

**NUTANIX**

# Data Engineering Interview Questions



Ankita Gulati

Shubh Goyal



# Job Details

- **Position:** Data Engineer
- **Experience:** 2+ years
- **Location:** Bangalore
- **Work mode:** Office
- **Compensation:** ₹25+ LPA
- **Total Rounds:** 6
- **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

# Online Assessment (HackerRank)

1. Optimize an inefficient data pipeline – identify bottlenecks and propose improvements.
2. Design a schema for a complex tracking system (covering normalization, indexing, and scalability).
3. Write a script to clean and transform messy data into a structured, usable format.

Ankita Gulati

Shubh Goyal

# Round 2

# Technical Phone Interview

1. Describe your experience with real-time data processing. Which tools and architectures have you worked with?
  
2. How would you design and build a data warehouse from scratch?
  - Discuss schema design, ETL strategies, and handling evolving requirements.
  
3. Explain MapReduce in simple, non-technical terms as if explaining to a business stakeholder.

# Round 3

## Data Modeling

1. Design a data model for Nutanix's global supply chain management system.
  - Include fact and dimension tables.
  - Discuss Slowly Changing Dimensions (SCDs).
  - Address handling of rapidly changing attributes.
2. Would you incorporate NoSQL elements for flexibility? Why or why not?

# Round 4

## Coding

1. Build a real-time analytics pipeline using Kafka, Spark Streaming, and Elasticsearch.
  - Handle ingestion, processing, and indexing.
  - Discuss trade-offs in latency vs throughput.

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?

# Round 5

## Team & Culture Fit

1. Share a challenging experience from your career. How did you overcome it, and what did you learn?
2. Suggest potential new features for Nutanix products.
3. How would you propose an AI-driven data governance tool? What problems would it solve, and how would you implement it?

Ankita Gulati

Shubh Goyal

Thank You

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



Ankita Gulati

Shubh Goyal