

# Google Dataflow and Matillion 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

Google Dataflow and Matillion training by Multisoft Systems equips professionals with the expertise to design and manage real-time data pipelines and ETL workflows. This comprehensive course focuses on Google Dataflow, a fully managed stream and batch processing service, and Matillion, a powerful cloud-native ETL tool used for seamless data integration.

## Module 1: Introduction to Google Dataflow and Matillion

- ✓ Understanding Dataflow and Matillion's roles in data processing
- ✓ Benefits of integrating Google Dataflow with Matillion

## Module 2: Introduction to Matillion and Snowflake

- ✓ Overview of Matillion ETL
- ✓ Overview of Snowflake data warehouse
- ✓ Benefits of using Matillion with Snowflake
- ✓ Key concepts and terminology

## Module 3: Setting up the Environment

- ✓ Configuring Snowflake environment
- ✓ Setting up Matillion instance for Snowflake
- ✓ Connecting Matillion to Snowflake
- ✓ Best practices for environment setup

## Module 4: Data Extraction

- ✓ Configuring Matillion to connect to various data sources
- ✓ Using Matillion's Extract components to pull data into Snowflake
- ✓ Handling different data formats (CSV, JSON, XML)
- ✓ Scheduling and managing data extraction jobs

## Module 5: Data Transformation in Matillion

- ✓ Overview of transformation components in Matillion
- ✓ Implementing common transformations (filtering, joining, aggregating)
- ✓ Advanced transformations (lookups, data cleansing, custom scripts)

## Module 6: Loading Data into Snowflake

- ✓ Configuring data load jobs to Snowflake
- ✓ Loading data efficiently and handling large datasets
- ✓ Managing data types and schema mapping
- ✓ Error handling and data quality considerations

## Module 7: Building and Managing Orchestration Jobs

- ✓ Introduction to orchestration components in Matillion
- ✓ Designing workflows and pipelines to automate ETL processes
- ✓ Scheduling and managing orchestration jobs
- ✓ Using Matillion variables and parameters

## Module 8: Advanced Snowflake Features and Integration

- ✓ Utilizing Snowflake's features (e.g., semi-structured data, streams, tasks)
- ✓ Leveraging Matillion's advanced components and functionalities with Snowflake
- ✓ Performance optimization techniques for ETL processes
- ✓ Advanced data integration patterns

## Module 9: Security and Governance

- ✓ Configuring user roles and permissions in Snowflake
- ✓ Implementing security best practices in Matillion and Snowflake
- ✓ Data encryption and compliance considerations
- ✓ Auditing and monitoring