Advanced Data Engineering Architecture and Operations Training Course

1 – 5 February 2027
Addis Ababa

AGILE LEADERS
Training Center

AGILE LEADERS
Training Center

Advanced Data Engineering Architecture and Operations Training Course

Ref: 778_223996 **Date:** 1 – 5 February 2027 **Location:** Addis Ababa **Fees:** 4500 Euro

Overview

Advanced Data Engineering Architecture and Operations Training Course is a one-week advanced course for data engineering, analytics platform, cloud data, and technical leadership functions who leave with a Production Data Platform Architecture and Operations Pack. Participants connect pipeline design, batch and streaming ingestion, lakehouse and warehouse patterns, orchestration, data quality, observability, security, cost controls, and incident response. Agile Leaders Training Center develops practical advanced data engineering architecture and operations capability.

Who Should Attend

- Data engineering functions responsible for production ingestion and transformation pipelines
- Analytics platform functions responsible for shared processing and storage services
- Cloud data functions responsible for scalable and cost-controlled workloads
- Technical leadership functions responsible for architecture decisions and operating standards
- Reliability functions responsible for monitoring, recovery, and incident coordination

The course assumes participants can build data transformations, use cloud data services, read pipeline logs, and work with version control, and leaves out introductory coding, basic database administration, and certification exam preparation.

Departments and Industries

The course supports departments and industries that design and operate production data platforms.

- Banking data platforms and risk analytics
- Retail customer and supply-chain data engineering
- Healthcare analytics infrastructure
- Manufacturing telemetry and operations data
- Telecommunications network data services

Learning Objectives

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

- Evaluate warehouse, lakehouse, batch, and streaming architecture choices
- Build governed ingestion and transformation pipelines
- Apply orchestration, retry, and dependency patterns
- Use data quality controls and lineage evidence
- Diagnose reliability, security, performance, and cost issues
- Build a production data platform operations pack

Course Agenda

Day 1: Design the Platform Architecture

- Data Platform Capability and Workload Map
- Warehouse and Lakehouse Architecture Decision Record
- Batch and Streaming Processing Selection Matrix
- Data Domain Ownership and Interface Canvas
- Storage Format and Lifecycle Decision Guide

Day 2: Engineer Ingestion and Transformation

- Source-to-Target Pipeline Design Blueprint
- Change Data Capture and Incremental Load Pattern
- Schema Evolution and Compatibility Contract
- Idempotent Transformation and Replay Method
- Pipeline Versioning and Deployment Control

Day 3: Orchestrate Quality and Reliability

- Directed Acyclic Graph Dependency Model
- Retry Backoff and Failure Routing Policy
- Data Quality Expectation and Quarantine Rule Set
- Lineage Evidence and Reconciliation Record
- Service-Level Indicator and Error Budget Dashboard

Day 4: Govern Production Operations

- Least-Privilege Data Access Matrix
- Encryption and Secret Management Control Map
- Pipeline Observability Signal Catalog
- Workload Performance and Cost Review
- Data Incident Triage and Recovery Runbook

Day 5: Practice Platform Operations

- Exercise: Compare Warehouse and Lakehouse Decisions
- Exercise: Reconcile an Incremental Ingestion Pipeline
- Exercise: Diagnose an Orchestration Failure
- Exercise: Prioritize Reliability and Cost Corrections
- Capstone: Production Data Platform Architecture and Operations Pack

Practical Exercises

The course uses suggested activities based on banking, retail, healthcare, manufacturing, and telecommunications data workloads.

- Suggested activity: select architecture patterns from workload, latency, governance, and recovery requirements.
- Suggested activity: design an idempotent pipeline with schema controls, quality expectations, and reconciliation evidence.
- Suggested activity: analyze logs, metrics, dependencies, access controls, and cost signals during a production incident.
- Suggested activity: assemble architecture decisions, operating controls, and recovery actions into the final pack.

FAQs

Who suits advanced data engineering architecture training, and what does it assume?

Data engineering, platform, cloud data, reliability, and technical leadership functions suit the training; it assumes practical experience with transformations, data services, pipeline logs, and version control.

How does advanced data engineering differ from introductory data engineering?

Advanced data engineering focuses on architecture tradeoffs, production reliability, orchestration, observability, governance, cost, and incident recovery, while introductory training focuses on basic storage, transformation, and pipeline construction.

When should data engineers choose batch or streaming processing?

Data engineers should compare latency needs, event volume, ordering, replay, state management, source behavior, operating complexity, and cost before selecting batch, streaming, or a combined pattern.

How should production data pipeline reliability be measured?

Production reliability should be measured through freshness, completeness, accuracy, availability, latency, failure rate, recovery time, lineage coverage, and the service indicators agreed with data consumers.

AGILE LEADERS
Training Center

What belongs in a data engineering incident runbook?

A data engineering incident runbook should define detection signals, severity, ownership, containment, replay decisions, reconciliation, stakeholder communication, recovery validation, and follow-up control improvements.

Conclusion

Participants take back a Production Data Platform Architecture and Operations Pack linking architecture decisions, pipelines, orchestration, quality, lineage, observability, security, cost, and incident recovery. It changes disconnected engineering choices into an operating model for production data services. The pack supports coordinated decisions across engineering, platform, reliability, governance, and business data functions.