SAS Certification Prep Guide: Base Programming for SAS 9, Third Edition

by SAS Institute

Glossary

A-C

aggregate storage location

a location in an operating system that can contain a group of distinct files. The exact name for this location varies by operating system; for example, directory, folder, or partitioned data set.

analysis variable

a numeric variable that is used to calculate statistics or to display values. Usually an analysis variable contains quantitative or continuous values, but this is not required.

arithmetic expression

See SAS expression

arithmetic operator

in SAS, any of the symbols (+, -, /, *, and **) that are used to perform addition, subtraction, division, multiplication, or exponentiation in arithmetic expressions. In SYSTEM 2000 software only, ** is not supported.

array

in the SAS programming language, a temporary grouping of SAS variables that have the same data type and that are arranged in a particular order and identified by an array name. The array exists only for the duration of the current DATA step.

array name

a name that is selected to identify a group of variables or temporary data elements. It must be a valid SAS name that is not the name of a variable in the same DATA step or SCL (SAS Component Language) program.

array reference

a reference to an element to be processed in an array.

attribute

See variable attribute

Base SAS

the core product that is part of SAS Foundation and is installed with every deployment of SAS software. Base SAS provides an information delivery system for accessing, managing, analyzing, and presenting data.

Base SAS software

See Base SAS

BY group

a group of observations or rows that have the same value for a variable that is specified in a BY statement. If more than one variable is specified in a BY statement, then the BY group is a group of observations that have a unique combination of values for those variables.

BY group variable

See BY variable

BY value

the value of a BY variable.

BY variable

a variable that is named in a BY statement and whose values define groups of observations to process.

BY-group processing

the process of using the BY statement to process observations that are ordered, grouped, or indexed according to the values of one or more variables. Many SAS procedures and the DATA step support BY-group processing. For example, you can use BY-group processing with the PRINT procedure to print separate reports for different groups of observations in a single SAS data set.

catalog

See SAS catalog

character format

a set of instructions that tell SAS to use a specific pattern for writing character data values.

character function

a type of function that enables you to manipulate, compare, evaluate, or analyze character strings.

character informat

a set of instructions that tell SAS to use a specific pattern for reading character data values into character variables.

character value

a value that can contain alphabetic characters, the numeric characters 0 through 9, and other special characters.

character variable

a variable whose values can consist of alphabetic characters and special characters as well as numeric characters.

class variable

See classification variable

classification variable

a variable whose values are used to group (or classify) the observations in a data set into different groups that are meaningful for analysis. A classification variable can have either character or numeric values. Classification variables include group, subgroup, category, and BY variables.

code editing window

a generic term for any window into which users can type or paste program code, or in which they can make changes to existing programs.

column input

in the DATA step, a style of input in which column numbers are included in the INPUT statement to tell SAS which columns contain the values for each variable. This style of input is useful when the values for each variable are in the same location in all records.

comment

See comment statement

comment statement

information that is embedded in a SAS program and that serves as explanatory text. SAS ignores comments during processing but writes them to the SAS log. Comment syntax has several forms. For example, a comment can appear as a statement that begins with an asterisk and ends with a semicolon, as in * message ;.

comparison operator

in programming languages, a symbol or mnemonic code that is used in expressions to test for a particular relationship between two values or text strings. For example, the symbol < and its corresponding mnemonic, LT, are used to determine whether one value is less than another.

compilation

See program compilation

compound expression

an expression that contains more than one operator.

concatenate

to join the contents of two or more elements, end to end, forming a separate element. Examples of elements are character values, tables, external files, SAS data sets, and SAS libraries.

condition

in a SAS program, one or more numeric or character expressions that result in a value on which some decision depends.

constant

in SAS software, a number or a character string that indicates a fixed value.

D-E

data error

a type of execution error that occurs when a SAS program analyzes data that contains invalid values. For example, a data error occurs if you specify numeric variables in the INPUT statement for character data. SAS reports these errors in the SAS log but continues to execute the program.

data set

See SAS data set

data set option

a SAS language element that specifies actions that apply to only one particular SAS data set. For example, data set options enable you to rename variables, to select only the first or last n observations for processing, to drop variables from processing or from the output data set, and to specify a password for a SAS data set.

DATA step

in a SAS program, a group of statements that begins with a DATA statement and that ends with either a RUN statement, another DATA statement, a PROC statement, or the end of the job. The DATA step enables you to read raw data or other SAS data sets and to create SAS data sets.

data value

a unit of character, numeric, or alphanumeric information. This unit is stored as one item in a data record, such as a person's height being stored as one variable (namely, a column or vertical component) in an observation (row).

data view

See SAS data view

date and time format

instructions that tell SAS how to write numeric values as dates, times, and datetimes.

date and time informat

the instructions that tell SAS how to read numeric values that are represented as dates, times, and datetimes.

date constant

See SAS date constant

date value

See SAS date value

datetime constant

See SAS datetime constant

datetime value

See SAS datetime value

delimiter

a character that serves as a boundary that separates the elements of a text string.

descriptor information

information about the contents and attributes of a SAS data set. For example, the descriptor information includes the data types and lengths of the variables, as well as which engine was used to create the data. SAS creates and maintains descriptor information within every SAS data set.

directory

a named subdivision on a computer disk, used in organizing files and metadata about files.

DO group

a sequence of statements that starts with a simple DO statement and that ends with a corresponding END statement.

DO loop

a sequence of statements that starts with an iterative DO, DO WHILE, or DO UNTIL statement and that ends with a corresponding END statement. The statements are executed (usually repeatedly) according to directions that are specified in the DO statement.

double trailing at sign

a special symbol @@ that is used to hold a line of data in the input buffer during multiple iterations of a DATA step.

engine

a component of SAS software that reads from or writes to a file. Various engines enable SAS to access different types of file formats.

external file

a file that is created and maintained by a host operating system or by another vendor's software application. An external file can read both data and stored SAS statements.

F-L

field

the smallest logical unit of data in a file.

file reference

See fileref

fileref

a name that is temporarily assigned to an external file or to an aggregate storage location such as a directory or a folder. The fileref identifies the file or the storage location to SAS.

FIRST. variable

a temporary variable that SAS creates to identify the first observation of each BY group. The variable is not added to the SAS data set.

floating-point representation

a compact form of storing real numbers on a computer, similar to scientific notation. Different operating systems use different techniques for floating-point representation.

format modifier

a special symbol that is used in the INPUT and PUT statements and which enables you to control how SAS reads input data and writes output data.

formatted input

a style of input that uses special instructions called informats in the INPUT statement to determine how values that are entered in data fields should be interpreted.

function

See SAS function

informat

See SAS informat

input buffer

a temporary area of memory into which each record of data is read when the INPUT statement executes.

integer

a number that does not contain a decimal point.

interleaving

a process in which SAS combines two or more sorted SAS data sets into one sorted SAS data set based on the values of the BY variables.

keyword

See SAS keyword

LAST. variable

a temporary variable that SAS creates to identify the last observation of each BY group. This variable is not added to the SAS data set.

length variable

a numeric variable that is used with the $VARYING. informat or format to specify the actual length of a character variable whose values do not all have the same length.

library member

any of several types of SAS file in a SAS library. A library member can be a data set, a view, a catalog, a stored program, or an access descriptor.

library reference

See libref

libref

a SAS name that is associated with the location of a SAS library. For example, in the name MYLIB.MYFILE, MYLIB is the libref, and MYFILE is a file in the SAS library.

line-hold specifier

a special symbol used in INPUT and PUT statements that enables you to hold a record in the input or output buffer for further processing. Line-hold specifiers include the trailing at sign (@) and the double trailing at sign (@@).

list input

a style of input in which names of variables, not column locations, are specified in the INPUT statement. List input scans input records for data values that are separated by at least one blank or by some other delimiter.

LISTING output

SAS procedure output that is in a monospace font. All text in listing output has the same font size, and no special font styles are applied to it.

log

See SAS log

logical data model

a framework into which engines fit information for processing by SAS. This data model is a logical representation of data or files, not a physical structure.

logical operator

an operator that is used in expressions to link sequences of comparisons. The logical operators are AND, OR, and NOT.

M-P

match-merging

a process in which SAS joins observations from two or more SAS data sets according to the values of the BY variables.

member name

a name that is assigned to a SAS file in a SAS library.

merging

the process of combining observations from two or more SAS data sets into a single observation in a new SAS data set.

missing value

a type of value for a variable that contains no data for a particular row or column. By default, SAS writes a missing numeric value as a single period and a missing character value as a blank space.

modified list input

a style of input that uses special instructions called informats and format modifiers in the INPUT statement. Modified list input scans input records for data values that are separated by at least one blank (or by some other delimiter), or in some cases, by multiple blanks.

numeric format

a set of instructions that tell SAS to use a specific pattern for writing the values of numeric variables.

numeric informat

a set of instructions that tell SAS to use a specific pattern for reading numeric data values.

numeric value

a value that usually contains only numbers, which can include numbers in E-notation and hexadecimal notation. A numeric value can sometimes contain a decimal point, a plus sign, or a minus sign. Numeric values are stored in numeric variables.

observation

a row in a SAS data set. All of the data values in an observation are associated with a single entity such as a customer or a state. Each observation contains either one data value or a missing-value indicator for each variable.

one-level name

a one-part name for a SAS file in which the libref (library reference) is omitted and only the filename is specified. When you specify a one-level name, the default temporary library, Work, is assumed.

one-to-one matching

the process of combining observations from two or more data sets into one observation, using two or more SET statements to read observations independently from each data set.

one-to-one merging

the process of using the MERGE statement (without a BY statement) to combine observations from two or more data sets based on the observations' positions in the data sets.

operand

any of the variables and constants in an expression that contain operators, variables, and constants.

operator

See SAS operator

page size

the number of bytes of data that SAS moves between external storage and memory in one input/output operation. Page size is analogous to buffer size for SAS data sets.

PDV

See program data vector

permanent SAS data library

a SAS library that is not deleted when a SAS session ends, and which is therefore available to subsequent SAS sessions.

permanent SAS file

a file in a SAS library that is not deleted when the SAS session or job terminates.

physical filename

the name that an operating system uses to identify a file.

pointer

in the DATA step, a programming tool that SAS uses to keep track of its position in the input or output buffer.

pointer control

the process of instructing SAS to move the pointer before reading or writing data.

PROC

See SAS procedure

PROC step

a group of SAS statements that call and execute a SAS procedure. A PROC step usually takes a SAS data set as input.

procedure

See SAS procedure

procedure output file

an external file that contains the result of the analysis that a SAS procedure performs or the report that the procedure produces. Most SAS procedures write output to the procedure output file by default. Reports that are produced by SAS DATA steps, using PUT statements and a FILE statement along with a PRINT destination, also go to this file.

program compilation

the process of checking syntax and translating a portion of a program into a form that the computer can execute.

program data vector

the temporary area of computer memory in which SAS builds a SAS data set, one observation at a time. The program data vector is a logical concept and does not necessarily correspond to a single contiguous area of memory. Short form: PDV.

programming error

a flaw in the logic of a SAS program that can cause the program to fail or to perform differently than the programmer intended.

R-S

raw data

in statistical analysis, data (including data in SAS data sets) that has not had a particular operation, such as standardization, performed on it.

SAS catalog

a SAS file that stores many different kinds of information in smaller units called catalog entries. A single SAS catalog can contain different types of catalog entries.

SAS data file

a type of SAS data set that contains data values as well as descriptor information that is associated with the data. The descriptor information includes information such as the data types and lengths of the variables, as well as the name of the engine that was used to create the data.

SAS data set

a file whose contents are in one of the native SAS file formats. There are two types of SAS data sets: SAS data files and SAS data views. SAS data files contain data values in addition to descriptor information that is associated with the data. SAS data views contain only the descriptor information plus other information that is required for retrieving data values from other SAS data sets or from files whose contents are in other software vendors' file formats.

SAS data set option

an option that appears in parentheses after a SAS data set name. Data set options specify actions that apply only to the processing of that SAS data set.

SAS data view

a type of SAS data set that retrieves data values from other files. A SAS data view contains only descriptor information such as the data types and lengths of the variables (columns) plus other information that is required for retrieving data values from other SAS data sets or from files that are stored in other software vendors' file formats. Short form: data view.

SAS date constant

a string in the form 'ddMMMyy'd or 'ddMMMyyyy'd that represents a date in a SAS statement. The string is enclosed in quotation marks and is followed by the character d (for example, '6JUL01'd, '06JUL01'd, '6 JUL2001'd, or '06JUL2001'd).

SAS date value

an integer that represents a date in SAS software. The integer represents the number of days between January 1, 1960, and another specified date. For example, the SAS date value 366 represents the calendar date January 1, 1961.

SAS datetime constant

a string in the form 'ddMMMyy:hh:mm:ss'dt or 'ddMMMyyyy:hh:mm:ss'dt that represents a date and time in SAS. The string is enclosed in quotation marks and is followed by the characters dt (for example, '06JUL2001:09:53:22'dt).

SAS datetime value

an integer that represents a date and a time in SAS software. The integer represents the number of seconds between midnight, January 1, 1960, and another specified date and time. For example, the SAS datetime value for 9:30 a.m., June 5, 2000, is 1275816600.

SAS Enterprise Guide

a software application with a point-and-click interface that gives users access to the functionality of many components of SAS software. Interactive dialog boxes guide users through data analysis tasks and reporting tasks, and users can easily export the results of those tasks to other Windows applications or servers. SAS Enterprise Guide provides access not only to SAS data files, but also to data that is in a wide variety of other software vendors' formats and in other operating system formats.

SAS expression

a type of macro expression consisting of a sequence of operands and arithmetic operators that form a set of instructions that are evaluated to produce a numeric value, a character value, or a Boolean value. Examples of operands are constants and system functions. SAS uses arithmetic expressions in program statements to create variables, to assign values, to calculate new values, to transform variables, and to perform conditional processing.

SAS file

a specially structured file that is created, organized, and maintained by SAS. A SAS file can be a SAS data set, a catalog, a stored program, an access descriptor, a utility file, a multidimensional database file, a financial database file, a data mining database file, or an item store file.

SAS function

a type of SAS language element that can be used in an expression or assignment statement to process zero or more arguments and to return a value. Examples of SAS functions are MEAN and SUM. Short form: function.

SAS informat

a type of SAS language element that applies a pattern to or executes instructions for a data value to be read as input. Types of informats correspond to the data's type: numeric, character, date, time, or timestamp. The ability to create user-defined informats is also supported. Examples of SAS informats are BINARY and DATE. Short form: informat.

SAS keyword

a literal that is a primary part of the SAS language. For example, SAS keywords include DATA, PROC, RUN, names of SAS language elements, names of SAS statement options, and system variables.

SAS library

one or more files that are defined, recognized, and accessible by SAS and that are referenced and stored as a unit. Each file is a member of the library.

SAS log

a file that contains a record of the SAS statements that you enter, as well as messages about the execution of your program.

SAS name

a name that is assigned to items such as SAS variables and SAS data sets. For most SAS names, the first character must be a letter or an underscore. Subsequent characters can be letters, numbers, or underscores. Blanks and special characters (except the underscore) are not allowed. However, the VALIDVARNAME= system option determines what rules apply to SAS variable names. The maximum length of a SAS name depends on the language element that it is assigned to.

SAS operator

in a SAS expression, any of several symbols that request a comparison, a logical operation, or an arithmetic calculation.

SAS procedure

a program that provides specific functionality and that is accessed with a PROC statement. For example, SAS procedures can be used to produce reports, to manage files, or to analyze data. Many procedures are included in SAS software.

SAS program

a group of SAS statements that guide SAS through a process or series of processes in order to read and transform input data and to generate output. The DATA step and the procedure step, used alone or in combination, form the basis of SAS programs.

SAS session

the activity between invoking and exiting a specific SAS software product.

SAS statement

a string of SAS keywords, SAS names, and special characters and operators that instructs SAS to perform an operation or that gives information to SAS. Each SAS statement ends with a semicolon.

SAS system option

an option that affects the processing of an entire SAS program or interactive SAS session from the time the option is specified until it is changed. Examples of items that are controlled by SAS system options include the appearance of SAS output, the handling of some files that are used by SAS, the use of system variables, the processing of observations in SAS data sets, features of SAS initialization, and the way SAS interacts with your host operating environment.

SAS time constant

a string in the form 'hh:mm:ss't that represents a time in a SAS statement. The string is enclosed in quotation marks and is followed by the character t (for example, '09:53:22't).

SAS time value

an integer that represents a time in SAS software. The integer represents the number of seconds between midnight of the current day and another specified time value. For example, the SAS time value for 9:30 a.m. is 34200.

SAS variable

a column in a SAS data set or in a SAS data view. The data values for each variable describe a single characteristic for all observations (rows).

Sashelp library

a SAS library supplied by SAS software that stores the text for Help windows, default function-key definitions and window definitions, and menus.

Sasuser library

a default, permanent SAS library that is created at the beginning of your first SAS session. The Sasuser library contains a PROFILE catalog that stores the customized features or settings that you specify for SAS.

statement

See SAS statement

statement option

a word that you specify in a particular SAS statement and which affects only the processing that that statement performs.

syntax checking

the process by which SAS checks each SAS statement for proper usage, correct spelling, proper SAS naming conventions, and so on.

syntax error

an error in the spelling or grammar of a SAS statement. SAS finds syntax errors as it compiles each SAS step before execution.

system option

See SAS system option

T-W

target variable

the variable to which the result of a function or expression is assigned.

taskbar

the bar at the bottom of the Windows desktop that displays active applications. The taskbar enables you to easily switch between applications and to restore, move, size, minimize, maximize, and close applications.

temporary SAS file

a SAS file in a SAS library (usually the Work library) that is deleted at the end of the SAS session or job.

temporary SAS library

a library that exists only for the current SAS session or job. The most common temporary library is the Work library.

text output

another term for listing output.

time constant

See SAS time constant

time value

See SAS time value

title

a heading that is printed at the top of each page of SAS output or of the SAS log.

toggle

an option, parameter, or other mechanism that enables you to turn on or turn off a processing feature.

trailing at sign

a special symbol @ that is used to hold a line of input or output so that SAS can read from it or write to it in a subsequent INPUT or PUT statement.

variable

See SAS variable

variable attribute

any of the following characteristics that are associated with a particular variable: name, label, format, informat, data type, and length.

variable label

up to 256 characters of descriptive text that can be printed in the output by certain procedures instead of, or in addition to, the variable name.

variable length

the number of bytes used to store each of a variable's values in a SAS data set.

variable type

the classification of a variable as either numeric or character. Type is an attribute of SAS variables.

Work library

a temporary SAS library that is automatically defined by SAS at the beginning of each SAS session or SAS job. Unless you have specified a User library, any newly created SAS file that has a one- level name will be placed in the Work library by default and will be deleted at the end of the current SAS session or job.