

Data Engineering with Spark Framework and Databricks Platform

Lyon (France)

24 - 28 August 2026

UK Training

PARTNER



Data Engineering with Spark Framework and Databricks Platform

Code: IT32 From: 24 - 28 August 2026 City: Lyon (France) Fees: 5900 Pound

Introduction

In a world where organizations generate and process vast volumes of data every second, data engineering has become the backbone of digital transformation. Data Engineering with Apache Spark and Databricks equips professionals with the ability to manage, transform, and optimize data pipelines at scale – ensuring accuracy, performance, and business agility.

This course provides an in-depth understanding of the tools and frameworks that power modern data platforms. It is tailored for executives, team leaders, and professionals working in sectors such as oil and gas, banking, telecommunications, government, and other data-intensive industries.

Participants will gain hands-on experience in building reliable data architectures, integrating multiple data sources, and leveraging the full potential of Apache Spark and Databricks to enable data-driven innovation and decision-making.

Course Objectives

By the end of this course, participants will be able to:

- Understand the architecture and ecosystem of Apache Spark and Databricks.
- Build and optimize scalable data pipelines for batch and real-time processing.
- Apply distributed computing principles for faster data handling.
- Integrate Databricks with various data sources and cloud platforms.
- Manage clusters, jobs, and notebooks efficiently.
- Ensure data quality, reliability, and governance across the data lifecycle.
- Use Delta Lake for structured data management and versioning.
- Implement performance optimization and troubleshooting techniques.
- Apply data engineering concepts to support analytics and AI initiatives.
- Build practical, production-ready solutions within Databricks.

Course Outlines

Day 1: Foundations of Data Engineering and Apache Spark

- Introduction to data engineering and its role in enterprise systems.
- Overview of Apache Spark and its ecosystem.
- Understanding Spark architecture: driver, executor, and cluster manager.
- Key Spark components: Spark SQL, Spark Streaming, and MLlib.
- Working with Resilient Distributed Datasets RDDs and DataFrames.
- Hands-on exercise: setting up Spark and running the first data transformation job.

Day 2: Working with Databricks Environment



- Introduction to Databricks and the concept of Unified Data Analytics.
- Creating and managing Databricks workspaces and clusters.
- Working with notebooks, jobs, and workflows.
- Integrating Databricks with cloud platforms Azure, AWS, or GCP.
- Connecting to external data sources and managing data ingestion.
- Practical exercise: building an ETL pipeline using Databricks.

Day 3: Advanced Data Processing and Optimization

- Advanced transformations using Spark SQL and DataFrame API.
- Real-time data processing with Structured Streaming.
- Implementing Delta Lake for data versioning and consistency.
- Performance tuning: caching, partitioning, and shuffle optimization.
- Managing job scheduling, cluster configuration, and monitoring.
- Lab session: optimizing a complex data pipeline for speed and reliability.

Day 4: Data Governance, Quality, and Security

- Understanding data governance principles and compliance requirements.
- Implementing data validation and data quality checks.
- Ensuring lineage, traceability, and auditing in Databricks.
- Managing access control and identity permissions.
- Encrypting data and securing sensitive information.
- Case study: enforcing governance policies in enterprise data environments.

Day 5: Real-World Implementation and Capstone Project

- Designing an end-to-end data engineering workflow.
- Integrating Databricks pipelines with analytics and BI tools.
- Streaming data from multiple sources and real-time dashboards.
- Capstone project: building and presenting a complete enterprise-grade data pipeline.
- Review session: lessons learned and best practices for production deployment.
- Final assessment and feedback discussion.

Why Attend this Course: Wins & Losses!

- Gain deep technical and strategic knowledge of Apache Spark and Databricks.
- Learn to build and optimize real-world, enterprise-scale data solutions.
- Improve your organization's ability to process and analyze data efficiently.
- Enhance your career with in-demand data engineering skills.
- Understand how to apply data governance and compliance best practices.
- Stay current with the latest cloud-based data technologies.
- Strengthen collaboration between IT, data, and business teams.
- Develop confidence in leading data transformation projects.

Conclusion





Data Engineering with Apache Spark and Databricks is more than a technical skill – it is a strategic capability for modern enterprises. This course equips professionals with the tools, methods, and frameworks to design robust, scalable, and secure data systems that empower intelligent decision-making.

Through practical exercises, real-world projects, and expert-led sessions, participants will develop the expertise needed to transform raw data into meaningful insights – driving innovation and sustainable digital growth within their organizations.

Head Office: +44 7480 775 526
Email: Sales@blackbird-training.com
Website: www.blackbird-training.com



Blackbird Training Clients



MANNAI Trading
Company WLL,
Qatar



Alumina Corporation
Guinea



Booking.com
Netherlands



Oxfam GB International
Organization,
Yemen



Capital Markets
Authority,
Kuwait



Waltersmith Petroman Oil Limited
Nigeria



Qatar National Bank
(QNB),
Qatar



Qatar Foundation,
Qatar



AFRICAN UNION ADVISORY
BOARD ON CORRUPTION,
Tanzania



KFAS
Kuwait



Reserve Bank of
Malawi,
Malawi



Central Bank of Nigeria
Nigeria



Ministry of Interior
Kingdom of Saudi Arabia
KSA



Mabruk Oil Company
Libya



Saudi Electricity
Company,
KSA



BADAN PENGELOLA
KEUANGAN Haji,
Indonesia



NATO
Italy



ENI CORPORATE
UNIVERSITY,
Italy



Gulf Bank
Kuwait



المؤسسة العامة للتأمينات الاجتماعية
General Organization for
Social Insurance
KSA



Defence Space Administration
Nigeria



National Industries
Group (Holding),
Kuwait



Hamad Medical
Corporation,
Qatar



USAID
Pakistan



STC Solutions,
KSA



North Oil company,



EKO Electricity



Oman Broadband



UNITED NATIONS
UN.



هيئة تنظيم الكهرباء - عمان
Authority for

UK Training
PARTNER



Blackbird Training Categories

Management & Admin

Entertainment & Leisure
Professional Skills
Finance, Accounting, Budgeting
Media & Public Relations
Project Management
Human Resources
Audit & Quality Assurance
Marketing, Sales, Customer Service
Secretary & Admin
Supply Chain & Logistics
Management & Leadership
Agile and Elevation

Technical Courses

Artificial Intelligence (AI)
Sustainability, ESG & Corporate Responsibility
Advanced Courses
Hospital Management
Public Sector
Special Workshops
Oil & Gas Engineering
Telecom Engineering
IT & IT Engineering
Health & Safety
Law and Contract Management
Customs & Safety
Aviation
C-Suite Training

