accenture

Data Engineering
Interview
Questions

Ankita Gulati

Shubh Goyal

Job Details

- **Position:** Azure Data Engineer
- **Experience:** 3+ years
- **Location:** Pan India
- **Work mode:** Hybrid
- **Compensation:** ₹15–20 LPA
- **Total Rounds:** 4
- **Top Required Skills:**
  1. SQL
  2. PySpark / Python
  3. Cloud Data Engineering
  4. ETL / Data Modeling
  5. Big Data & Streaming
  6. System Design

Ankita Gulati   Shubh Goyal

Round 1
Foundations & ADF Basics

1. Can you briefly introduce yourself and walk us through your journey as a Data Engineer so far?

2. Tell us about your current project—technologies used, data architecture, and your daily responsibilities.

3. What's the average data volume you handle in your current project? How do you ensure efficient processing of such data?

4. Can you explain the challenges you've faced in your projects and how you overcame them?

5. What is Azure Data Factory (ADF)? Would you classify it as an ETL or ELT tool? Why?

6. What is a linked service in ADF? How do you create one?

7. Can you explain the difference between a linked service and a dataset in ADF?

8. What are Integration Runtimes (IR) in ADF? Can you explain the different types and their use cases?

Ankita Gulati   Shubh Goyal

9. What are triggers in ADF, and how have you used them in your project? Specifically, explain tumbling window triggers.

10. How do you migrate ADF pipelines from development to production? Have you used ARM templates for deployment?

11. Write an SQL query to find the second-highest salary from a table.

12. Given a Python list of numbers, write a function to return the top 3 largest numbers.

13. Can you explain how you perform data validation using SQL or Python in your projects?

Ankita Gulati   Shubh Goyal

Round 2
Spark, PySpark & Databricks

1. How do you manage schema changes in PySpark when processing data over time?
2. Why is RDD considered resilient and fault-tolerant? How does Spark ensure reliability?
3. What is lazy evaluation in Spark, and how does it impact performance?
4. Differentiate between persist() and cache() in Spark. When would you use each?
5. What is a mount point in Azure Databricks? How do you mount ADLS Gen2 to Databricks?
6. What is the difference between reduceByKey() and groupByKey() in Spark?
7. What are the key differences between DataFrames and RDDs in PySpark?
8. Write an SQL query to fetch the top 3 products with the highest sales per region from a sales table.
9. Write a Python program to count the frequency of words in a text file.

Ankita Gulati   Shubh Goyal

10. Write PySpark code to read data from a Parquet file, perform aggregations, and save it to ADLS Gen2.

11. Write Python code to generate all permutations of a given list.

12. Implement a Python program to process large CSV files efficiently (memory optimization).

13. Explain the CAP theorem in distributed systems.

14. What is eventual consistency and when is it acceptable?

Ankita Gulati

Shubh Goyal

Round 3
Technical / Managerial Discussion

1. Can you introduce yourself again and provide a brief overview of your professional background as a Data Engineer?

2. How large is your team, and what are the roles of other members?

3. Describe the architecture of your current project. What was your role in designing or implementing it?

4. What is the difference between a Star Schema and a Snowflake Schema? Which one have you implemented in your data warehouse projects, and why?

5. Explain the difference between ROW_NUMBER(), RANK(), and DENSE_RANK() in SQL. Can you provide use case scenarios for each?

6. How do you ensure smooth communication and collaboration between different stakeholders such as data scientists, business teams, and developers?

Ankita Gulati   Shubh Goyal

Round 4
HR & Behavioral

1. What are your salary expectations?
2. What is your notice period, and are you open to negotiations?
3. Why do you want to join Accenture as a Data Engineer?

Ankita Gulati   Shubh Goyal

Thank You

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

HIRED

Ankita Gulati   Shubh Goyal