Tiger Analytics Data Engineer Interview Guide – Experienced 3+

This interview experience details my journey from application to offer for a Data Engineer role, highlighting each step of the process. The experience was both challenging and insightful, testing a broad range of skills from coding and problem-solving to architectural understanding and technical communication. Here's a breakdown of each stage of the process.

Application Stage

I applied for this role through Naukri.com, and to my surprise, I received a call from the recruiter the very next day. The initial conversation was straightforward, focusing on collecting key information such as:

- Total and relevant work experience
- Current and expected compensation
- Notice period availability

The recruiter provided a high-level overview of the role and next steps, which gave me a clear sense of the expectations and timeline.

Round 1: Coding Round (60 Minutes)

This was a time-pressured technical assessment with coding questions that spanned Python, SQL, and Apache Spark. The level of difficulty was above moderate, with an emphasis on both problem-solving efficiency and code optimization.

Key Challenges and My Approach

- **Spark Coding**:
    - One task required counting occurrences of elements in a list of tuples using Spark RDDs.
    - Instead of groupByKey (which can be inefficient), I leveraged the more efficient combineByKey transformation to compute averages of key-value pairs without sacrificing performance.
    - The challenge tested not only my knowledge of Spark but also my ability to use transformations optimally.
- **SQL Problems**:
    - I solved two complex SQL queries involving multiple table joins, Common Table Expressions (CTEs), and window functions.
    - One query required optimizing performance with window functions for calculating running totals, while another focused on joining tables with common single-column keys.
    - These problems demonstrated the importance of clear logic and efficient query design.

- **Python Coding**:
    - One question tested my ability to flatten nested lists recursively, which I implemented using recursion for dynamic depth handling.
    - Another required concatenating lists within a range, which I solved using concise list comprehensions.

In addition to the coding tasks, theory-based questions covered Hadoop architecture, Spark transformations and optimizations, and general SQL best practices. This round pushed my ability to think quickly and write clean, efficient code under time constraints.

Round 2: Technical Discussion (60 Minutes)

This round, conducted by an architect-level interviewer, delved deep into both conceptual understanding and practical applications.

Structure and Key Areas Covered

1. **Project Architecture Discussion**

   The interviewer asked me to walk through my current project architecture. I outlined the data flow from ingestion to transformation and storage, explaining how we handle scalability and fault tolerance.

Probing questions focused on:

- Why I chose specific technologies (e.g., Spark over traditional ETL tools)
- How we manage dependencies and retries in data pipelines
- Techniques for optimizing Spark jobs

2. **SQL Challenges**

   I encountered two SQL problems similar to those in the coding round, but this time, the focus was on explaining my thought process.

   I broke down each step of my query design, justifying choices like CTEs for readability and window functions for cumulative calculations.

3. **PySpark Coding**

   The next segment tested my PySpark knowledge with real-world transformation scenarios.

   - I demonstrated how to partition data effectively and avoid shuffling for better performance.
   - Discussed repartition vs. coalesce and how to optimize stages in a Spark DAG (Directed Acyclic Graph).

4. **Apache Spark Architecture**
   We had a detailed discussion on:
   - DAG creation and task execution
   - Stages, tasks, and partitioning strategies
   - Lineage graphs for fault tolerance

The interviewer was thorough but encouraging, providing constructive feedback and highlighting areas of strength.

Round 3: HR Round

The final step was a relaxed yet professional conversation with HR.

Topics covered:

- Basic introduction and career aspirations
- Location preferences and salary negotiations
- A walkthrough of the company's policies, culture, and benefits

The HR representative was clear about the offer timeline and next steps. I received the official job offer within a week, with a joining date scheduled to align with my notice period.

Glassdoor Tiger Analytics Review –

https://www.glassdoor.co.in/Reviews/Tiger-Analytics-Reviews-E717029.htm

Tiger Analytics Careers –

https://www.tigeranalytics.com/about-us/current-openings/

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