AMERICAN EXPRESS

Data Engineering
Interview
Questions

Ankita Gulati

Shubh Goyal

Job Details

- **Position:** Data Engineer
- **Experience:** 2-3 years
- **Location:** Gurgoan
- **Work mode:** Office
- **Compensation:** ₹20+ 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 & Python Coding

SQL Questions:

1. Consecutive Numbers:
• Table: id | num
• Write a query to find numbers that appear 3 times consecutively.
• Example Output: 1 (since "1" appears three times in a row).

2. Core SQL Topics Tested:
• Window functions (ROW_NUMBER(), DENSE_RANK()).
• Recursive CTEs.
• Subqueries and joins.

Python Questions:

3. Coding Tasks:
• Manipulation of lists and dictionaries (lookups, counting, transformations).

4. Theoretical Questions:
• What are decorators in Python?
• Explain multiprocessing with an example.

Ankita Gulati   Shubh Goyal

Round 2
Big Data & PySpark

PySpark Coding:
 1. Given a dataset, write transformations using DataFrame API (not Spark SQL).
 • Apply filters, joins, and window functions.

Spark Concepts:
 2. Optimization techniques:
 • Repartition vs. Coalesce.
 • Broadcast joins.
 • Salting for skew handling.
 • Debugging OutOfMemory (OOM) issues (partitioning, executor memory tuning, GC settings).

Advanced Spark Topics:
 3. Bloom Filters: What they are and how they work in Spark.
 4. Difference between Spark Context vs. Spark Session.
 5. Spark memory management: handling overhead heap memory issues.

Ankita Gulati   Shubh Goyal

Round 3
Techno-Managerial

Project Discussion:
 1. Explain one project in detail, using tools like Snowflake, Databricks, Spark, Airflow, AWS.
 • Draw the ETL flow: Ingestion → Transformation → Loading into Snowflake.

Scenario-Based Questions:
 2. What challenges did you face?
 • Infrastructure-related vs. data-volume-related issues.
 3. How did you optimize Spark jobs? (Broadcast joins, repartitioning, caching).
 4. How do you handle auto-scaling in Databricks clusters?

Behavioral / Leadership:
 5. How do you manage multitasking when tasks exceed available time?

Ankita Gulati   Shubh Goyal

6. If a critical dependency is blocked - (teammate on leave), how do you still deliver?

7. Cloud Cost Optimization:

• Share strategies to reduce cost while still meeting SLAs.

Ankita Gulati   Shubh Goyal

Round 4
HR Discussion

Topics Discussed:
1. Current salary and expected salary negotiation.
2. Notice period and joining flexibility.
3. Cultural alignment with American Express values.

Ankita Gulati   Shubh Goyal

3. Behavioral Questions:
 • Tell me about a time you handled production failure under tight deadlines.
 • How do you prioritize when multiple teams demand data simultaneously?

Ankita Gulati

Shubh Goyal

Round 5
HR Round

Discussion Topics:
- Career aspirations and long-term goals.
- Fit within Adobe's data platform teams.
- Salary expectations and negotiation.
- Work culture, flexibility, and alignment with Adobe values.

Ankita Gulati

Shubh Goyal

Thank You

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

HIRED

Ankita Gulati   Shubh Goyal