



London TDM

Artificial Intelligence and Data Science Training Courses

Course Venue: United Kingdom - London

Course Date: From 19 October 2025 To 23 October 2025

Course Place: London Paddington

Course Fees: 7,500 USD

Introduction

In today's data-driven world, efficient data engineering and robust pipeline architecture are critical for extracting value from data at scale. This 5-day professional course is designed to provide participants with the foundational knowledge and practical skills required to design, build, and manage data pipelines that are scalable, reliable, and optimized for performance. Participants will learn about the tools, technologies, and best practices that are essential for successful data engineering and pipeline development.

Objectives

- Understand the fundamentals of data engineering and pipeline architecture.
- Learn to design scalable and efficient data pipelines.
- Gain hands-on experience with modern data processing tools and technologies.
- Master techniques for data transformation, cleansing, and enrichment.
- Ensure data quality and integrity throughout pipeline processes.

Course Outlines

Day 1: Introduction to Data Engineering

- Overview of data engineering and its role in data-driven organizations.
- Key components of a data pipeline and their functions.
- Introduction to data ingestion: Batch vs. streaming data.
- Data storage solutions: Relational databases, NoSQL, and data lakes.
- Tools and technologies: An exploration of essential data engineering tools.

Day 2: Designing Data Pipelines

- Best practices for designing scalable and reliable data pipelines.
- Understanding data flow and dependencies in pipelines.
- Pipeline orchestration with tools like Apache Airflow and Luigi.
- Managing and monitoring data pipelines for performance and reliability.
- Case studies of successful data pipeline architectures.

Day 3: Data Transformation and Processing

- Introduction to data processing frameworks: Apache Spark and Hadoop.
- Techniques for data cleansing, transformation, and enrichment.
- Working with ETL (Extract, Transform, Load) and ELT (Extract, Load, Transform) processes.
- Optimizing data processing for performance and cost efficiency.
- Hands-on lab: Building a data transformation pipeline.

Day 4: Ensuring Data Quality and Integrity

- Importance of data quality in data engineering and analytics.
- Techniques for validating and testing data across the pipeline.
- Implementing data governance and compliance measures.
- Using tools for data quality management and monitoring.

- Real-world examples of maintaining high data quality standards.

Day 5: Advanced Topics and Emerging Trends

- Exploring serverless data engineering and pipeline solutions.
- Introduction to real-time analytics and stream processing.
- Leveraging machine learning within data pipelines.
- Future trends in data engineering: Automation and AI-driven pipelines.
- Course wrap-up: Discussion, Q&A session, and next steps in your data engineering journey.