Long Variable Names

In some instances, data files with variable names longer than eight bytes require special consideration:

- If you save a data file in portable format (see EXPORT on p. 601) , variable names that exceed eight bytes are converted to unique eight-character names. For example, *mylongrootname1*, *mylongrootname2*, and *mylongrootname3* would be converted to *mylongro*, *mylong_2*, and *mylong_3*, respectively.
- When using data files with variable names longer than eight bytes in SPSS 10.x or 11.x, unique, eight-byte versions of variable names are used; however, the original variable names are preserved for use in release 12.0 or later. In releases prior to SPSS 10.0, the original long variable names are lost if you save the data file.
- Matrix data files (commonly created with the `MATRIX OUT` subcommand, available in some procedures) in which the *VARNAME_* variable is longer than an eight-byte string cannot be read by releases of SPSS prior to release 12.0.

Variables

The columns in an SPSS data file are **variables**. Variables are similar to fields in a database table.

- Variable names can be defined with numerous commands, including DATA LIST, GET DATA, NUMERIC, STRING, VECTOR, COMPUTE, and RECODE. They can be changed with the RENAME VARIABLES command.
- Optional variable attributes can include descriptive variable labels (VARIABLE LABELS command), value labels (VALUE LABELS command), and missing value definitions (MISSING VALUES command).

The following sections provide information on variable naming rules, syntax for referring to inclusive lists of variables (keywords `ALL` and `TO`), scratch (temporary) variables, and system variables.

Variable Names

Variable names are stored in the dictionary of an SPSS-format data file or active dataset. Observe the following rules when establishing variable names or referring to variables by their names on commands:

- Each variable name must be unique; duplication is not allowed.
- Variable names can be up to 64 bytes long, and the first character must be a letter or one of the characters @, #, or $. Subsequent characters can be any combination of letters, numbers, a period (.), and nonpunctuation characters. Sixty-four bytes typically means 64 characters in single-byte languages (e.g., English, French, German, Spanish, Italian, Hebrew, Russian, Greek, Arabic, and Thai) and 32 characters in double-byte languages (e.g., Japanese, Chinese, and Korean).

  (*Note*: Letters include any nonpunctuation characters used in writing ordinary words in the languages supported in the character set of the platform on which SPSS is running.)

- Variable names cannot contain spaces.
- A # character in the first position of a variable name defines a scratch variable. You can only create scratch variables with command syntax. You cannot specify a # as the first character of a variable in dialog boxes that create new variables. For more information, see Scratch Variables on p. 34.
- A $ sign in the first position indicates that the variable is a system variable. For more information, see System Variables on p. 34. The $ sign is not allowed as the initial character of a user-defined variable.
- The period, underscore, and the characters $, #, and @ can be used within variable names. For example, *A._$@#1* is a valid variable name.
- Variable names ending with a period should be avoided, since the period may be interpreted as a command terminator. You can only create variables that end with a period in command syntax. You cannot create variables that end with a period in dialog boxes that create new variables.
- Variable names ending in underscores should be avoided, since such names may conflict with names of variables automatically created by commands and procedures.
- Reserved keywords cannot be used as variable names. Reserved keywords are: ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO, and WITH.
- Variable names can be defined with any mixture of uppercase and lowercase characters, and case is preserved for display purposes.
- When long variable names need to wrap onto multiple lines in output, SPSS attempts to break the lines at underscores, periods, and where content changes from lower case to upper case.

Mixed Case Variable Names

Variable names can be defined with any mixture of upper- and lowercase characters, and case is preserved for display purposes.

- Variable names are stored and displayed exactly as specified on commands that read data or create new variables. For example, `compute NewVar = 1` creates a new variable that will be displayed as *NewVar* in the Data Editor and in output from any procedures that display variable names.
- Commands that refer to existing variable names are not case sensitive. For example, `FREQUENCIES VARIABLES = newvar`, `FREQUENCIES VARIABLES = NEWVAR`, and `FREQUENCIES VARIABLES = NewVar` are all functionally equivalent.
- In languages such as Japanese, where some characters exist in both narrow and wide forms, these characters are considered different and are displayed using the form in which they were entered.
- When long variable names need to wrap onto multiple lines in output, SPSS attempts to break lines at underscores, periods, and changes from lower to upper case.

You can use the `RENAME VARIABLES` command to change the case of any characters in a variable name.

Example

```
RENAME VARIABLES (newvariable = NewVariable).
```

- For the existing variable name specification, case is ignored. Any combination of upper and lower case will work.
- For the new variable name, case will be preserved as entered for display purposes.

For more information, see the RENAME VARIABLES command.

Long Variable Names

In some instances, data files with variable names longer than eight bytes require special consideration:

- If you save a data file in portable format (see EXPORT on p. 601) , variable names that exceed eight bytes are converted to unique eight-character names. For example, *mylongrootname1*, *mylongrootname2*, and *mylongrootname3* would be converted to *mylongro*, *mylong_2*, and *mylong_3*, respectively.
- When using data files with variable names longer than eight bytes in SPSS 10.x or 11.x, unique, eight-byte versions of variable names are used; however, the original variable names are preserved for use in release 12.0 or later. In releases prior to SPSS 10.0, the original long variable names are lost if you save the data file.
- Matrix data files (commonly created with the MATRIX OUT subcommand, available in some procedures) in which the *VARNAME_* variable is longer than an eight-byte string cannot be read by releases of SPSS prior to release 12.0.

Keyword TO

You can establish names for a set of variables or refer to any number of consecutive variables by specifying the beginning and the ending variables joined by the keyword TO.

To establish names for a set of variables with the keyword TO, use a character prefix with a numeric suffix.

- The prefix can be any valid name. Both the beginning and ending variables must use the same prefix.
- The numeric suffix can be any integer, but the first number must be smaller than the second. For example, ITEM1 TO ITEM5 establishes five variables named *ITEM1*, *ITEM2*, *ITEM3*, *ITEM4*, and *ITEM5*.
- Leading zeros used in numeric suffixes are included in the variable name. For example, V001 TO V100 establishes 100 variables—*V001*, *V002*, *V003*, ..., *V100*. V1 TO V100 establishes 100 variables—*V1*, *V2*, *V3*, ..., *V100*.

The keyword TO can also be used on procedures and other commands to refer to consecutive variables on the active dataset. For example, AVAR TO VARB refers to the variables *AVAR* and all subsequent variables up to and including *VARB*.