

How to Easily Remove Special Characters from Strings in SAS

Authored by
stats writer

November 24, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Easily Remove Special Characters from Strings in SAS*.
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=100230>

In SAS, the COMPRESS function can be used to remove special characters from a string. This function takes two arguments, a source string and a list of characters to remove. The list of characters can be specified as a string of characters or as a character set. Once the function is applied, it will return a new string with all of the specified characters removed.

The easiest way to remove special characters from a string in SAS is to use the function with the 'kas' modifier.

This function uses the following basic syntax:

```
data new_data;  
set original_data;  
remove_specials = compress(some_string, 'kas');  
run;
```

The following example shows how to use this syntax in practice.

Example: Remove Special Characters from String in SAS

Suppose we have the following dataset in SAS that contains the names of various employees and their total sales:

```
/*create dataset*/  
data data1;  
input name $ sales;  
datalines;  
Bob&%^ 45  
M&$#@ike 50  
Randy)) 39  
Chad!? 14  
Dan** 29  
R 44  
;  
run;
```

```
/*view dataset*/  
proc print data=data1;
```

Obs	name	sales
1	Bob&%^	45
2	M&\$#@ike	50
3	Randy))	39
4	Chad!?	14
5	Dan**	29
6	R[on]	44

Notice that the values in the **name** column contain several special characters.

We can use the **COMPRESS** function to remove these special characters:

```
/*create second dataset with special characters removed from names*/
```

```
data data2;
```

```
set data1;
```

```
new_name=compress(name, , 'kas');
```

```
run;
```

```
/*view dataset*/
```

```
proc print data=data2;
```

Obs	name	sales	new_name
1	Bob&%^	45	Bob
2	M&\$#@ike	50	Mike
3	Randy))	39	Randy
4	Chad!?	14	Chad
5	Dan**	29	Dan
6	R[on]	44	Ron

Notice that the **new_name** column contains the values in the **name** column with the special characters removed.

Here's exactly what the **COMPRESS** function did to remove these special characters:

k specifies that we would like to 'keep' certain characters

a specifies to keep alphabetic characters

s specifies to keep space characters

Note: You can find a complete list of modifiers for the **COMPRESS** function on this .

ARABPSYCHOLOGY.COM