

Optimizing SAS Programs

This Level IV course is designed for SAS programmers who want to maximize the efficiency of their SAS programs. These programmers may write or maintain production programs, and they often work with large data sets. Whatever their job functions, they are resource-constrained and particularly interested in writing efficient SAS code.

Duration: 2.0 days

Course Description [\[Click to register ONLINE \]](#)

This course focuses on the following key areas: establishing and using a benchmark environment, improving the speed of data access, reducing program development time, conserving data storage space, and reducing memory requirements.

Prerequisites

This course is **not** for beginning SAS software users. Before attending this course, you should have completed the [SAS Programming II: Manipulating Data with the DATA Step](#) course. You should have at least nine months of experience writing SAS programs. You should have a thorough understanding of Base SAS software as described in the [SAS Programming II: Manipulating Data with the DATA Step](#) prerequisites. In addition, you should understand

- general programming logic and the compilation and execution process of the DATA step
- how to read, concatenate, merge, match-merge, and interleave SAS data sets by using the SET, MERGE, and BY statements appropriately
- how to create indexes on SAS data files
- how to use array processing and DO loops to process data iteratively
- basic SQL processing.

Course Contents

Measuring Efficiency

- measuring machine resources
- setting up a benchmark environment

Controlling Memory Usage

- controlling page size and the number of buffers
- using the SASFILE statement
- using system options to control memory usage (Self-Study)

Utilizing Best Practices to Improve Efficiency

- executing only necessary statements
- eliminating unnecessary passes through the data
- reading and writing only essential data
- storing data in permanent SAS data sets
- avoiding unnecessary PROC invocation (Self-Study)

Querying Data Efficiently

- indexing
- using the IN operator versus the OR operator
- creating a query with PROC SQL versus PROC PRINT
- summarizing data

Sorting Data Sets

- avoiding unnecessary sorts
- strategies for handling large data sets
- eliminating duplicates
- calculating and allocating sort resources
- choosing the right sort routine (Self-Study)

Combining Data Vertically

- concatenating data
- interleaving data
- comparison of techniques

Combining Data Horizontally

- one-to-one and one-to-many matches
- non-matches
- retrieving data from multiple SAS data sets
- many-to-many matches (Self-Study)
- comparison of techniques

Controlling Data Storage Space

- handling long character data
- reducing the length of numeric variables
- compressing data files
- using SAS data views
- updating in place (Self-Study)
- WORK library (Self-Study)
- accessing DBMS tables or spreadsheets (Self-Study)