

TOP 50



INTERVIEW QUESTIONS FOR DATA ENGINEERS

Preparing for a data engineering interview?
These SQL questions will help you stand out!



Abhishek Agrawal
Data Engineer



SQL Basics

Q1. What is SQL? Explain its basic structure.

Q2. What are the differences between **SQL** and **NoSQL** databases?

Q3. Explain the difference between **DDL**, **DML**, and **DCL** commands.

Q4. How do you use the **SELECT** statement in SQL?

Q5. What is the difference between **WHERE** and **HAVING** clauses?

Q6. How do you use the **GROUP BY** clause in SQL?

Q7. Explain **NULL** values in SQL. How do you handle NULLs in queries?

Q8. What are **primary keys** and **foreign keys** in SQL?

Q9. What is the difference between **INNER JOIN** and **OUTER JOIN**?

Q10. How do you use **UNION** and **UNION ALL** in SQL?



SQL Functions and Operators

Q11. What are **aggregate functions** in SQL? Give examples.

Q12. How do you use **string functions** like **CONCAT()**, **UPPER()**, and **LOWER()** in SQL?

Q13. What is the difference between **RANK()**, **DENSE_RANK()**, and **ROW_NUMBER()**?

Q14. Explain the use of **COALESCE()** and **NULLIF()** functions.

Q15. What is the difference between **COUNT()**, **SUM()**, **AVG()**, **MIN()**, and **MAX()**?

Q16. How do you use **CASE** statements in SQL?

Q17. How do you use **LIKE** and **ILIKE** for pattern matching?

Q18. What are **date functions** in SQL? Provide examples.

Q19. How do you extract **year, month, and day** from a date in SQL?

Q20. What is the difference between **BETWEEN** and **IN** operators in SQL?



SQL Joins and Subqueries

Q21. What are the different types of **joins** in SQL?

Q22. Explain the difference between **CROSS JOIN** and **SELF JOIN**.

Q23. How do you write a **subquery** in SQL?

Q24. What is the difference between **correlated** and **non-correlated subqueries**?

Q25. How do you optimize **subqueries** for better performance?

Q26. How do you perform a **join on multiple tables**?

Q27. What are **common table expressions (CTEs)** in SQL?

Q28. How do you use **nested subqueries**? Provide an example.

Q29. What is the **EXISTS** operator in SQL?

Q30. What is the difference between **INNER JOIN** and **LEFT JOIN** with examples?



SQL Performance Optimization

Q31. What are **indexes** in SQL, and how do they improve performance?

Q32. What are the differences between **clustered and non-clustered indexes**?

Q33. How do you **optimize a slow SQL query**?

Q34. What is the **EXPLAIN or EXPLAIN PLAN** command in SQL?

Q35. How do you reduce **query execution time**?

Q36. What is **query optimization**, and why is it important?

Q37. How do you handle **large datasets** in SQL?

Q38. What is **sharding** in SQL databases?

Q39. How do you avoid **deadlocks** in SQL transactions?

Q40. What are **best practices for writing efficient SQL queries**?



SQL Advanced Topics

Q41. What are **stored procedures** in SQL? How do you create them?

Q42. How do you use **triggers** in SQL? Provide an example.

Q43. What are **views** in SQL? How do they improve query performance?

Q44. What is the difference between temporary **tables and CTEs**?

Q45. How do you use **transactions** in SQL?

Q46. What is **ACID compliance** in SQL databases?

Q47. Explain the concept of **data normalization**.

Q48. What is **denormalization**, and when should you use it?

Q49. How do you handle **recursive queries** in SQL?

Q50. What are **materialized views**, and how are they different from regular views?



Conclusion

Bonus Tip:

Make sure you know how to write queries in real-time during interviews.

✅ "Ready to ace your next SQL interview?"
"Bookmark this cheat sheet to practice frequently asked questions and impress your interviewer!"

**Follow for more
content like this**



Abhishek Agrawal

Azure Data Engineer

