Paytm

Data Engineering
Interview Questions

Ankita Gulati

Shubh Goyal

Job Details

- **Position:** Data Engineer
- **Experience:** 2+ 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
Introduction & Background

Introduction & Background
 1. Walk me through your past experience, background, and projects.

SQL Questions
 2. Given two tables, perform an INNER JOIN on columns A & B and return the total row count of the output.
 • (Hint: Consider Null Safe Join for handling null values)
 3. A cricket scoreboard schema was provided. Write a set of ad-hoc SQL queries to extract insights (e.g., highest scorer, strike rate, team total).

DSA Questions
 4. Given a list of numbers, find all pairs whose sum equals a given value X.
 5. Given a list of strings, find all anagrams within the list.

Ankita Gulati   Shubh Goyal

Round 2
Spark + PySpark + DSA

Project & Spark Fundamentals
1. Explain your previous projects in detail. What technologies and tools did you use?
2. Explain fundamentals of Apache Spark.
 • What happens in case of job failure?
 • How do you perform Spark job optimization?

PySpark Coding Exercise
3. Given a dataset with 4–5 columns and a problem statement:
 • Write PySpark code (on paper) to solve the problem using transformations and actions.

DSA Question
4. Write code to reverse a linked list.

Ankita Gulati   Shubh Goyal

Round 3
Techno-Managerial

Architecture Discussion

1. Walk me through the architecture of projects you have worked on.
 • Explain components, scalability, and design trade-offs.

Behavioral Questions

2. If you faced a conflict with a teammate during a project, how did you resolve it?
3. Have you ever missed project delivery timelines?
 • If yes, how did you communicate it to leadership and stakeholders?

System Design

4. Design an Online Ticket Booking System.
 • Cover system components, database schema, concurrency handling, and fault tolerance.

Ankita Gulati   Shubh Goyal

Round 4
HR Discussion

1. Discussion about salary expectations and negotiations.

2. Why are you switching jobs?

3. What are you looking for in your next role at Paytm?

Ankita Gulati   Shubh Goyal

Round 5
System Design

1. Design a data platform to ingest, process, and analyze telemetry data from millions of Nutanix clusters worldwide.
 • What ingestion strategies would you use (batch vs streaming)?
 • Which storage solutions would you recommend (HDFS, S3, NoSQL, data lake)?
 • How would you structure real-time vs offline analytics?
 • How would you integrate machine learning for predictive maintenance?

Ankita Gulati   Shubh Goyal

Thank You

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

HIRED

Ankita Gulati   Shubh Goyal