

Accenture Azure Data Engineer Interview Guide – Experienced 3+

Interview Rounds Overview

1. Round 1 & 2 – Technical

- Focus: Technical skills in SQL, Python, Azure Data Factory, and PySpark.

2. Round 3 – Managerial

- Focus: Team dynamics, project ownership, and stakeholder collaboration.

3. Round 4 – HR Discussion

- Focus: Salary negotiation and notice period.

Detailed Breakdown of Rounds

Round 1 & Round 2 – Technical

Key Questions and Topics:

1. Introduction and Experience

- Briefly introduce yourself and walk us through your journey as a Data Engineer so far.
- Current project details: technologies, data architecture, and responsibilities.
- Average data volume handled and strategies for efficient processing.
- Challenges faced in your projects and how you overcame them.

2. Azure Data Factory (ADF)

- What is ADF? Is it an ETL or ELT tool, and why?
- Explain a linked service and how to create one.
- Difference between linked service and dataset in ADF.
- Integration Runtimes (IR): Types and use cases.
- Triggers in ADF, especially tumbling window triggers.
- Moving pipelines from development to production: ARM templates for deployment.

3. SQL

- Write a query to find the second-highest salary in a table.

Solution:

```
SELECT MAX(salary) AS second_highest_salary
FROM employees
WHERE salary < (SELECT MAX(salary) FROM employees);
```

4. Python

- Function to find the top 3 largest numbers in a list.

Solution:

```
def top_3_largest(numbers):
    return sorted(numbers, reverse=True)[:3]
```

5. Data Validation and Schema Management

- Handling data validation using SQL or Python.
- Managing schema changes in PySpark over time.

6. PySpark and Spark

- Why is RDD considered resilient and fault-tolerant?
- Lazy evaluation in Spark and its impact on performance.
- Difference between persist() and cache() in Spark.
- Difference between reduceByKey() and groupByKey().
- DataFrames vs. RDDs in PySpark.
- What are the key differences between DataFrames and RDDs in PySpark?
- How do you manage schema changes in PySpark when processing data over time?

7. Azure Databricks

- What is a mount point in Azure Databricks? How to mount ADLS Gen2 to Databricks?

Round 3 – Managerial

This round assessed my ability to work within a team, take ownership of projects, and collaborate effectively with stakeholders.

Key Questions:

1. Professional Background

- Can you introduce yourself and provide a brief overview of your career?
- Team size and roles of other members in your project.

2. Project Architecture

- Describe the architecture of your current project and your role in its design and implementation.

3. Data Warehouse Design

- Star Schema vs. Snowflake Schema: Which one you implemented and why?

4. SQL Analytical Functions

- Differences between ROW_NUMBER(), RANK(), and DENSE_RANK() and their usage in project scenarios.

4. Stakeholder Collaboration

- Strategies for smooth communication between data scientists, business teams, and developers.

Round 4 – HR Discussion

The HR discussion was primarily focused on administrative and personal aspects.

Key Topics:

- Salary negotiations.
- Notice period and availability.

Glassdoor Accenture Review –

<https://www.glassdoor.co.in/Reviews/Accenture-Reviews-E4138.htm>

Accenture Careers –

<https://www.accenture.com/in-en/careers>

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

© Shubham Wadekar