

Using SAS Data Integration Studio to Create Efficient ETL Processes

Duration: 2.0 days

This course provides critical ETL performance, tuning, and capacity information for SAS 9 and SAS Data Integration Studio. The course introduces ETL concepts and covers ETL performance topics in depth and with examples as needed.

Learn how to

- analyze and debug ETL flows
- gain efficiency quickly
- set up the system environment
- customize advanced performance techniques.

Who should attend

SAS Data Integration Studio users who want to learn to improve performance of ETL processes

Prerequisites

Before attending this course, you should have attended the *SAS Programming 1: Essentials* course (or have equivalent experience). Knowledge of the SAS macro language is helpful. Also, a good understanding of SAS Data Integration Studio is necessary. This knowledge can be obtained by attending the *Using SAS Data Integration Studio to Build Data Marts from Enterprise Data Sources* course.

Course Contents

Tracking Efficiency in ETL Flows

- what is the SAS Intelligence Platform?
- important concepts for SAS Data Integration Studio
- measuring efficiencies

Basic ETL Performance Tuning

- working with intermediate files
- views or physical tables for ETL flows
- data quality considerations
- transformation and column optimization
- status handling and return code check

Working with Data

- keys
- indexes
- data functions
- counting rows
- load techniques

Sorts and Joins

- overview of process intensive tasks
- tuning sort performance
- optimizing join performance

Dimensional Models

- what is a star schema?
- defining slowly changing dimensions
- building the dimensions
- building the fact tables

Working with RDBMs

- optimizing processing with RDBMSs

Optimizing the Physical Environment

- overview
- hardware and memory optimization
- monitoring performance

Classroom Course Materials

Students attend classroom courses in one of our public training centers. You receive a hardcopy of the course notes.