A1. A join selects columns from 2 or more tables. A union selects rows.

Q2. What is normalization and what are the five normal forms?

A2. Normalization is a design procedure for representing data in tabular format. The five normal forms are progressive rules to represent the data with minimal redundancy.

Q3. What are foreign keys?

A3. These are attributes of one table that have matching values in a primary key in another table, allowing for relationships between tables.

Q4. Describe the elements of the SELECT query syntax.

A4. SELECT element FROM table WHERE conditional statement.

Q5. Explain the use of the WHERE clause.

A5. WHERE is used with a relational statement to isolate the object element or row.

Q6. What techniques are used to retrieve data from more than one table in a single SQL statement?

A6. Joins, unions and nested selects are used to retrieve data.

Q7. What is a view? Why use it?

A7. A view is a virtual table made up of data from base tables and other views, but not stored separately.

Q8. Explain an outer join.

A8. An outer join includes rows from tables when there are no matching values in the tables.

Q9. What is a subselect? Is it different from a nested select?

A9. A subselect is a select which works in conjunction with another select. A nested select is a kind of subselect where the inner select passes to the where criteria for the outer select.

Q10. What is the difference between group by and order by?

A10. Group by controls the presentation of the rows, order by controls the presentation of the columns for the results of the SELECT statement.

Q11. What keyword does an SQL SELECT statement use for a string search?

A11. The LIKE keyword allows for string searches. The % sign is used as a wildcard.

Q12. What are some sql aggregates and other built-in functions?

A12. The common aggregate, built-in functions are AVG, SUM, MIN, MAX, COUNT and DISTINCT.

Q13. How is the SUBSTR keyword used in sql?

A13. SUBSTR is used for string manipulation with column name, first position and string length used as arguments. Eg. SUBSTR (NAME, 1 3) refers to the first three characters in the column NAME.

Q14. Explain the EXPLAIN statement.

A14. The explain statement provides information about the optimizer's choice of access path of the sql.

Q15. What is referential integrity?

A15. Referential integrity refers to the consistency that must be maintained between primary and foreign keys, ie every foreign key value must have a corresponding primary key value.

Q16. What is a NULL value? What are the pros and cons of using NULLS?

A16. A NULL value takes up one byte of storage and indicates that a value is not present as opposed to a space or zero value. It's the DB2 equivalent of TBD on an organizational chart and often correctly portrays a business situation. Unfortunately, it requires extra coding for an application program to handle this situation.

Q17. What is a synonym? How is it used?

A17. A synonym is used to reference a table or view by another name. The other name can then be written in the application code pointing to test tables in the development stage and to production entities when the code is migrated. The synonym is linked to the AUTHID that created it.

Q18. What is an alias and how does it differ from a synonym?

A18. An alias is an alternative to a synonym, designed for a distributed environment to avoid having to use the location qualifier of a table or view. The alias is not dropped when the table is dropped.

Q19. When can an insert of a new primary key value threaten referential integrity?

A19. Never. New primary key values are not a problem. However, the values of foreign key inserts must have corresponding primary key values in their related tables. And updates of primary key values may require changes in foreign key values to maintain referential integrity.

Q20. What is the difference between static and dynamic sql?

A20. Static sql is hard-coded in a program when the programmer knows the statements to be executed. For dynamic sql the program must dynamically allocate memory to receive the query results.

Q21. Compare a subselect to a join.

A21. Any subselect can be rewritten as a join, but not vice versa. Joins are usually more efficient as join rows can be returned immediately, subselects require a temporary work area for inner selects results while processing the outer select.

Q22. What is the difference between IN subselects and EXISTS subselect?

A22. If there is an index on the attributes tested an IN is more efficient since DB2 uses the index for the IN. (IN for index is the mnemonic).

Q23. What is a Cartesian product?

A23. A Cartesian product results from a faulty query. It is a row in the results for every combination in the join tables.

Q24. What is a tuple?

A24. A tuple is an instance of data within a relational database.

Q25. What is the difference between static and dynamic sql?

A25. Static sql is compiled and optimized prior to its execution; dynamic is compiled and optimized during execution.

Q26. Any SQL implementation covers data types in couple of main categories. Which of the following are those data types ? (Check all that apply) A. NUMERIC B. CHARACTER C. DATE AND TIME D. BLOBS E. BIT

A26. A,B,C. Not all SQL implementations have a BLOB or a BIT data types.

Q27. We have a table with a CHARACTER data type field. We apply a ">" row comparison between this field and another CHARACTER field in another table. What will be the results for records with field value of NULL ? (Check one that applies the best) A. TRUE B. FALSE C. UNKNOWN D. Error. E. Those records will be ignored

A27. C. NULL in a row when compared will give an UNKNOWN result.

Q28. Any database needs to go through a normalization process to make sure that data is represented only once. This will eliminate problems with creating or destroying data in the database. The normalization process is done usually in three steps which results in first, second and third normal forms. Which best describes the process to obtain the third normal form? (Check one that applies the best) A. Each table should have related columns. B. Each separate table should have a primary key. C. We have a table with multi-valued key. All columns that are dependent on only one or on some of the keys should be moved in a different table. D. If a table has columns not dependent on the primary keys, they need to be moved in a separate table. E. Primary key is always UNIQUE and NOT NULL.

A28. D. All columns in a table should be dependent on the primary key. This will eliminate transitive dependencies in which A depends on B, and B depends on C, but we're not sure how C depends on A.

Q29. SQL can be embedded in a host program that uses a relational database as a persistent data repository. Some of the most important pre-defined structures for this mechanism are SQLDA ("SQL Descriptor Area") and SQLCA ("SQL Communications Area") SQLCA contains two structures - SQLCODE and SQLSTATE. SQLSTATE is a standard set of error messages and warnings in which the first two characters defines the class and the last three defines the subclass of the error. Which of the following SQLSTATE codes is interpreted as "No data returned"? (Check one that applies the best) A. 00xxx B. 01xxx C. 02xxx D. 22xxx E. 2Axxx

A29. C. 00 - is successful completion, 01 - warnings, 22 - is data exception and 2A is syntax error. The SQLSTATE code format returned for "No data returned" is "02xxx".