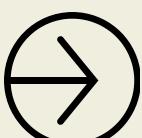# Grab

Data Engineering
Interview Questions

Ankita Gulati

Shubh Goyal

Job Details

- **Position:** Lead Data Engineer
- **Experience:** 8+ years
- **Location:** Singapore (HQ)
- **Work mode:** Hybrid
- **Compensation:** $150K
- **Total Rounds:** 6
- **Top Required Skills:**

1. Advanced SQL
2. Python / Scala
3. Big Data Frameworks
4. Cloud Platforms
5. Data Modeling & Warehousing
6. ETL Orchestration
7. System Design for high-scale pipelines
8. Leadership & stakeholder management

Ankita Gulati   Shubh Goyal

Round 1
Initial Screening

1. Tell me about yourself and your current role.
2. Why do you want to join Grab? How do you connect with our mission (heart, hunger, honor, humility)?
3. What's your experience in building scalable data pipelines?
4. Which cloud technologies are you strongest in?

Ankita Gulati

Shubh Goyal

Round 2
Technical Assessment

1. SQL:
   a. Write a query to find the top 3 users by ride count in the last month.
   b. Write a query to calculate the conversion rate for Grab Ads.
   c. Identify high-value users based on ride frequency (use window functions).
2. Coding:
   a. Optimize a Python script that processes billions of ride records (vectorization, multiprocessing).
   b. Write a Python function to detect anomalies in ride data (e.g., unusually long trips).
   c. Write a Scala program for parallel data processing on Spark.

Ankita Gulati   Shubh Goyal

Round 3
System Design Interview

1. Design a system to calculate the average speed of all cars in a city in real-time.
    a. Which data ingestion service? (Kafka/Kinesis)
    b. How to store aggregated metrics (Redshift/Snowflake/BigQuery).
2. How would you structure a data warehouse for Grab's food delivery service?
    a. Fact: Orders (order_id, user_id, restaurant_id, delivery_time, order_value).
    b. Dimensions: Customer, Restaurant, Driver, Date, Payment.
3. Design a fraud detection pipeline for Grab rides in near real-time.
    a. Must handle high throughput, ensure fault tolerance.

Ankita Gulati   Shubh Goyal

Round 4
Deep Dive Technical Interview

1. Spark Optimization:
   a. How would you optimize a Spark job processing terabytes of data?
   b. Difference between Spark persist vs checkpoint – when to use which?
   c. Handling OutOfMemory errors in Spark executors.
2. Kafka/Streaming:
   a. How do you manage streaming data pipelines with Kafka?
   b. How does Kafka ensure consistency in distributed systems?
3. Cloud & Architecture:
   a. Which AWS services would you use for a real-time analytics pipeline?
   b. Pros & cons of serverless data processing (AWS Lambda vs EMR/Spark).

Ankita Gulati

Shubh Goyal

Round 5
Behavioral & Leadership Round

1. Tell me about a time you faced conflict within a team. How did you resolve it?
2. How do you prioritize competing business requirements from product vs. analytics teams?
3. Describe a challenging project and how you ensured its success.
4. What strategies do you use to mentor junior engineers?

Ankita Gulati   Shubh Goyal

Round 6
Final Interview

1. What do you see as future challenges in data engineering (scaling, governance, cost optimization)?
2. How would you contribute to Grab's mission using data?
3. If you join, what first 90 days roadmap would you set as Lead Data Engineer?

Ankita Gulati

Shubh Goyal

Thank You

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

HIRED

Ankita Gulati   Shubh Goyal