

Modifying Character Values with Functions

Starting from this video, I will introduce you how to use SAS functions to manipulate the values of character variables.



Selected Character Functions

Function	Purpose
SCAN	returns a specified word from a character value.
SUBSTR	extracts a substring or replaces character values.
TRIM	trims trailing blanks from character values.
CATX	concatenates character strings, removes leading and trailing blanks, and inserts separators.
INDEX	searches a character expression for a string of characters, and returns the position of the string's first character for the first occurrence of the string.
FIND	searches for a specific substring of characters within a character string.
UPCASE	converts all letters in a value to uppercase.
LOWCASE	converts all letters in a value to lowercase.
PROPCASE	converts all letters in a value to proper case.
TRANWRD	replaces or removes all occurrences of a pattern of characters within a character string.



SCAN Function

The SCAN function enables you to separate a character value into words and to return a specified word. Let's look at the following example to see how the SCAN function works.

The SCAN function uses delimiters, which are characters that are specified as word separators, to separate a character string into words. For example, if you are working with the character string below and you specify the comma as a delimiter, the SCAN function separates the string into three words.

LOW,MODERATE,HIGH
↑ ↑ ↑
1 2 3

Then the function returns whichever word you specify. In this example, if you specify the third word, the SCAN function returns the word HIGH.



Specifying Multiple Delimiters

When using the SCAN function, you can specify as many delimiters as needed to correctly separate the character expression. When you specify multiple delimiters, SAS uses any of the delimiters, singly or in any combination, as word separators. For example, if you specify both the slash and the hyphen as delimiters, the SCAN function separates the following text string into three words:

607/555-1273
↑ ↑ ↑
1 2 3

The SCAN function treats two or more continuous delimiters as one delimiter. Also, leading delimiters have no effect.



Default Delimiters

If you do not specify delimiters when using the SCAN function, default delimiters are used. The default delimiters are shown as following

blank . < (+ | & ! \$ *) ; ^ - / , %



SCAN Function Syntax

Now that you are familiar with how the SCAN function works, let's examine the syntax of the function.

General form, SCAN function:

`SCAN(argument, n, 'delimiters')`

Example: `LastName=scan(name,1,',');`

where

- argument specifies the character variable or expression to scan.
- n specifies which word to return.
- delimiters are special characters that must be enclosed in single quotation marks (' '). If you do not specify delimiters, default delimiters are used.

