

Matillion (Cloud Data Integration & ETL) Training

COURSE CONTENT

GET IN TOUCH



Multisoft Systems
B - 125, Sector - 2, Noida



(+91) 9810-306-956



info@multisoftsystems.com



www.multisoftsystems.com

About Multisoft

Train yourself with the best and develop valuable in-demand skills with Multisoft Systems. A leading certification training provider, Multisoft collaborates with top technologies to bring world-class one-on-one and certification trainings. With the goal to empower professionals and business across the globe, we offer more than 1500 training courses, which are delivered by Multisoft's global subject matter experts. We offer tailored corporate training; project Based Training, comprehensive learning solution with lifetime e-learning access, after training support and globally recognized training certificates.

About Course

Matillion Training (Cloud Data Integration & ETL) by Multisoft Systems is designed to equip professionals with the skills required to build, manage, and optimize modern data pipelines in cloud environments. This training focuses on leveraging Matillion's powerful ETL capabilities to extract, transform, and load data efficiently into leading cloud data warehouses such as Snowflake, Amazon Redshift, and Google BigQuery.

Module 1: Introduction to Cloud Data Integration & Matillion

- ✓ Overview of ETL and ELT concepts
- ✓ Cloud data integration fundamentals
- ✓ Introduction to Matillion platform
- ✓ Matillion architecture and components
- ✓ Use cases and industry applications

Module 2: Matillion Environment Setup

- ✓ Matillion instance setup on AWS/Azure/GCP
- ✓ User interface and navigation
- ✓ Project creation and environment configuration
- ✓ Managing users and roles
- ✓ Connecting to cloud data warehouses (Snowflake, Redshift, BigQuery)

Module 3: Data Extraction Techniques

- ✓ Extracting data from databases and flat files
- ✓ API-based data extraction
- ✓ Incremental and full data loads
- ✓ Handling different data formats (CSV, JSON, XML)
- ✓ Best practices for data ingestion

Module 4: Orchestration Jobs

- ✓ Introduction to orchestration jobs
- ✓ Scheduling and triggering workflows
- ✓ Using components like Run Transformation, Python Script, API Query
- ✓ Managing dependencies between jobs
- ✓ Logging and monitoring orchestration jobs

Module 5: Transformation Jobs

- ✓ Introduction to transformation jobs
- ✓ Data cleansing and enrichment techniques
- ✓ Using transformation components (Join, Aggregate, Filter, Calculator)
- ✓ Handling large datasets efficiently
- ✓ Designing reusable transformation logic

Module 6: Working with Cloud Data Warehouses

- ✓ Integration with Snowflake, Redshift, and BigQuery
- ✓ Data loading and staging strategies
- ✓ Optimizing queries and performance
- ✓ Managing schemas and tables
- ✓ Data partitioning and indexing concepts

Module 7: Advanced Data Integration

- ✓ API integrations and third-party connectors
- ✓ Handling semi-structured data (JSON, XML)
- ✓ Data blending from multiple sources
- ✓ Implementing CDC (Change Data Capture)
- ✓ Real-time and near real-time data processing

Module 8: Automation & Scheduling

- ✓ Job scheduling and dependency management
- ✓ Automating ETL workflows
- ✓ Event-based triggers
- ✓ Error handling and retry mechanisms
- ✓ Notifications and alerts setup

Module 9: Performance Optimization & Best Practices

- ✓ Performance tuning techniques
- ✓ Optimizing transformation jobs
- ✓ Resource management and scaling
- ✓ Data quality and validation checks
- ✓ Best practices for ETL pipeline design

Module 10: Security, Governance & Monitoring

- ✓ Data security in Matillion
- ✓ Role-based access control (RBAC)
- ✓ Data governance practices
- ✓ Monitoring job performance
- ✓ Audit logs and compliance

Module 11: Real-Time Project & Use Cases

- ✓ End-to-end ETL project implementation
- ✓ Building a complete data pipeline
- ✓ Industry-specific use cases
- ✓ Debugging and troubleshooting
- ✓ Deployment and production readiness