



Big Data Analysis and Decision Intelligence Training Course

23 – 27 May 2027
Johannesburg

Big Data Analysis and Decision Intelligence Training Course

Ref: 786_159761 **Date:** 23 – 27 May 2027 **Location:** Johannesburg **Fees:** 4500 Euro

Overview

Big Data Analysis and Decision Intelligence Training Course is a five-day intermediate course for data analysts, business intelligence functions, data engineers, transformation teams, and decision-support specialists who leave with a Big Data Analysis and Decision Evidence Pack. Participants connect big data characteristics, distributed data processing, batch analytics, streaming analytics, data pipelines, scalable queries, data quality, visualization, and big data governance. Agile Leaders Training Center develops practical big data analysis capability.

Who Should Attend

- Data analysis functions responsible for large-scale preparation and analytical evidence
- Business intelligence functions responsible for scalable queries and decision reporting
- Data engineering functions responsible for distributed pipelines and processing patterns
- Digital transformation functions responsible for analytical use cases and platform choices
- Decision-support functions responsible for interpreting and communicating large-data insights

The course assumes participants can query structured data, interpret descriptive analytics, and follow data-flow logic, and leaves out introductory programming, infrastructure administration, platform certification, and production system deployment.

Departments and Industries

The course supports departments and industries that analyze high-volume, high-velocity, or varied data.

- Financial services data and risk analytics
- Retail and digital customer analytics
- Manufacturing and sensor-data operations
- Telecommunications network analytics
- Healthcare service data analysis
- Public-service digital transformation

Learning Objectives

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



- Analyze big data characteristics and analytical use cases
- Compare distributed storage and processing patterns
- Build batch and streaming pipeline designs
- Apply scalable preparation and query methods
- Evaluate data quality, governance, and processing risks
- Build visual decision evidence from large datasets

Course Agenda

Day 1: Frame Big Data Analysis

- Big Data Characteristics and Use Case Canvas
- NIST Big Data Reference Architecture Stakeholder Map
- Distributed Storage and Compute Pattern Matrix
- Analytical Workload and Latency Profile
- Decision Question and Evidence Definition Sheet

Day 2: Design Data Pipelines

- Source-to-Consumption Data Flow Map
- Batch Ingestion and Transformation Pipeline
- Streaming Event and Processing Pipeline
- Schema and Metadata Definition Register
- Data Lake Zone and Lifecycle Model

Day 3: Perform Scalable Analysis

- Distributed Partitioning and Parallel Processing Plan
- Spark SQL Scalable Query Pattern
- Spark Structured Streaming Analysis Pattern
- Kafka Streams Transformation and Aggregation Pattern
- Analytical Result Validation Checklist

Day 4: Govern and Communicate Evidence

- Big Data Quality Dimension Scorecard
- Lineage and Processing Traceability Map
- Security and Privacy Risk Register
- Large-Scale Visualization Selection Matrix
- Decision Insight and Limitation Brief



Day 5: Practice Big Data Decisions

- Exercise: Classify a Big Data Use Case and Architecture
- Exercise: Design Batch and Streaming Data Pipelines
- Exercise: Select Scalable Query and Processing Patterns
- Exercise: Validate Quality, Governance, and Visual Evidence
- Capstone: Big Data Analysis and Decision Evidence Pack

Practical Exercises

The course uses suggested activities based on financial transactions, retail journeys, industrial sensors, and service operations.

- Suggested activity: classify volume, velocity, variety, processing, and decision requirements for a selected use case.
- Suggested activity: map batch and streaming data from source through transformation, analysis, and consumption.
- Suggested activity: compare distributed query and stream-processing patterns for latency, state, and validation needs.
- Suggested activity: assemble quality, lineage, risk, visualization, and decision limitations in the final evidence pack.

FAQs

Who suits big data analysis training, and what does it assume?

Data analysis, business intelligence, data engineering, transformation, and decision-support functions suit the course; it assumes structured-query experience, descriptive analytics, and data-flow literacy.

How does big data analysis differ from general data analytics training?

Big data analysis emphasizes distributed storage, parallel processing, batch and streaming pipelines, scalable queries, and governance challenges that arise when traditional single-system approaches are insufficient.

When should big data analysis use batch or streaming processing?

Batch processing suits bounded data and scheduled analysis, while streaming processing suits continuously arriving events and decisions that require lower latency or ongoing state.

How do Spark and Kafka support big data analysis?

Spark supports distributed queries, transformations, and structured streaming analysis, while Kafka supports event transport and stream-processing patterns; the course uses them as examples rather than product-certification subjects.



How is big data decision evidence validated?

Big data decision evidence is validated through schema controls, quality checks, lineage, processing tests, analytical reconciliation, visual review, risk documentation, and explicit interpretation limits.

Conclusion

Participants take back a Big Data Analysis and Decision Evidence Pack linking use cases, architecture, pipelines, scalable queries, streaming patterns, quality, governance, visualization, and interpretation. It changes disconnected large-data activities into a traceable analysis workflow. The pack supports coordinated decisions across analysis, engineering, business intelligence, transformation, and operational functions.