An Introduction to SAS® Programming

SAS programs consists of two major type of steps

> The DATA step – used to create or modify a SAS dataset – [Contents > SAS Products > Base SAS > SAS Language concepts > Data Step Concepts]
>
>> SAS dataset – a file containing a collection of similar information
>>
>> Can be visualized as a two-dimensional array (table) that looks like spreadsheet (e.g. like EXCEL)
>>
>> Observation – each row represents the information for a single item
>>
>> Variable – each column contains one type of item
>>
>> In addition to data values, variable names and types, lengths, labels and formats are stored in this file
>>
>> Source of data for the data step
>>
>>> Convert a raw data file to a SAS data set. This can be either stored in a separate file or included in the SAS program itself
>>>
>>> The SAS datasets are initially created from some source of data. The SAS dataset, once created, can be stored as a "permanent" file or recreated each time SAS is run on that dataset.
>>>
>>> When the SAS dataset is created we can assign formats and labels
>>>
>>> When the SAS dataset is created we can perform modifications to the data, transformations and calculations

> The PROC step – (PROC is from PROCedure) used to process SAS data sets
>
>> Allows File manipulation - SORTS
>>
>> Report preparation – PRINT and tabulation procedures
>>
>> Analysis – MEANS, FREQ, various statistical analyses
>>
>> Graphics – CHART, PLOT, TIMEPLOT
>>
>> Utilities – CONTENTS, DATASETS, FORMAT, file import and compare utilities

SAS Display Manager (in SAS help see "Using SAS Software in Your Operating Environment > Using SAS in Windows > Running programs in the SAS Windowing environment"). Here there are descriptions of the SAS interface, including graphics of SAS windows and menu options.

> Program editor window (= Editor) - type, edit, save and submit SAS programs
>
> Log window (= Log) – Displays the SAS log (notes and messages produced when programs run)
>
> Output window (= Output) – Displays output from SAS program runs

General rules about SAS programs

SAS statements can begin on any line.

SAS statements can be continued on another line as long as no word is split.

More than one statement can be written on a single line.

At least one blank must separate each word or item in a SAS statement, except for mathematical operators (e.g. +, -, *, /, =).

Each SAS statement must end with a semicolon, ";".

In PC SAS it is a good practice to end each DATA step or PROC step with a "RUN;" statement.

Statements surrounded by a pair of "/*" to start and "*/" to end can be included anywhere in a SAS statement that a blank would appear. The enclosed section is a comment. This can also be used to turn any segment of a SAS program into an inactive comment section.

The order of many statements is not important. This is true in both DATA steps and PROC steps. There are some logical exceptions. For example, you cannot process data until after an "INPUT" statement.

Introduction to the SAS DATA step – creating a SAS data set from raw data.

The DATA step starts with the word DATA followed by a data-set-name.

There are some limits on what can be used as the data-set-name. The SAS help (9.1.3) states that "A SAS name can be up to eight characters long. The first character must be a letter (A,B,C,...,Z) or underscore (_). Subsequent characters can be letters, numbers (0 to 9), or underscores. Note that no blanks are allowed. Two names (_N_ and _ERROR_) are reserved by the SAS System."

Names longer than eight characters are acceptable in SAS 9.1.3, up to 32 characters.

Lengths of variables and names in SAS

Maximum Length of SAS Names	
SAS Application	**Max Length**
Arrays	32
CALL routines	16
Catalog entries	32
DATA step statement labels	32
DATA step variable labels	256
DATA step variables	32
DATA step windows	32
Engines	8
Filerefs	8
Formats, character	31
Formats, numeric	32
Functions	16
Generation data sets	28
Informats, character	30
Informats, numeric	31
Librefs	8
Macro variables	32
Macro windows	32
Macros	32
Members of SAS data libraries (SAS data sets, views, catalogs, indexes) except for generation data sets	32
Passwords	8
Procedure names (first 8 characters must be unique, and may not begin with "SAS")	16
SCL variables	32