How can I use PROC FORMAT in SAS and what are some examples?

Authored by

stats writer

June 25, 2024

PROC FORMAT is a SAS procedure that allows users to define custom formats for variables in their datasets. This can be useful for converting numerical values into more easily interpretable categories, such as converting a numeric code for gender into "Male" and "Female." To use PROC FORMAT, users must first create a format catalog which stores the defined formats. They can then use the FORMAT statement in a DATA step or PROC step to apply the desired format to a variable. Examples of using PROC FORMAT include creating a custom format for age ranges, converting numerical grades into letter grades, and categorizing income levels. This procedure provides a flexible and efficient way to manipulate and present data in a more user-friendly manner.

Use PROC FORMAT in SAS (With Examples)

You can use PROC FORMAT in SAS to create a mapping of data values into data labels.

This procedure uses the following basic syntax:

```
proc format;
value points_range
25-high='High'
15-<25='Medium'
other ='Low';
run;
```

This particular example creates the following mapping:

Values equal to 25 or greater will be shown as 'High'Values between 15 and 25 will be shown as

'Medium'All other values will be shown as 'Low'

The following examples show how to use PROC FORMAT with the following dataset in SAS:

```
/*create dataset*/
data my_data;
input team $ position $ points;
datalines;
A Guard 25
A Guard 20
A Guard 30
A Forward 25
A Forward 10
B Guard 10
B Guard 22
B Forward 30
B Forward 10
B Forward 10
B Forward 25
;
run;

/*view dataset*/
```

```
proc printdata=my_data;
```

Obs	team	position	points
1	A	Guard	25
2	A	Guard	20
3	A	Guard	30
4	A	Forward	25
5	A	Forward	10
6	B	Guard	10
7	B	Guard	22
8	B	Forward	30
9	B	Forward	10
10	B	Forward	10
11	B	Forward	25

Example 1: Use PROC FORMAT to Format Values as Labels in Frequency Table

Suppose we use PROC FREQ to create a frequency table of values in the points column of the dataset:

```
/*calculate frequency of values in points column*/
proc freqdata = my_data;
table points;
run;
```

The FREQ Procedure

points	Frequency	Percent	Cumulative Frequency	Cumulative Percent
10	4	36.36	4	36.36
20	1	9.09	5	45.45
22	1	9.09	6	54.55
25	3	27.27	9	81.82
30	2	18.18	11	100.00

The output displays the frequency of each individual value in the points column.

However, suppose we would like to format the values as follows:

Values equal to 25 or greater will be shown as 'High'Values between 15 and 25 will be shown as 'Medium'All other values will be shown as 'Low'

We can use PROC FORMAT to do so:

```
/*define formatting for points variable*/proc format;
value points_range
25-high='High'
15-<25='Medium'
other ='Low';
```

```
run;

/*create frequency table for points variable, using formatting defined above*/proc freqdata = my_data;
table points;
format points points_range.;
run;
```

The FREQ Procedure

points	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Low	4	36.36	4	36.36
Medium	2	18.18	6	54.55
High	5	45.45	11	100.00

The frequency table now groups the values of the points variable into the labels that we specified using the PROC FORMAT statement.

Example 2: USE PROC FORMAT to Create New Variable

We can also use PROC FORMAT to create a new variable in a dataset that converts data values into data labels.

The following syntax shows how to do so:

```
/*define formatting for points variable*/proc format;
value points_range
25-high='High'
15-<25='Medium'
other ='Low';
run;

/*create new dataset with points_range variable*/
data new_data;
set my_data;
point_range = put(points, points_range.);
run;

/*view dataset*/
proc printdata=new_data;
```

Obs	team	position	points	point_range
1	A	Guard	25	High
2	A	Guard	20	Medium
3	A	Guard	30	High
4	A	Forward	25	High
5	A	Forward	10	Low
6	B	Guard	10	Low
7	B	Guard	22	Medium
8	B	Forward	30	High
9	B	Forward	10	Low
10	B	Forward	10	Low
11	B	Forward	25	High

The new variable called 'point_range' takes on a value of Low, Medium or High depending on the corresponding value for the 'points' variable.

Note: You can find the complete documentation for PROC FORMAT .

The following tutorials explain how to perform other common tasks in SAS: