

CloudPandith

Azure Databricks Intermediate | Delta Lake & Optimization

Real-Time Projects | Interview Ready Training

www.cloudpandith.com | 8904424822 | cloudpandith@gmail.com

🌀 Section 1: Core Spark Concepts & Architecture

1. Lecture 1: Spark Partitions Explained End-to-End | Architecture, Shuffle, Repartition vs Coalesce
2. Lecture 2: Spark Execution Plan Explained | Lazy Evaluation | Catalyst & Tungsten Optimizer
3. Lecture 3: Spark Join Algorithms Explained | Sort Merge vs Broadcast vs Shuffle Hash

🌀 Section 2: Spark Performance Optimization

4. Lecture 4: Spark Cache vs Persist vs Unpersist Explained
5. Lecture 5: Spark Cache vs Persist Hands-On | Tungsten Optimization | File Formats

🌀 Section 3: Python for PySpark ETL

6. Lecture 6: Python Control Flow in PySpark | If Else, Loops
7. Lecture 7: Python File Handling & Loops | break, continue
8. Lecture 8: Exception Handling in Python | try except finally

🌀 Section 4: Data Formats in Databricks

9. Lecture 9: Excel vs XML in Databricks | Read & Write in PySpark
10. Lecture 10: JSON Deep Dive | Flatten Nested JSON | explode, arrays_zip

www.cloudpandith.com | 8904424822 | cloudpandith@gmail.com

● Section 5: Delta Lake Fundamentals

11. Lecture 11: Delta Lake Explained | ACID Transactions
12. Lecture 12: Delta Table Creation | Managed vs External
13. Lecture 13: Delta Transformations | INSERT, UPDATE, DELETE, MERGE

● Section 6: Delta Lake Advanced Concepts

14. Lecture 14: Time Travel, Restore & Vacuum Explained
15. Lecture 15: Partitioning Explained | Best Practices
16. Lecture 16: Z-Ordering Explained | Performance Tuning
17. Lecture 17: Liquid Clustering Explained

● Section 7: SCD & Data Modeling

18. Lecture 18: SCD Type 1 Explained | MERGE, Hash Key
19. Lecture 19: SCD Type 2 Explained | Full History
20. Lecture 20: Star Schema Fact & Dimension Tables

● Section 8: Data Ingestion Pipelines

21. Lecture 21: Auto Loader Explained | Incremental Processing
22. Lecture 22: Idempotent Pipelines | COPY INTO vs MERGE

● Section 9: Cluster & Architecture

23. Lecture 23: Cluster Types Explained | Job, All-Purpose, Serverless
24. Lecture 24: ADLS Connection Explained | Optimization
25. Lecture 25: Cluster Sizing Explained | Memory & Partitions
26. Lecture 26: Medallion Architecture Explained | Bronze, Silver, Gold

● Section 10: Logging & Monitoring

27. Lecture 27: Databricks Logging to Azure SQL | ETL Logging

● Section 11: CI/CD

28. Lecture 28: Databricks CI/CD | Azure DevOps Integration

● Section 12: Interview Preparation

29. Lecture 29: Databricks Interview Questions & Answers

30. Lecture 30: PySpark Interview Questions | Real Scenarios