# SQL and Database Knowledge: Questions and Answers

61. What is SQL, and how is it used in data science?

SQL (Structured Query Language) is a standard language used to communicate with relational databases. In data science, SQL is used to extract, manipulate, and analyze data stored in databases. It allows data scientists to query large datasets efficiently and perform operations like filtering, joining, grouping, and aggregating data.

62. Explain the difference between SQL's INNER JOIN and LEFT JOIN.

An INNER JOIN returns only the rows that have matching values in both tables. A LEFT JOIN returns all rows from the left table and the matched rows from the right table; if there is no match, NULL values are returned for columns from the right table.

63. What is a primary key and a foreign key in a relational database?

A primary key is a column or a set of columns that uniquely identifies each row in a table. A foreign key is a column or a set of columns in one table that refers to the primary key in another table, establishing a relationship between the two tables.

64. How do you write a SQL query to retrieve data from a database table?

A basic SQL query to retrieve data is: SELECT column1, column2 FROM table\_name WHERE condition; This retrieves specific columns from a table where the condition is met.

65. What is the purpose of the GROUP BY clause in SQL?

The GROUP BY clause is used to group rows that have the same values in specified columns into summary rows, like calculating aggregates (e.g., COUNT, SUM, AVG) for each group.

66. Explain the concept of indexing in databases.

Indexing is a technique used to speed up the retrieval of records from a database table. An index is created on one or more columns of a table, allowing the database to find data faster without scanning the entire table.

67. What are NoSQL databases, and how are they different from SQL databases?

NoSQL databases are non-relational databases designed to handle large volumes of unstructured or semi-structured data. Unlike SQL databases, which use structured schemas and tables, NoSQL databases use flexible data models such as key-value pairs, documents, graphs, or wide-columns. They are often used for big data and real-time web applications.