Fanatics

Data Engineering Interview Questions

Ankita Gulati

Shubh Goyal

Job Details

- **Position:** Senior Data Engineer
- **Experience:** 6+ years
- **Location:** Bangalore
- **Work mode:** Hybrid
- **Compensation:** ₹50+ LPA
- **Total Rounds:** 4
- **Top Required Skills:**

1. SQL
2. PySpark / Python
3. Cloud Data Engineering
4. ETL / Data Modeling
5. Big Data & Streaming
6. System Design

Ankita Gulati

Shubh Goyal

Round 1
Technical Screening

1. Walk me through your current role and responsibilities as a Data Engineer.
2. Can you describe an end-to-end data pipeline you've built recently? What were the challenges?
3. What types of datasets do you typically work with (volume, variety, velocity)?
4. Write a query to find the second highest salary from an employee table.
5. Given an orders table, find all customers who placed orders in consecutive months.
6. Write a query to calculate the rolling 7-day average of sales.
7. Reverse a string in Python without using in-built functions.
8. Implement a function to check if two strings are anagrams without sorting.
9. Write a Python script to count the frequency of words in a list of strings.

Ankita Gulati

Shubh Goyal

10. Explain what an ETL pipeline is and how you monitor failures.

11. What are some best practices in designing transformations in SQL?

12. Difference between batch processing and streaming.

Ankita Gulati

Shubh Goyal

Round 2
SQL, Python & Architecture

1. Explain one of your major projects in detail. What was the architecture, and why did you choose those tools?
2. How do you handle large datasets efficiently (partitioning, bucketing, sharding)?
3. What are the trade-offs between star schema and snowflake schema?
4. How do you ensure scalability when query volumes increase significantly?
5. From the orders table, write a query to find the number of unique customers who placed more than two orders in the last month.
6. Given customers and orders tables, write a query to fetch customer names who haven't placed any orders.
7. Write a query to find employees whose salary is higher than their manager's salary.

Ankita Gulati Shubh Goyal

8. Find the first non-repeating character in a string.
9. Implement a Python generator that yields numbers in the Fibonacci sequence.
10. Given a list of integers, group them into sublists where each sublist has numbers with the same frequency count.
11. What is data partitioning in cloud warehouses like Snowflake or BigQuery? How does it improve performance?
12. Explain how you would optimize queries in Redshift or BigQuery.
13. How do you handle schema evolution in streaming pipelines?

Round 3
System Design

1. Design a task scheduler that manages task execution, retries, and failures.
2. How would you design the database schema to store task metadata and states (pending, running, completed, failed)?
3. How would you design the system to handle dependencies between tasks?
4. What would change if the system had to handle 10x more tasks?
5. How do you ensure high availability and fault tolerance?
6. Design a scalable data ingestion system to collect real-time clickstream data and make it available for analytics in near real-time.
7 How would you design a data lineage and governance framework in a large organization?

Ankita Gulati

Shubh Goyal

8. What caching strategies would you use for frequently accessed but rarely updated data?

9. How would you design a fraud detection pipeline for online transactions using both batch and streaming?

Ankita Gulati

Shubh Goyal

Round 4
HR & Behavioral

1. Tell me about a time when you faced a production failure. How did you handle it?
2. How do you prioritize tasks when you have multiple urgent requests?
3. Describe a project where you worked closely with data scientists or business stakeholders. How did you manage communication?
4. How do you ensure data quality in your pipelines? (validations, monitoring, tests)
5. Have you mentored junior engineers? How do you approach knowledge sharing within your team?
6. Why do you want to join Fanatics?
7. What do you expect from this role in terms of challenges and growth?
8. Where do you see yourself in 3–5 years as a Data Engineer/Lead?

Ankita Gulati

Shubh Goyal

Thank You

Best of luck with your upcoming interviews — you've got this!

Ankita Gulati

Shubh Goyal