

How can I use PROC FREQ to analyze data by group using the SAS software?

Authored by
stats writer

June 26, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I use PROC FREQ to analyze data by group using the SAS software?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=154259>

PROC FREQ is a statistical procedure in the SAS software that allows for the analysis of data by group. It is a useful tool for examining the frequency distribution of categorical variables and identifying any patterns or trends within different groups. By specifying the relevant group variables, PROC FREQ can provide detailed summary statistics, such as counts, percentages, and measures of association, for each group. This allows for a comprehensive understanding of the data and can aid in making informed decisions based on the results. Additionally, PROC FREQ offers various options for customizing the analysis, making it a versatile tool for data exploration and comparison by group.

SAS: Use PROC FREQ by Group

You can use the following basic syntax to calculate frequencies by group in SAS:

```
proc freqdata=my_data;  
by var1;  
tables var2;  
run;
```

This particular syntax creates a frequency table for the values of the variable called var2, grouped by the variable called var1.

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

Example: Using Proc FREQ by Group in SAS

Suppose we have the following dataset in SAS:

```
/*create dataset*/  
data my_data;  
input team $ position $ points;  
datalines;  
A Guard 22  
A Guard 20  
A Guard 30  
A Forward 14  
A Forward 11  
B Guard 12  
B Guard 22  
B Forward 30  
B Forward 9  
B Forward 12  
B Forward 25  
;  
run;  
  
/*view dataset*/  
proc printdata=my_data;
```

Obs	team	position	points
1	A	Guard	22
2	A	Guard	20
3	A	Guard	30
4	A	Forward	14
5	A	Forward	11
6	B	Guard	12
7	B	Guard	22
8	B	Forward	30
9	B	Forward	9
10	B	Forward	12
11	B	Forward	25

We can use the following PROC FREQ statement to calculate the frequency of the position values, grouped by team:

```
/*calculate frequency of position, grouped by team*/  
proc freqdata = my_data;  
by team;  
tables position;  
run;
```

The FREQ Procedure

team=A

position	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Forward	2	40.00	2	40.00
Guard	3	60.00	5	100.00

The FREQ Procedure

team=B

position	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Forward	4	66.67	4	66.67
Guard	2	33.33	6	100.00

The output displays the frequency of the values for the position variable, grouped by the team variable.

For example, we can see:

The value "Forward" occurred 2 times for team A. The value "Guard" occurred 3 times for team A. The value "Forward" occurred 4 times for team B. The value "Guard" occurred 2 times for team B.

Note that in this example, we used the tables statement to calculate the frequencies for just one variable, but we could type out the names of multiple variables to calculate frequencies for more than one variable.

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

ARABPSYCHOLOGY.COM