

# DataBricks Lakehouse Training

*COURSE CONTENT*

## GET IN TOUCH



Multisoft Systems  
B - 125, Sector - 2, Noida



(+91) 9810-306-956



info@multisoftsystems.com



www.multisoftsystems.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

The DataBricks Lakehouse Training by Multisoft Systems offers a cutting-edge educational experience designed to empower professionals with the knowledge and skills necessary to leverage the full potential of the DataBricks Lakehouse platform.

## **Module 1: Understand how to use and the benefits of using the Databricks Lakehouse Platform and its tools**

- ✓ Data Lakehouse (architecture, descriptions, benefits)
- ✓ Data Science and Engineering workspace (clusters, notebooks, data storage)
- ✓ Delta Lake (general concepts, table management and manipulation, optimizations)

## **Module 2: Build ETL pipelines using Apache Spark SQL and Python**

- ✓ Relational entities (databases, tables, views)
- ✓ ELT (creating tables, writing data to tables, cleaning data, combining and reshaping tables, SQL UDFs)
- ✓ Python (facilitating Spark SQL with string manipulation and control flow, passing data between PySpark and Spark SQL)

## **Module 3: Incrementally process data**

- ✓ Structured Streaming (general concepts, triggers, watermarks)
- ✓ Auto Loader (streaming reads)
- ✓ Multi-hop Architecture (bronze-silver-gold, streaming applications)
- ✓ Delta Live Tables (benefits and features)

## **Module 4: Build production pipelines for data engineering applications and Databricks SQL queries and dashboards**

- ✓ Jobs (scheduling, task orchestration, UI)
- ✓ Dashboards (endpoints, scheduling, alerting, refreshing)

## **Module 5: Understand and follow best security practices**

- ✓ Unity Catalog (benefits and features)
- ✓ Entity Permissions (data objects Privileges)