



Myntra

Data Engineering Interview Questions



Ankita Gulati

Shubh Goyal



Job Details

- **Position:** Senior Data Engineer
- **Experience:** 5+ years
- **Location:** Bangalore
- **Work mode:** Office
- **Compensation:** ₹40+ 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

Round 1

Telephonic Screening

Interviewer: Data Engineering Manager

1. How would you optimize a data pipeline for performance in Spark?
2. Tell me about a time you had to work with incomplete or ambiguous data. How did you handle it?
3. What strategies do you use to ensure data accuracy in large-scale systems?

Round 2

Technical Interview

Interviewer: Senior Data Engineer

1. Design a data pipeline to process terabytes of user interaction data on Spark. How would you ensure scalability?

2. What are the most common performance tuning techniques in Spark?

- Example: shuffling optimization, broadcast joins, caching, Catalyst optimizer.

3. Walk through the steps to create a fault-tolerant data pipeline.

4. How would you extend your pipeline design to handle streaming data?

- Discuss Spark Streaming, checkpointing, and idempotent writes.

Round 3

Data Pipeline Debugging

1. If a pipeline breaks at a specific stage, how would you locate and fix the issue?
2. How would you automate monitoring and alerting for your data pipelines?
3. Design a system to identify and fix backend source pipe issues that affect downstream pipelines.
4. Build a Spark job that can process both batch and streaming data, and explain how you'd incorporate validation checks, logging, and unit tests.

Round 4

Behavioral Interview

1. Tell me about a time when you had to explain technical concepts to non-technical stakeholders. How did you ensure clarity?
2. How do you handle pressure in a fast-paced environment like Myntra?
3. Give an example of a time you failed. What did you learn from the experience, and how did you improve afterward?
4. How do you approach collaboration with product managers, data scientists, and business teams to ensure alignment?

Thank You

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

