adidas

Data Engineering Interview Questions

Ankita Gulati

Shubh Goyal

Job Details

- **Position:** Data Engineer
- **Experience:** 3+ years
- **Location:** Gurgaon
- **Work mode:** Hybrid
- **Compensation:** ₹17–22LPA
- **Total Rounds:** 4
- **Top Required Skills:**

1. Apache Spark Internals
2. Advanced SQL (Window Functions, Joins, Aggregations)
3. Core Python (Data Structures & Problem Solving)
4. Data Engineering Best Practices
5. ETL Pipeline Design & Scalability

Ankita Gulati

Shubh Goyal

Round 1
HR Screening

Details:

- Duration: 30 minutes
- Mode: Telephonic

Topics:

1. The interviewer asked the candidate to walk through their professional background, covering current role, past experiences, and major projects.
2. The candidate was asked to describe how their technical skills and prior experiences align with the responsibilities of the Data Engineer role at Adidas.

Difficulty Level: Easy

Ankita Gulati

Shubh Goyal

Round 2
Technical Deep Dive

1. Explain in detail how a Spark job executes internally. Start from when an action is triggered and describe the creation of the DAG, how stages are formed, how tasks are scheduled across executors, and how Spark handles failures and retries.
2. Write an SQL query to find the top three selling products for each brand in the last 30 days. You must partition the data by brand and rank the products within each brand based on total sales, using window functions.
3. Write an SQL query to calculate the average salary percentage change for each department over the past year. Ensure the solution accounts for departments with missing monthly records and handles division-by-zero safely.
4. Given a list of integers, write a Python function that moves all zero values to the end of the list while maintaining the relative order of the non-zero elements. For example, [0, 1, 0, 3, 12] should be transformed to [1, 3, 12, 0, 0].
5. Given a list of floating-point numbers, write a Python function that rounds each number to the nearest integer. For example, [1.2, 2.7, 3.5] should become [1, 3, 4].

Ankita Gulati

Shubh Goyal

Round 3
Scenario-Based Problem Solving

1. Suppose your data ingestion pipeline normally ingests 100 GB of data per hour, but suddenly the incoming data volume spikes to 1 TB per hour. How would you redesign your ingestion strategy to ensure stable performance and prevent pipeline failure? Discuss whether you would scale horizontally, partition data differently, or optimize transformations.
2. Imagine that business requirements for your ETL pipelines change frequently, such as new fields being added to the schema, transformations being updated, or downstream tables needing modifications. How would you design a resilient pipeline that can handle these changes with minimal downtime?
3. How would you implement monitoring and alerting for a production pipeline at Adidas? What metrics (e.g., throughput, latency, error rates) would you monitor, and what tools or frameworks would you use to trigger alerts and scale resources dynamically?

Ankita Gulati

Shubh Goyal

Round 4

Final HR + Director Discussion

1. The HR and Director asked the candidate to explain their understanding of the Data Engineer role at Adidas and what they expect in terms of day-to-day responsibilities.
2. The conversation included an overview of Adidas' work culture, team collaboration practices, and leadership expectations.
3. HR asked the candidate to confirm their genuine interest in the position and alignment with their career goals.
4. Salary expectations, notice period, and a potential joining date were discussed to align on logistics.

Ankita Gulati

Shubh Goyal

Thank You

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

Ankita Gulati

Shubh Goyal