Data Engineering
Interview Questions

Ankita Gulati

Shubh Goyal

Job Details

- **Position:** Data Engineer
- **Experience:** 4+ years
- **Location:** Remote
- **Work mode:** –
- **Compensation:** ₹35+ 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 Interview 1

SQL & Query Optimization
1. Write a SQL query to retrieve data from a database.
2. Optimize SQL queries for time and space complexity.

Python & Algorithms
3. Implement a Python function to manipulate data or solve a basic algorithm problem.

Resume-Based Questions
4. Explain key technical experiences and challenges mentioned in your resume.

Ankita Gulati   Shubh Goyal

Round 2
Technical Interview II

SQL Challenges
1. Write a SQL query to calculate the average salary of employees by department.

Python Problem-Solving
2. Implement a Python function to find the unique element in a sorted array.

Project-Based Discussion
3. Discuss your previous projects in detail, including:
- Challenges faced.
- How you diagnosed and resolved them.
- Performance and scalability improvements achieved.

Ankita Gulati   Shubh Goyal

Round 3
System Design

Data Pipeline Design
 1. Design a data pipeline to ingest livestream data.
 • Consider scalability, latency, and reliability.

IoT Data Processing
 2. Propose a scalable pipeline for analyzing large-scale IoT datasets.
 • Cover ingestion, storage, and real-time/near real-time processing.

Ankita Gulati   Shubh Goyal

Round 4
Managerial Discussion

Project & Problem-Solving
1. Walk me through a complex project you worked on and how you solved key problems.

Behavioral Insights
2. Share an example of working in a cross-functional team or resolving conflict.

Cross-AZ Data Transfer
3. What strategies would you use for cross-AZ data transfer?
 • Focus: low latency, fault tolerance, encryption, and access control.

Spark Optimizations
4. Explain key Spark optimization techniques, including:
 • Partitioning , Caching, Join optimization, Memory tuning & Reducing shuffle overhead.

Ankita Gulati   Shubh Goyal

Thank You

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

Ankita Gulati   Shubh Goyal