Myntra

Data Engineering
Interview
Questions

Ankita Gulati   Shubh Goyal

Job Details

- **Position:** Data Engineer
- **Experience:** 3+ years
- **Location:** Bangalore
- **Work mode:** Office
- **Compensation:** ₹30+ 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
SQL, Data Modeling & Python

1. Introduce yourself and explain your role in recent projects.
2. What motivated you to apply for Myntra?
3. What is the difference between the HAVING clause and the WHERE clause?
4. Explain the use of a SELF JOIN with an example.
5. Provide use cases for WINDOW functions in SQL.
6. True/False: Indexing always improves query performance. Justify your answer.
7. In a given scenario, how would you decide between using fact tables and dimension tables?
8. Explain the implementation of Slowly Changing Dimension (SCD) Type 4.
9. Write a SQL query to track historical changes in data for a given table.
10. Discuss different data loading strategies for incoming data (e.g., full load, incremental load, CDC).

Ankita Gulati   Shubh Goyal

11. Write a Python program to:
• Calculate the total amount spent by each user.
• Identify the top 5 users based on spending.
• Find the most frequently purchased product.

Ankita Gulati   Shubh Goyal

Round 2
Technical Discussion

1. Walk me through a past project you've worked on.
   • What were the main challenges?
   • How did you diagnose and resolve them?

2. Let's revisit some of your unsolved technical test questions — explain your solution approach for each

.

3. Solve three medium-level SQL problems, including one involving two tables with values and NULLs.
   • Show the outputs for LEFT JOIN, RIGHT JOIN, and INNER JOIN operations.
   • Explain why the results differ.

Ankita Gulati   Shubh Goyal

Round 3
System Design & Engineering

Apache Spark Fundamentals
1. Explain the role of cores, executors, jobs, and stages in Spark.
2. Differentiate between transformations vs. actions in Spark with examples.
3. What are the differences between REPARTITION and COALESCE?
4. Explain optimization techniques in Spark for large datasets.

File Formats & Storage
5. Compare Delta Lake vs. Parquet file formats.
6. What is Z-Ordering, and how does it improve query performance?
7. How would you implement incremental loading using Delta file formats?

Ankita Gulati   Shubh Goyal

Join Strategies & SQL
 8. Compare different JOIN strategies in Spark (broadcast join, shuffle join, etc.).
 9. Solve an advanced SQL scenario using Common Table Expressions (CTEs) and conditional joins.

Ankita Gulati   Shubh Goyal

Round 4
Hiring Manager Discussion

1. Discuss your past projects and explain the biggest challenges you faced.
2. How have you leveraged Databricks services in your projects? Provide specific examples.
3. Why do you want to switch from your current company to Myntra?
4. How do you see yourself contributing to Myntra's data platform and analytics initiatives?

Ankita Gulati   Shubh Goyal

Thank You

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

HIRED

Ankita Gulati   Shubh Goyal