

# Google Data Engineer Interview Experience Guide

A data engineering candidate recently received an offer from Google's Bangalore office for an L3/L4 Data Engineer role. This document outlines their background, application process, interview experience, and preparation tips.

## Candidate Background

- **Total Experience:** 3 years post-graduation
- **Domain Expertise:** Data Engineering
- **Skillset:** ETL Pipelines, Python, SQL, Airflow, Snowflake, GCP
- **Previous CTC:** ₹9 LPA (Fixed)

## Application Process

The candidate applied via the Google Careers Portal. Shortly afterward, they were contacted on LinkedIn by a Google recruiter regarding the same opportunity.

## Offer Details

Component	Value
Base Salary	₹20 LPA
Annual Bonus	₹3 LPA
Stocks (RSUs)	₹15–20 LPA per year (vested over 4 years)
<b>Total Compensation</b> ₹35–40 LPA (Realistic for L3/L4 in India)	

## Interview Process Overview

The candidate went through a **6-round interview process**, covering technical depth, problem-solving, system design, and behavioral fit.

## Round 1: Online Assessment (90 Minutes)

### Areas Covered:

- DSA: One medium-level LeetCode-style problem involving sliding window error processing
- SQL: Joining large tables using window functions for aggregation and filtering
- Python: Writing a function to detect data schema anomalies in logs

### Preparation Tips:

- Focus on LeetCode (medium-level) and HackerRank (Advanced SQL)
- Understand window functions, schema validation, and log-based anomaly detection

## Round 2: Data Engineering Round

### Topics:

- Design a daily ETL pipeline to ingest API data into BigQuery
- Handle schema evolution in production
- Compare Airflow's @daily vs once trigger scheduling
- Optimize a slow SQL query provided during the interview

### Preparation Tips:

- Be ready to sketch data pipelines with failure handling, partitioning, and retry logic
- Practice ETL architecture using cloud-native tools like BigQuery, Airflow, and GCS

## Round 3: Coding Round

### Task:

- Process a large log file (in GBs) to identify the top 10 users by event frequency
- Optimize for memory efficiency and handle streaming input

### Follow-up:

- Handle midstream schema changes gracefully

### Preparation Tips:

- Practice generator functions, external sorting, and stream-based transformations
- Ensure familiarity with large file parsing and efficient in-memory computation

## Round 4: System Design Round

### Scenario:

- Design a real-time data pipeline for clickstream events

### Questions:

- How to ensure fault tolerance
- Where to implement deduplication logic
- How to efficiently store 1 billion+ rows

### Relevant Tools & Concepts:

- Pub/Sub, Kafka
- Dataflow, Spark
- BigQuery, Z-ordering, partition strategies

## Round 5: Behavioral (Googliness + Stakeholder Management)

### Questions Asked:

- A situation where the candidate took ownership of a failing project
- Handling conflict with a product manager
- Actions taken when a deadline was missed due to code errors

### Preparation Tips:

- Use the STAR format (Situation, Task, Action, Result)
- Show humility, collaboration, and a growth mindset

## Round 6: Hiring Manager Round

### Focus Areas:

- Deep dive into the candidate's recent projects
- Building a data pipeline from scratch
- Onboarding and mentoring junior engineers
- Defining success metrics for pipelines

### Final Questions:

- Motivation to join Google
- What excites the candidate about working there

**Preparation Tips:**

- Clearly articulate reasons for joining Google (scale, impact, learning)
- Align project outcomes with business impact and operational excellence

**Glassdoor Google Review –**

<https://www.glassdoor.co.in/Reviews/Google-Reviews-E9079.htm>

**Google Careers –**

<https://www.google.com/about/careers/applications/>

**Subscribe to my YouTube Channel for Free Data Engineering Content –**

<https://www.youtube.com/@shubhamwadekar27>

**Connect with me here –**

<https://bento.me/shubhamwadekar>

**Checkout more Interview Preparation Material on –**

[https://topmate.io/shubham\\_wadekar](https://topmate.io/shubham_wadekar)