Base SAS Programming

1. SAS Programming 1: Essentials

Duration: 2.5 Days

This course is for users who want to learn how to write SAS programs to access, explore, prepare, and analyze data. It is the entry point to learning SAS programming for data science, machine learning, and artificial intelligence. It is a prerequisite to many other SAS courses.

Learn how to

use SAS Studio and SAS Enterprise Guide to write and submit SAS programs
access SAS, Microsoft Excel, and text data
explore and validate data
prepare data by subsetting rows and computing new columns
analyze and report on data
export data and results to Excel, PDF, and other formats
use SQL in SAS to query and join tables.

Who should attend
Anyone starting to write SAS programs

Prerequisites
Before attending this course, you should have experience using computer software. Specifically, you should be able to

- understand file structures and system commands on your operating systems
- access data files on your operating systems.

No prior SAS experience is needed.

Course Contents

Essentials

- the SAS programming process
- using SAS programming tools
- understanding SAS syntax

Accessing Data

- understanding SAS data
- accessing data through libraries

importing data into

SAS Exploring and

Validating Data

- exploring data
- filtering rows
- formatting columns
- sorting data and removing duplicates

Preparing Data

- reading and filtering data
- computing new columns
- conditional processing

Analyzing and Reporting on Data

- enhancing reports with titles, footnotes, and labels
- creating frequency reports
- creating summary statistics reports

Exporting Results

- exporting data
- exporting reports

Using SQL in SAS

- using Structured Query Language in SAS
- joining tables using SQL in SAS

2. SAS Programming 2: Data Manipulation Techniques

Duration: 2.5 Days

This course is for those who need to learn data manipulation techniques using the SAS DATA step and procedures to access, transform, and summarize data. The course builds on the concepts that are presented in the *SAS Programming 1: Essentials* course and is not recommended for beginning SAS software users.

Learn how to

understand and control DATA step processing
create an accumulating column and process data in groups
manipulate data with functions
convert column type
create custom formats
concatenate and merge tables
process repetitive code
restructure tables.

Who should attend
Business analysts and SAS programmers

Prerequisites
Before attending this course, you should be able to do the following:

write DATA step code to subset rows and columns, compute new columns, and process data conditionally
sort tables using the SORT procedure
apply SAS formats

Course Contents

Controlling DATA Step Processing

setting up for this course
understanding DATA step processing
directing DATA step output

Summarizing Data

creating an accumulating column
processing data in groups

Manipulating Data with Functions

understanding SAS functions and CALL routines
using numeric and date functions
using character functions
using special functions to convert column type

Creating Custom Formats

creating and using custom formats
creating custom formats from tables

Combining Tables

concatenating tables
merging tables
identifying matching and nonmatching rows

Processing Repetitive Code

using iterative DO loops
using conditional DO loops

Restructuring Tables

restructuring data with the DATA step
restructuring data with the TRANSPOSE procedure

SAS Macro Language 1: Essentials

Duration: 14.0 hours

This course focuses on the components of the SAS macro facility and how to design, write, and debug macro systems. Emphasis is placed on understanding how programs with macro code are processed.

Learn how to

- perform text substitution in SAS code
- automate and customize the production of SAS code
- conditionally or iteratively construct SAS code
- use macro variables and macro functions.

Who should attend: Experienced SAS programmers who have a sound understanding of DATA step processing and who want to write SAS programs that are reusable and dynamic

Prerequisites

Before attending this course, you should have completed the *SAS Programming 2: Data Manipulation Techniques* course or have equivalent knowledge. Specifically, you should be able to

- use a DATA step to read from or write to a SAS data set or external file
- use DATA step programming statements such as IF-THEN/ELSE, DO WHILE, DO UNTIL, and iterative DO
- use SAS data set options such as DROP=, KEEP=, and OBS=
- use character functions such as SUBSTR, SCAN, INDEX, and UPCASE
- form subsets of data using the WHERE clause
- create and use SAS date values and constants
- use SAS procedures such as SORT, PRINT, CONTENTS, MEANS, FREQ, TABULATE, and CHART.

Software Addressed

This course addresses Base SAS software. This course is appropriate for students who are using SAS 9 software.

Course Contents

Introduction

- overview of SAS Foundation
- course logistics
- course data files

- purpose of the macro facility
- program flow

Macro Variables

- introduction to macro variables
- automatic macro variables
- macro variable references
- user-defined macro variables
- delimiting macro variable references
- macro functions

Macro Definitions

- defining and calling a macro
- macro parameters

DATA Step and SQL Interfaces

- creating macro variables in the DATA step
- indirect references to macro variables
- creating macro variables in SQL

Macro Programs

- conditional processing
- parameter validation
- iterative processing
- global and local symbol tables

Learning More

- SAS resources
- beyond this course

Supplemental Materials

- program flow

SAS SQL 1: Essentials

Duration: 17.5 hours

This outline is provisional and subject to change.

This course teaches you how to process SAS data using Structured Query Language (SQL).

Learn how to

- query and subset data
- summarize and present data
- combine tables, including complex joins and merges
- create and modify table views and indexes
- replace multiple DATA and PROC steps with one SQL query.

Who should attend: SAS programmers and business analysts

Prerequisites

Before attending this class, you should be able to

- submit SAS programs on your operating system
- create and access SAS data sets
- use arithmetic, comparison, and logical operators
- invoke SAS procedures.

You can gain this experience from the *SAS Programming 1: Essentials* course. No knowledge of SQL is necessary.

Software Addressed

This course addresses Base SAS software. This course is appropriate for students who are using SAS 9 software.

Course Contents

Introduction

- overview of SAS Foundation
- course logistics
- course data files
- introducing the Structured Query Language

Basic Queries

overview of the SQL procedure
specifying columns
specifying rows

Displaying Query Results

presenting data
summarizing data

SQL Joins

introduction to SQL joins
inner joins
outer joins
complex SQL joins

Subqueries

noncorrelated subqueries
in-line views

Set Operators

introduction to set operators
the UNION operator
the OUTER UNION operator
the EXCEPT operator
the INTERSECT operator

Creating Tables and Views

creating tables with the SQL procedure
creating views with the SQL procedure

Advanced PROC SQL Features

dictionary tables and views
using SQL procedure options
interfacing PROC SQL with the macro language

Learning More

SAS resources
beyond this course

SAS Programming 3: Advanced Techniques and Efficiencies

Duration: 21.0 hours

This course is for SAS programmers who prepare data for analysis. The comparisons of manipulation techniques and resource cost benefits are designed to help programmers choose the most appropriate technique for their data situation.

Learn how to compare various SAS programming techniques that enable you to

- benchmark computer resource usage
- control memory, I/O, and CPU resources
- create and use indexes
- combine data horizontally
- use hash and hiter DATA step component objects and arrays as lookup tables
- compress SAS data sets
- sample your SAS data sets
- create and use SAS data views
- safely reduce the length of numeric variables
- create user-defined functions and informats.

Who should attend: Experienced SAS programmers

Prerequisites

This course is **not** appropriate for beginning SAS software users. Before attending this course, you should have at least nine months of SAS programming experience and should have completed the *SAS Programming 2: Data Manipulation Techniques* course. Specifically, you should be able to do the following:

- understand your operating system file structures and perform basic operating system tasks
- understand programming logic concepts
- understand the compilation and execution processes of the DATA step
- use different varieties of input to create SAS data sets from external files
- use SAS software to access SAS libraries
- create and use SAS date values
- read, concatenate, merge, match-merge, and interleave SAS data sets
- use the DROP=, KEEP=, and RENAME= data set options
- create multiple output data sets
- use one-dimensional array processing and DO loops to process data iteratively
- use SAS functions to perform data manipulation and transformations
- use the FORMAT procedure to create user-defined formats.

Software Addressed

This course addresses Base SAS software. This course is appropriate for students who are using SAS 9 software.

Course Contents

Introduction

- overview of SAS Foundation
- course logistics
- creating the course data

Efficient SAS Programming

- identifying computer resources related to efficiency

Controlling I/O Processing and Memory

- SAS DATA step processing
- controlling I/O
- reducing the length of numeric variables
- compressing SAS data sets
- using SAS views

Accessing Observations

- access methods
- accessing observations by number
- creating an index
- using an index

DATA Step Arrays

- introduction to lookup techniques
- one-dimensional arrays
- multidimensional arrays
- loading a multidimensional array from a SAS data set

DATA Step Hash and Hiter Objects

- introduction
- hash object methods
- loading a hash object from a SAS data set
- DATA step hiter object

Combining Data Horizontally

- DATA step merges and SQL procedure joins
- using an index to combine data
- combining summary and detail data
- combining data conditionally

User-Defined Functions and Formats

- user-defined functions
- user-defined formats

Learning More

- areas of support from SAS
- other courses to consider

Combining Raw Data Files Vertically

- combining raw data files vertically

Statistics 1: Introduction to ANOVA, Regression, and Logistic Regression

Duration: 21.0 hours

This course is for SAS software users who perform statistical analyses using SAS/STAT software. The focus is on *t* tests, ANOVA, linear regression, and logistic regression. This course (or equivalent knowledge) is a prerequisite to many of the courses in the statistical analysis curriculum.

Learn how to

- generate descriptive statistics and explore data with graphs
- perform analysis of variance and apply multiple comparison techniques
- perform linear regression and assess the assumptions
- use regression model selection techniques to aid in the choice of predictor variables in multiple regression
- use diagnostic statistics to assess statistical assumptions and identify potential outliers in multiple regression
- use chi-square statistics to detect associations among categorical variables
- fit a multiple logistic regression model.

Who should attend: Statisticians, researchers, and business analysts who use SAS programming to generate analyses using either continuous or categorical response (dependent) variables

Prerequisites

Before attending this course, you should

have completed the equivalent of an undergraduate course in statistics covering *p*-values, hypothesis testing, analysis of variance, and regression
be able to execute SAS programs and create SAS data sets. You can gain this experience by completing the *SAS Programming 1: Essentials* course.

Course Contents

Prerequisite Basic Concepts

descriptive statistics
inferential statistics
steps for conducting a hypothesis test
basics of using your SAS software

Introduction to Statistics

examining data distributions
obtaining and interpreting sample statistics using the UNIVARIATE and MEANS procedures
examining data distributions graphically in the UNIVARIATE and SGPLOT procedures
constructing confidence intervals
performing simple tests of hypothesis

t Tests and Analysis of Variance

performing tests of differences between two group means using PROC TTEST
performing one-way ANOVA with the GLM procedure
performing post-hoc multiple comparisons tests in PROC GLM
performing two-way ANOVA with and without interactions

Linear Regression

producing correlations with the CORR procedure
fitting a simple linear regression model with the REG procedure
understanding the concepts of multiple regression
using automated model selection techniques in PROC REG to choose from among several candidate models
interpreting models

Linear Regression Diagnostics

examining residuals
investigating influential observations
assessing collinearity

Categorical Data Analysis

producing frequency tables with the FREQ procedure
examining tests for general and linear association using the FREQ procedure
understanding exact tests
understanding the concepts of logistic regression
fitting univariate and multivariate logistic regression models using the LOGISTIC procedure

Predictive Modeling Using Logistic Regression

Duration: 14.0 hours

This course covers predictive modeling using SAS/STAT software with emphasis on the LOGISTIC procedure. This course also discusses selecting variables, assessing models, treating missing values and using efficiency techniques for massive data sets.

Learn how to

- use logistic regression to model an individual's behavior as a function of known inputs
- create effect plots and odds ratio plots using ODS Statistical Graphics
- handle missing data values
- tackle multicollinearity in your predictors
- assess model performance and compare models.

Who should attend: Modelers, analysts and statisticians who need to build predictive models, particularly models from the banking, financial services, direct marketing, insurance and telecommunications industries

Prerequisites

Before attending this course, you should

have experience executing SAS programs and creating SAS data sets, which you can gain from the *SAS Programming 1: Essentials* course
have experience building statistical models using SAS software
have completed a statistics course that covers linear regression and logistic regression, such as the *Statistics 1: Introduction to ANOVA, Regression, and Logistic Regression* course.

Course Contents

Predictive Modeling

- business applications
- analytical challenges

Fitting the Model

- parameter estimation
- adjustments for oversampling

Preparing the Input Variables

- missing values
- categorical inputs
- variable clustering
- variable screening
- subset selection

Classifier Performance

- ROC curves and Lift charts
- optimal cutoffs
- K-S statistic
- *c* statistic
- profit
- evaluating a series of models