

Microsoft Certified Azure Databricks Data Engineer Associate Certification (Exam DP-750)

COURSE CONTENT

GET IN TOUCH



Multisoft Systems
B - 125, Sector - 2, Noida



(+91) 9810-306-956



info@multosoftsystems.com



www.multosoftsystems.com

About Multisoft

Train yourself with the best and develop valuable in-demand skills with Multisoft Systems. A leading certification training provider, Multisoft collaborates with top technologies to bring world-class one-on-one and certification trainings. With the goal to empower professionals and business across the globe, we offer more than 1500 training courses, which are delivered by Multisoft's global subject matter experts. We offer tailored corporate training; project Based Training, comprehensive learning solution with lifetime e-learning access, after training support and globally recognized training certificates.

About Course

Microsoft Certified: Azure Databricks Data Engineer Associate Certification (Exam DP-750) Training is designed to help professionals build scalable data engineering solutions using Azure Databricks and Apache Spark. The course covers Delta Lake, data ingestion, ETL/ELT pipelines, Spark SQL, streaming data, workflow orchestration, Unity Catalog, security, performance optimization, and data governance. Participants will gain practical experience in designing and managing enterprise data pipelines while preparing for the DP-750 certification.

Module 1: Introduction to Azure Databricks

- ✓ DP-750 Certification Overview
- ✓ Azure Databricks Architecture
- ✓ Lakehouse Fundamentals
- ✓ Workspace Configuration
- ✓ Databricks User Interface
- ✓ Workspace Administration
- ✓ Databricks Runtime
- ✓ Enterprise Use Cases

Module 2: Apache Spark Fundamentals

- ✓ Apache Spark Architecture
- ✓ Spark Clusters
- ✓ Spark DataFrames
- ✓ Spark SQL
- ✓ Spark Transformations
- ✓ Spark Actions
- ✓ Partitioning
- ✓ Performance Basics

Module 3: Azure Databricks Development

- ✓ Databricks Notebooks
- ✓ Python for Databricks
- ✓ SQL Development

- ✓ Notebook Collaboration
- ✓ Widgets
- ✓ Magic Commands
- ✓ Workspace Assets
- ✓ Source Control Integration

Module 4: Data Ingestion and ETL

- ✓ Batch Data Ingestion
- ✓ Incremental Loading
- ✓ Auto Loader
- ✓ Data Validation
- ✓ Data Cleansing
- ✓ ETL Pipelines
- ✓ ELT Workflows
- ✓ Data Transformation
- ✓ Error Handling
- ✓ Data Quality Checks

Module 5: Delta Lake

- ✓ Delta Lake Architecture
- ✓ ACID Transactions
- ✓ Schema Enforcement
- ✓ Schema Evolution
- ✓ Time Travel
- ✓ Delta Tables

- ✓ Delta Optimization
- ✓ Vacuum Operations
- ✓ Data Versioning
- ✓ Best Practices

Module 6: Advanced Data Processing

- ✓ Complex Transformations
- ✓ Joins and Aggregations
- ✓ Window Functions
- ✓ User Defined Functions (UDFs)
- ✓ Data Serialization
- ✓ Nested Data Processing
- ✓ Semi-Structured Data
- ✓ Performance Optimization

Module 7: Structured Streaming

- ✓ Streaming Architecture
- ✓ Structured Streaming
- ✓ Streaming Sources
- ✓ Streaming Sinks
- ✓ Checkpointing
- ✓ Watermarking
- ✓ Event Processing
- ✓ Fault Tolerance
- ✓ Streaming Optimization

Module 8: Workflow Orchestration

- ✓ Databricks Workflows
- ✓ Job Scheduling
- ✓ Pipeline Automation
- ✓ Task Dependencies
- ✓ Parameter Management
- ✓ Notifications
- ✓ Retry Policies
- ✓ Workflow Monitoring

Module 9: Data Governance and Security

- ✓ Unity Catalog
- ✓ Data Access Control
- ✓ Identity Management
- ✓ Credential Management
- ✓ Data Lineage
- ✓ Audit Logging
- ✓ Encryption
- ✓ Governance Best Practices
- ✓ Regulatory Compliance

Module 10: Performance Optimization

- ✓ Cluster Configuration
- ✓ Cluster Policies

- ✓ Spark Optimization
- ✓ Query Optimization
- ✓ Caching
- ✓ Partition Strategies
- ✓ Resource Scaling
- ✓ Cost Optimization
- ✓ Performance Monitoring

Module 11: Azure Integration

- ✓ Azure Data Lake Storage
- ✓ Azure Data Factory
- ✓ Azure Synapse Analytics
- ✓ Azure Event Hubs
- ✓ Azure Key Vault
- ✓ Power BI Integration
- ✓ Azure Monitor
- ✓ Microsoft Fabric Connectivity

Module 12: Production Operations

- ✓ Deployment Strategies
- ✓ Monitoring
- ✓ Logging
- ✓ Troubleshooting
- ✓ Backup Strategies
- ✓ Disaster Recovery

- ✓ Operational Best Practices
- ✓ Production Readiness

Module 13: Enterprise Data Engineering Project

- ✓ End-to-End Data Pipeline
- ✓ Lakehouse Implementation
- ✓ Streaming Analytics Project
- ✓ Enterprise Data Governance
- ✓ Performance Tuning
- ✓ Security Implementation
- ✓ DP-750 Certification Practice Labs
- ✓ Capstone Data Engineering Project