Data Engineering Interview Questions

Ankita Gulati

Shubh Goyal

Job Details

- **Position:** Senior Data Engineer
- **Experience:** 4+ years
- **Location:** Pan India
- **Work mode:** Remote
- **Compensation:** ₹16-20 LPA
- **Total Rounds:** 4
- **Top Required Skills:**

1. SQL
2. PySpark / Python / Databricks
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 most recent project using Azure and Spark. What was the architecture and your specific contributions?
2. How did you implement data ingestion, transformation, and orchestration in your project?
3. What were the key challenges you faced while using Azure Data Lake and Spark?
4. Write a SQL query using window functions to fetch the top 3 salaries in each department.
5. Given a sales table, write a query to calculate the running total of sales per region using window functions.
6. Write a query to find customers who purchased in 3 consecutive months.
7. Explain how window functions are used in PySpark. Can you provide an example?

Ankita Gulati Shubh Goyal

8. How would you implement a CASE WHEN statement in PySpark?

9. What is the difference between selectExpr() and withColumn() in PySpark?

10. Difference between DataFrame and RDD in Spark.

11. Explain the role of Azure Data Factory in an end-to-end pipeline.

12. What are some key differences between Delta Lake and a traditional data lake?

Ankita Gulati

Shubh Goyal

Round 2
SQL, Spark & Azure

1. Explain the roles and responsibilities you handled in your previous projects.
2. Which orchestration tool did you use? Why did you choose Azure Data Factory?
3. How do you handle data quality and validation in Azure pipelines?
4. Difference between INNER JOIN, LEFT JOIN, and FULL OUTER JOIN with examples.
5. Write a query to find employees who earn more than the average salary in their department.
6. Write a query using window functions to rank orders by sales within each category.
7. What are broadcast joins in Spark? When would you use them?
8. Explain repartition vs coalesce in Spark. Which one is more efficient and why?

Ankita Gulati Shubh Goyal

9. How do you handle data skew in Spark jobs? Provide real-world techniques.

10. What are data partitioning strategies in Azure Data Lake?

11. How do you optimize ADF pipelines for large-scale data processing?

12. How do you monitor and debug failed jobs in Databricks or ADF?

Ankita Gulati

Shubh Goyal

Round 3
Databricks + Scenario-Based

1. What is Unity Catalog in Databricks?
2. How does it help with data governance and auditing?
3. How do you implement access control at the table, schema, and row level?
4. Explain how schema evolution works in Delta Lake.
5. What are MERGE operations in Delta Lake? Provide an example use case.
6. How does Delta Lake ensure ACID transactions in a data lake environment?
7. You have a pipeline failing due to schema mismatch between raw and processed data. How would you fix it?
8. You are asked to provide role-based access to different teams in Databricks. How would you design the governance layer?
9. How would you build a slowly changing dimension (SCD Type 2) table in Delta Lake?

Ankita Gulati Shubh Goyal

10. Design a scalable ingestion pipeline to bring streaming data (IoT logs) into Azure Data Lake and make it available for analytics.

11. How would you design a multi-region data pipeline to ensure low latency for global users?

12. What caching strategy would you use for frequently accessed but rarely updated data in Databricks?

Ankita Gulati

Shubh Goyal

Round 4
HR Discussion

1. Walk me through your resume and highlight the most relevant experiences.
2. What excites you about working at Wipro as a Senior Data Engineer?
3. What are your role expectations for this position?
4. If you receive multiple offers, how will you evaluate them?
5. What is your preferred location and joining timeline?
6. Where do you see yourself in the next 3–5 years in terms of career growth?

Ankita Gulati

Shubh Goyal

Thank You

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

Ankita Gulati

Shubh Goyal