**Bitwise Data Engineer Interview Guide – Experienced 2+**

**Round 1: Technical Interview**

**1. Introduction & Project Overview**

The interviewer asked me to introduce myself and provide an overview of my completed projects, focusing on the tools, technologies, and challenges.

**Example Questions:**

 Describe a data pipeline you built and optimized.

 What challenges did you encounter when scaling your project?

 How did you ensure data quality and integrity?

**2. SQL Questions**

 **Question 1:** Find the 2nd highest salary for each department using the

DENSE\_RANK() function.

 **Follow-up:** Why not use ROW\_NUMBER() instead? Discuss pros and cons.

 **Question 2:** Calculate the cumulative transaction amount for each month using a transaction table.

**Key Concepts Tested:**

 DENSE\_RANK() vs. ROW\_NUMBER() usage.

 CTEs and window functions (SUM() OVER for cumulative sums).

 Handling NULL values in joins.

**3. Spark & PySpark Questions**

 Discussed Spark cluster management and optimization techniques.

 Solved the 2nd highest salary problem using PySpark.

**Example Questions:**

 How do you manage memory allocation in Spark?

 How would you handle skewed data in a Spark job?

 Write PySpark code to filter and count records.

**Round 2: Client Interview (Gartner)**

**1. Introduction & In-Depth Project Discussion**

The focus was on the architecture, design choices, and contributions to my projects.

**Example Questions:**

 Why did you choose a particular data storage solution?

 How did you ensure fault tolerance in your pipelines?

**2. AWS Questions**

 **Lambda vs. Glue:** Discussed use cases for both services.

 **S3 Storage Options:** Described Standard, Intelligent-Tiering, and Glacier.

 **Step Functions:** Explained orchestration of workflows.

 Optimizing long-running PySpark scripts on EMR.

**Key Concepts:**

 AWS services for data processing.

 Cost optimization strategies and performance tuning.

**3. SQL Questions**

 Predicted outputs for different join types using two sample tables with NULL values.

 Query to find employees earning more than their managers.

**Round 3: Managerial Interview**

**1. Introduction & Project Summary**

The manager asked for a summary of my experience and a deeper discussion about key projects.

**Example Questions:**

 What motivates you to pursue a change in your career?

 Can you share a time you faced a significant challenge and how you overcame it?

 What lessons have you learned from project failures?

**Preparation Tips:**

1. **Master SQL:** Focus on window functions, complex joins, and handling NULL values.

2. **Hands-on with PySpark:** Practice transformations, optimizations, and writing efficient code.

3. **AWS Knowledge:** Understand core services like Lambda, Glue, EMR, and cost management strategies.

4. **Behavioral Skills:** Be ready to discuss challenges, teamwork, and leadership experiences.