verizon

# Data Engineering Interview Questions

Ankita Gulati

Shubh Goyal

# Job Details

- **Position:** Senior Data Engineer
- **Experience:** 5+ years
- **Location:** Pune
- **Work mode:** Bangalore
- **Compensation:** ₹25+ LPA
- **Total Rounds:** 3
- **Top Required Skills:**

1. SQL
2. PySpark / Python / Databricks
3. Cloud Data Engineering
4. ETL / Data Modeling
5. Big Data & Streaming
6. System Design

Ankita Gulati

Shubh Goyal

# Round 1
# Technical Discussion

1. Explain the difference between clustering and partitioning in data warehousing. Provide use cases for each.
2. What are Slowly Changing Dimensions (SCDs)? Explain the different types with real-world examples.
3. Compare data lakes vs. data warehouses. What are the advantages and disadvantages of each?
4. Design a schema for a swipe-payment API to support high availability and scalability.
5. Describe the ETL process end-to-end (Extract, Transform, Load) with an example pipeline.
6. How do you optimize SQL queries for better performance? Discuss indexing, partitioning, and execution plans.
7. SQL Coding: Write a query to find the top 3 highest revenue-generating customers in the past 6 months.

Ankita Gulati

Shubh Goyal

8. SQL Coding: Given an orders table with order_id, customer_id, order_date, amount, write a query to calculate month-over-month revenue growth.

9. Implement the Next Greater Element algorithm for an array.

10. Write an algorithm to search in a rotated sorted array.

11. Find the K-th element across two sorted arrays efficiently.

12. Traverse a binary tree in zigzag (spiral) order.

Ankita Gulati

Shubh Goyal

# Round 2
# Advanced Technical Discussion

1. How does Hadoop MapReduce work? What are its limitations?
2. Explain Apache Spark and its advantages over Hadoop MapReduce.
3. How would you handle data skew in Spark joins?
4. Explain Spark optimizations (caching, partitioning, broadcast joins).
5. Explain Apache Kafka and its role in real-time streaming pipelines.
6. How would you design a Kafka + Spark Streaming pipeline for clickstream analytics?
7. How do you ensure exactly-once message processing in a streaming architecture?
8. Walk me through designing a data pipeline and schema to support address lookup or payment flow.
9. Given a case study (e.g., loading pricing options), break it down into ingestion, transformations, storage, and serving layers.

Ankita Gulati Shubh Goyal

10. How would you design a fault-tolerant, highly available data pipeline on AWS/GCP/Azure?
11. Explain how you would implement data governance and access control in a multi-tenant data lake (e.g., Unity Catalog, Lake Formation, Ranger).
12. Python Coding: Write a function to flatten a deeply nested JSON into key-value pairs.
13. Python DSA: Write a function to find the first non-repeating character in a string.
14. SQL Problem: Write a query to get the customers who placed orders in every month of the last year.
15. System Design: How would you design a recommendation system using customer transaction and browsing data?

Ankita Gulati

Shubh Goyal

# Round 3
# HR Discussion

1. Tell me about yourself and your journey as a Data Engineer.
2. Why do you want to join Verizon?
3. Describe a time when you handled a production failure in a data pipeline. How did you resolve it?
4. How do you prioritize tasks when managing multiple urgent requests?
5. Share an example of when you collaborated with cross-functional teams (data scientists, business stakeholders).
6. What are your short-term and long-term career goals?
7. What are your salary expectations and notice period?

# *Thank You*

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


Ankita Gulati

Shubh Goyal