is a thing or object of importance about which data must be captured.

10. What is DDL (Data Definition Language)?

A data base schema is specifies by a set of definitions expressed by a special language called DDL.

11. What is DML (Data Manipulation Language)?

This language that enable user to access or manipulate data as organised by appropriate data model. Procedural DML or Low level: DML requires a user to specify what data are needed and how to get those data. Non-Procedural DML or High level: DML requires a user to specify what data are needed without specifying how to get those data

12. What is DML Compiler?

It translates DML statements in a query language into low-level instruction that the query evaluation engine can understand.

13. What is Query evaluation engine?

It executes low-level instruction generated by compiler.

14. What is Functional Dependency?

Functional Dependency is the starting point of normalization. Functional Dependency exists when a relation between two attributes allows you to uniquely determine the corresponding attribute's value.

15. What is 1 NF (Normal Form)?

The first normal form or 1NF is the first and the simplest type of normalization that can be implemented in a database. The main aims of 1NF are to:

1. Eliminate duplicative columns from the same table.
2. Create separate tables for each group of related data and identify each row with a unique column (the primary key).

16. What is Fully Functional dependency?

A functional dependency $X \rightarrow Y$ is full functional dependency if removal of any attribute A from X means that the dependency does not hold any more.

17. What is 2NF?

A relation schema R is in 2NF if it is in 1NF and every non-prime attribute A in R is fully functionally dependent on primary key.

18. What is 3NF?

A relation is in third normal form if it is in Second Normal Form and there are no functional (transitive) dependencies between two (or more) non-primary key attributes.

19. What is BCNF (Boyce-Codd Normal Form)?

A table is in Boyce-Codd normal form (BCNF) if and only if it is in 3NF and every determinant is a candidate key.

20. What is 4NF?

Fourth normal form requires that a table be BCNF and contain no multi-valued dependencies.

21. What is 5NF?

A table is in fifth normal form (5NF) or Project-Join Normal Form (PJNF) if it is in 4NF and it cannot have a lossless decomposition into any number of smaller tables.

22. What is a query?

A query with respect to DBMS relates to user commands that are used to interact with a data base.

23. What is meant by query optimization?

The phase that identifies an efficient execution plan for evaluating a query that has the least estimated cost is referred to as query optimization.

24. What is an attribute?

It is a particular property, which describes the entity.

25. What is RDBMS?

Relational Data Base Management Systems (RDBMS) are database management systems that maintain data records and indices in tables.

26. What's difference between DBMS and RDBMS?

DBMS provides a systematic and organized way of storing, managing and retrieving from collection of logically related information. RDBMS also provides what DBMS provides but above that it provides relationship integrity.

27. What is SQL?

SQL stands for Structured Query Language. SQL is an ANSI (American National Standards Institute) standard computer language for accessing and manipulating database systems. SQL statements are used to retrieve and update data in a database.

28. What is Stored Procedure?

A stored procedure is a named group of SQL statements that have been previously created and stored in the server database.

29. What is a view?

A view may be a subset of the database or it may contain virtual data that is derived from the database files but is not explicitly stored.

30. What is Trigger?

A trigger is a SQL procedure that initiates an action when an event (INSERT, DELETE or UPDATE) occurs.

31. What is Index?

An index is a physical structure containing pointers to the data.

32. What is extension and intension?

Extension -It is the number of tuples present in a table at any instance. This is time dependent.

Intension -It is a constant value that gives the name, structure of table and the constraints laid on it.

33. What do you mean by atomicity and aggregation?

Atomicity-Atomicity states that database modifications must follow an “all or nothing” rule. Each transaction is said to be “atomic.” If one part of the transaction fails, the entire transaction fails.

Aggregation - A feature of the entity relationship model that allows a relationship set to participate in another relationship set. This is indicated on an ER diagram by drawing a dashed box around the aggregation.

34. What is RDBMS KERNEL?

Two important pieces of RDBMS architecture are the kernel, which is the software, and the data dictionary, which consists of the system- level data structures used by the kernel to manage the database.

35. Name the sub-systems of a RDBMS?

I/O, Security, Language Processing, Process Control, Storage Management, Logging and Recovery, Distribution Control, Transaction Control, Memory Management, Lock Management.

36. How do you communicate with an RDBMS?

You communicate with an RDBMS using Structured Query Language (SQL)

37. Disadvantage in File Processing System?

- Data redundancy & inconsistency.
- Difficult in accessing data.
- Data isolation.
- Data integrity.
- Concurrent access is not possible.
- Security Problems.

38. What is VDL (View Definition Language)?

It specifies user views and their mappings to the conceptual schema.

39. What is SDL (Storage Definition Language)?

This language is to specify the internal schema. This language may Specify the mapping between two schemas.

40. Describe concurrency control?

Concurrency control is the process managing simultaneous operations against a database so that database integrity is not compromised. There are two approaches to concurrency control.

The pessimistic approach involves locking and the optimistic approach involves versioning.

41. Describe the difference between homogeneous and heterogeneous distributed database?

A homogeneous database is one that uses the same DBMS at each node. A heterogeneous database is one that may have a different DBMS at each node.

42. What is a distributed database?

A distributed database is a single logical database that is spread across more than one node or locations that are all connected via some communication link.

43. Explain the difference between two and three-tier architectures?

Three-tier architecture includes a client and two server layers.

The application code is stored on the application server and the database is stored on the database server. A two-tier architecture includes a client and one server layer. The database is stored on the database server.

44. Briefly describe the three types of SQL commands?

Data definition language commands are used to create, alter, and drop tables. Data manipulation commands are used to insert, modify, update, and query data in the database. Data control language commands help the DBA to control the database.

45. List some of the properties of a relation?

Relations in a database have a unique name and no multivalued attributes exist. Each row is unique and each attribute within a relation has a unique name. The sequence of both columns and rows is irrelevant.

46. Explain the differences between an intranet and an extranet?

An Internet database is accessible by everyone who has access to a Web site. An intranet database limits access to only people within a given organization.

47. What is SQL Deadlock?

Deadlock is a unique situation in a multi user system that causes two or more users to wait indefinitely for a locked resource.

48. What is a Catalog?

A catalog is a table that contains the information such as structure of each file, the type and storage format of each data item and various constraints on the data .The information stored in the catalog is called Metadata.

49. What is data ware housing & OLAP?

Data warehousing and OLAP (online analytical processing) systems are the techniques used in many companies to extract and analyze useful information from very large databases for decision making .

50. Describe the three levels of data abstraction?

Physical level: The lowest level of abstraction describes how data are stored.

Logical level: The next higher level of abstraction, describes what data are stored in database and what relationship among those data.

View level: The highest level of abstraction describes only part of entire database.

51. What is Data Independence?

Data independence means that the application is independent of the storage structure and access strategy of data.

52. How many types of relationship exist in database designing?

There are three major relationship models:-

One-to-one

One-to-many

Many-to-many

53. What is order by clause?

ORDER BY clause helps to sort the data in either ascending order to descending

54. What is the use of DBCC commands?

DBCC stands for database consistency checker. We use these commands to check the consistency of the databases, i.e., maintenance, validation task and status checks.

55. What is Collation?

Collation refers to a set of rules that determine how data is sorted and compared.

56. What is difference between DELETE & TRUNCATE commands?

Delete command removes the rows from a table based on the condition that we provide with a WHERE clause. Truncate will actually remove all the rows from a table and there will be no data in the table after we run the truncate command.

57. What is Hashing technique?

This is a primary file organization technique that provides very fast access to records on certain search conditions.

58. What is a transaction?

A transaction is a logical unit of database processing that includes one or more database access operations.

59. What are the different phases of Transaction?

Analysis phase

Redo phase

Undo phase

60. What is “transparent dbms”?

It is one, which keeps its physical structure hidden from user.

61. What are the primitive operations common to all record management System?

Addition, deletion and modification.

62. Explain the differences between structured data and unstructured data.

Structured data are facts concerning objects and events. The most important structured data are numeric, character, and dates.

Structured data are stored in tabular form. Unstructured data are multimedia data such as documents, photographs, maps, images, sound, and video clips. Unstructured data are most commonly found on Web servers and Web-enabled databases.

63. What are the major functions of the database administrator?

Managing database structure, controlling concurrent processing, managing processing rights and responsibilities, developing database security, providing for database recovery, managing the DBMS and maintaining the data repository.

64. What is a dependency graph?

A dependency graph is a diagram that is used to portray the connections between database elements.

65. Explain the difference between an exclusive lock and a shared lock?

An exclusive lock prohibits other users from reading the locked resource; a shared lock allows other users to read the locked resource, but they cannot update it.

66. Explain the "paradigm mismatch" between SQL and application programming languages.

SQL statements return a set of rows, while an application program works on one row at a time. To resolve this mismatch the results of SQL statements are processed as pseudofiles, using a cursor or pointer to specify which row is being processed.

67. Name four applications for triggers.

- (1) Providing default values, (2) enforcing data constraints,
- (3) Updating views and (4) enforcing referential integrity

68. What are the advantages of using stored procedures?

The advantages of stored procedures are (1) greater security, (2) decreased network traffic, (3) the fact that SQL can be optimized and (4) code sharing which leads to less work, standardized processing, and specialization among developers.

69. Explain the difference between attributes and identifiers.

Entities have attributes. Attributes are properties that describe the entity's characteristics. Entity instances have identifiers. Identifiers are attributes that name, or identify, entity instances.

70. What is Enterprise Resource Planning (ERP), and what kind of a database is used in an ERP application?

Enterprise Resource Planning (ERP) is an information system used in manufacturing companies and includes sales, inventory, production planning, purchasing and other business functions. An ERP system typically uses a multiuser database.

71. Describe the difference between embedded and dynamic SQL?

Embedded SQL is the process of including hard coded SQL statements. These statements do not change unless the source code is modified. Dynamic SQL is the process of generating SQL on the fly. The statements generated do not have to be the same each time.

72. Explain a join between tables

A join allows tables to be linked to other tables when a relationship between the tables exists. The relationships are established by using a common column in the tables and often uses the primary/foreign key relationship.

73. Describe a subquery.

A subquery is a query that is composed of two queries. The first query (inner query) is within the WHERE clause of the other query (outer query).

74. Compare a hierarchical and network database model?

The hierarchical model is a top-down structure where each parent may have many children but each child can have only one parent. This model supports one-to-one and one-to-many relationships.

The network model can be much more flexible than the hierarchical model since each parent can have multiple children but each child can also have multiple parents. This model supports one-to-one, one-to-many, and many-to-many relationships.

75. Explain the difference between a dynamic and materialized view.

A dynamic view may be created every time that a specific view is requested by a user. A materialized view is created and or updated infrequently and it must be synchronized with its associated base table(s).

76. Explain what needs to happen to convert a relation to third normal form.

First you must verify that a relation is in both first normal form and second normal form. If the relation is not, you must convert into second normal form. After a relation is in second normal form, you must remove all transitive dependencies.

77. Describe the four types of indexes?

A unique primary index is unique and is used to find and store a row. A nonunique primary index is not unique and is used to find a row but also where to store a row (based on its unique primary index). A unique secondary index is unique for each row and used to find table rows. A nonunique secondary index is not unique and used to find table rows.

78. Explain minimum and maximum cardinality?

Minimum cardinality is the minimum number of instances of an entity that can be associated with each instance of another entity. Maximum cardinality is the maximum number of instances of an entity that can be associated with each instance of another entity.

79. What is deadlock? How can it be avoided? How can it be resolved once it occurs?

Deadlock occurs when two transactions are each waiting on a resource that the other transaction holds. Deadlock can be prevented by requiring transactions to acquire all locks at the same time; once it occurs, the only way to cure it is to abort one of the transactions and back out of partially completed work.

80. Explain what we mean by an ACID transaction.

An ACID transaction is one that is atomic, consistent, isolated, and durable. Durable means that database changes are permanent. Consistency can mean either statement level or transaction level consistency. With transaction level consistency, a transaction may not see its own changes. Atomic means it is performed as a unit.

81. Under what conditions should indexes be used?

Indexes can be created to enforce uniqueness, to facilitate sorting, and to enable fast retrieval by column values. A good candidate for an index is a column that is frequently used with equal conditions in WHERE clauses.

82. What is difference between SQL and SQL SERVER?

SQL is a language that provides an interface to RDBMS, developed by IBM. SQL SERVER is a RDBMS just like Oracle, DB2.

83. What is Specialization?

It is the process of defining a set of subclasses of an entity type where each subclass contain all the attributes and relationships of the parent entity and may have additional attributes and relationships which are specific to itself.

84. What is generalization?

It is the process of finding common attributes and relations of a number of entities and defining a common super class for them.

85. What is meant by Proactive, Retroactive and Simultaneous Update?

Proactive Update: The updates that are applied to database before it becomes effective in real world.

Retroactive Update: The updates that are applied to database after it becomes

effective in real world.

Simultaneous Update: The updates that are applied to database at the same time when it becomes effective in real world.

86. What is RAID Technology?

Redundant array of inexpensive (or independent) disks. The main goal of raid technology is to even out the widely different rates of performance improvement of disks against those in memory and microprocessor. Raid technology employs the technique of data striping to achieve higher transfer rates.

87. What are serial, non serial schedule?

A schedule S is serial if, for every transaction T participating in the schedule, all the operations of T is executed consecutively in the schedule, otherwise, the schedule is called non-serial schedule.

88. What are conflict serializable schedules?

A schedule S of n transactions is serializable if it is equivalent to some serial schedule of the same n transactions.

89. What is view serializable?

A schedule is said to be view serializable if it is view equivalent with some serial schedule.

90. What is a foreign key?

A key of a relation schema is called as a foreign key if it is the primary key of some other relation to which it is related to.

91. What are the disadvantages of using a dbms?

- 1) High initial investments in h/w, s/w, and training.
- 2) Generality that a DBMS provides for defining and processing data.
- 3) Overhead for providing security, concurrency control, recovery, and integrity functions.

92. What is Lossless join property?

It guarantees that the spurious tuple generation does not occur with respect to relation schemas after decomposition.

93. What is a Phantom Deadlock?

In distributed deadlock detection, the delay in propagating local information might cause the deadlock detection algorithms to identify deadlocks that do not really exist. Such situations are called phantom deadlocks and they lead to unnecessary aborts.

94. What is a checkpoint and When does it occur?

A Checkpoint is like a snapshot of the DBMS state. By taking checkpoints, the DBMS can reduce the amount of work to be done during restart in the event of subsequent crashes.

95. What is schema?

The description of a data base is called the database schema , which is specified during database design and is not expected to change frequently . A displayed schema is called schema diagram .We call each object in the schema as schema construct.