Real-Time Data Pipelines Training Course

Professional Development Training Course Outline

Category	Business Intelligence	**Duration**	5 Days
Base Fee	$1,500 USD	**Delivery Mode**	Online / On-site Hybrid

Course Introduction

In today’s fast-paced digital world, organizations face the challenge of processing massive volumes of data in real time to derive actionable insights. Real-Time Data Pipelines Training Course equips participants with advanced knowledge and practical skills to design, implement, and manage end-to-end data pipelines that support dynamic analytics and data-driven decision-making. Participants will learn cutting-edge technologies, tools, and frameworks that streamline data ingestion, transformation, and delivery, ensuring high performance and scalability. This course is ideal for professionals looking to bridge the gap between data engineering and real-time analytics.

This comprehensive training emphasizes hands-on experience, integrating theoretical concepts with practical applications. Learners will explore case studies, real-world scenarios, and industry best practices to understand the critical role of data pipelines in modern organizations. By the end of the program, participants will possess the expertise to optimize real-time data flows, improve data quality, and enable faster insights, making them invaluable assets to any data-driven organization.

Programme Curriculum

Real-Time Data Pipelines Training Course

Introduction

In today’s fast-paced digital world, organizations face the challenge of processing massive volumes of data in real time to derive actionable insights. Real-Time Data Pipelines Training Course equips participants with advanced knowledge and practical skills to design, implement, and manage end-to-end data pipelines that support dynamic analytics and data-driven decision-making. Participants will learn cutting-edge technologies, tools, and frameworks that streamline data ingestion, transformation, and delivery, ensuring high performance and scalability. This course is ideal for professionals looking to bridge the gap between data engineering and real-time analytics.

This comprehensive training emphasizes hands-on experience, integrating theoretical concepts with practical applications. Learners will explore case studies, real-world scenarios, and industry best practices to understand the critical role of data pipelines in modern organizations. By the end of the program, participants will possess the expertise to optimize real-time data flows, improve data quality, and enable faster insights, making them invaluable assets to any data-driven organization.

Course Objectives

1. Understand the architecture and components of real-time data pipelines using trending frameworks.
2. Learn to ingest, process, and deliver streaming data efficiently.
3. Explore advanced data processing techniques with Apache Kafka, Spark Streaming, and Flink.
4. Gain expertise in designing scalable and fault-tolerant data pipelines.
5. Implement ETL and ELT workflows in real-time scenarios.
6. Optimize pipeline performance using monitoring and tuning strategies.
7. Integrate real-time data pipelines with cloud-based platforms like AWS, Azure, and GCP.
8. Apply data governance, quality, and security best practices.
9. Leverage machine learning models in streaming data applications.
10. Build dashboards and visualization for real-time insights.
11. Solve complex data engineering challenges using open-source tools.
12. Explore containerization and orchestration for pipeline deployment using Docker and Kubernetes.
13. Analyze real-world case studies to implement data-driven solutions effectively.

Organizational Benefits

- Accelerated data-driven decision-making.
- Enhanced operational efficiency through optimized data pipelines.
- Reduced downtime with fault-tolerant and scalable architectures.
- Improved data quality and governance.
- Cost-efficient management of streaming data workflows.
- Faster deployment of analytics and reporting solutions.
- Increased competitiveness through real-time insights.
- Empowered teams with hands-on skills in modern data tools.
- Streamlined integration with cloud infrastructure.
- Ability to implement advanced machine learning in real-time analytics.

Target Audiences

1. Data Engineers seeking to enhance real-time processing skills.
2. Data Analysts aiming to integrate streaming data into analytics workflows.
3. Business Intelligence professionals.
4. IT Professionals managing enterprise data solutions.
5. Machine Learning Engineers working with real-time models.
6. Cloud Architects implementing scalable pipelines.
7. Software Developers focused on data-centric applications.
8. Project Managers overseeing data integration and analytics projects.

Course Duration: 5 days

Course Modules

Module 1: Introduction to Real-Time Data Pipelines

- Overview of streaming data and batch processing differences
- Components of a modern data pipeline
- Key technologies and tools in the real-time ecosystem
- Challenges and best practices in pipeline design
- Case study: Real-time analytics implementation for e-commerce platform
- Hands-on exercise: Building a simple data ingestion pipeline

Module 2: Data Ingestion Techniques

- Understanding streaming vs batch ingestion
- Apache Kafka fundamentals and architecture
- Data ingestion from multiple sources
- Handling schema evolution in streams
- Case study: IoT data ingestion pipeline
- Hands-on exercise: Implementing Kafka producer and consumer

Module 3: Real-Time Data Processing

- Introduction to Spark Streaming and Apache Flink
- Data transformation and aggregation in real time
- Windowing and event time processing
- Handling late-arriving data and duplicates
- Case study: Fraud detection in financial transactions
- Hands-on exercise: Stream processing with Spark Structured Streaming

Module 4: Pipeline Orchestration and Workflow Management

- Airflow and Luigi for workflow automation
- Scheduling and monitoring pipelines
- Dependency management and pipeline failure handling
- Implementing retry and alert mechanisms
- Case study: Automated ETL pipeline orchestration
- Hands-on exercise: Setting up DAGs in Apache Airflow

Module 5: Cloud Integration and Deployment

- Real-time pipelines on AWS, GCP, and Azure
- Serverless data processing with AWS Lambda & Azure Functions
- Cloud storage and streaming services
- Deployment strategies for production-ready pipelines
- Case study: Cloud-based streaming analytics for retail
- Hands-on exercise: Deploying a pipeline in AWS

Module 6: Data Governance, Security, and Compliance

- Data privacy and compliance standards
- Role-based access control and encryption
- Monitoring and auditing data pipelines
- Data lineage and metadata management
- Case study: GDPR-compliant pipeline implementation
- Hands-on exercise: Securing Kafka topics and Spark jobs

Module 7: Real-Time Analytics and Visualization

- Building dashboards with Tableau, Power BI, and Grafana
- Integration of streaming data with visualization tools
- KPI tracking and alerting systems
- Case study: Real-time monitoring of manufacturing data
- Hands-on exercise: Creating live dashboards from streaming data
- Hands-on exercise: Analyzing pipeline performance metrics

Module 8: Advanced Topics and Case Studies

- Machine learning in streaming pipelines
- Anomaly detection and predictive analytics
- Event-driven architectures and microservices integration
- Optimization and scaling of real-time pipelines
- Case study: Predictive maintenance in IoT pipelines
- Hands-on exercise: Deploying ML models in a streaming pipeline

Training Methodology

- Interactive instructor-led sessions with practical demonstrations
- Hands-on labs with real-world datasets
- Group discussions and collaborative problem-solving
- Real-time case studies analysis and solutions
- Continuous assessment through exercises and mini-projects
- Personalized feedback and troubleshooting sessions

Register as a group from 3 participants for a Discount

Send us an email: info@fineskilltrainingcenter.org or call *+254769199797*

Certification

Upon successful completion of this training, participants will be issued with a globally- recognized certificate.

Tailor-Made Course

We also offer tailor-made courses based on your needs.

Key Notes

a. The participant must be conversant with English.

b. Upon completion of training the participant will be issued with an Authorized Training Certificate

c. Course duration is flexible and the contents can be modified to fit any number of days.

d. The course fee includes facilitation training materials, 2 coffee breaks, buffet lunch and A Certificate upon successful completion of Training.

e. One-year post-training support Consultation and Coaching provided after the course.

f. Payment should be done at least a week before commence of the training, to Fineskill Training Center account, as indicated in the invoice so as to enable us prepare better for you.

Upcoming Scheduled Sessions

Start Date	End Date	Location	Price
23 Nov 2026	27 Nov 2026	Online	$1,000
23 Nov 2026	27 Nov 2026	Physical	$1,500
23 Nov 2026	27 Nov 2026	Physical	$3,000
23 Nov 2026	27 Nov 2026	Physical	$3,000
23 Nov 2026	27 Nov 2026	Physical	$6,000
23 Nov 2026	27 Nov 2026	Physical	$4,500
23 Nov 2026	27 Nov 2026	Physical	$6,000
23 Nov 2026	27 Nov 2026	Physical	$6,000

Ready to enroll or request a customized program? Book online or contact us:

Register Online Now

Email: info@fineskilltrainingcenter.com | Website: www.fineskilltrainingcenter.com