

How can I calculate percentiles in SAS and what are some examples?

Authored by
stats writer

July 1, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I calculate percentiles in SAS and what are some examples?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=164368>

Percentiles are statistical measures that are used to divide a data set into 100 equal parts. They are commonly used in data analysis to understand the distribution of a dataset and identify outliers. In SAS, percentiles can be calculated using the PROC UNIVARIATE or PROC MEANS procedures. These procedures allow you to specify the percentile that you want to calculate and provide the corresponding value. For example, if you want to calculate the 25th percentile of a dataset, you would use the syntax "P25" in PROC UNIVARIATE or the option "PCTLPTS=25" in PROC MEANS. Other examples of percentiles include the 50th percentile (also known as the median) and the 75th percentile. By calculating percentiles in SAS, you can gain valuable insights into your data and make informed decisions based on the distribution of your data.

Calculate Percentiles in SAS (With Examples)

Here are the three most common ways to calculate the percentiles of a dataset in SAS:

Method 1: Calculate One Specific Percentile Value

```
/*calculate 70th percentile value for var1*/  
proc univariate data=original_data;  
var var1;  
output out=percentile_data  
pctlpts = 70 pctlpre = P_;  
run;
```

Method 2: Calculate Multiple Specific Percentile Values

```
/*calculate 70th, 80th, and 90th percentile value for  
var1*/
```

```
proc univariate data=original_data;  
var var1;  
output out=percentile_data  
pctlpts = 70 80 90 pctlpre = P_;  
run;
```

Method 3: Calculate Percentiles by Group

```
/*sort original data by var2*/  
proc sort data=original_data;  
by var2;  
run;  
  
/*calculate percentiles for var1 grouped by var2*/  
proc univariate data=original_data;  
var var1;  
by var2;  
output out=percentile_data  
pctlpts = 70, 80, 90  
pctlpre = P_;  
run;
```

Note: The pctlpts statement specifies which percentiles to calculate and the pctlpre statement specifies the

prefix to use for the percentiles in the output.

The following examples show how to use each method with the following dataset in SAS:

```
/*create dataset*/  
data original_data;  
input team $ points;  
datalines;  
A 12  
A 15  
A 16  
A 21  
A 22  
A 25  
A 29  
A 31  
B 16  
B 22  
B 25  
B 29  
B 30  
B 31  
B 33
```

B 38

;

run;

/*view dataset*/

proc printdata=original_data;

| Obs | team | points |
|-----|------|--------|
| 1 | A | 12 |
| 2 | A | 15 |
| 3 | A | 16 |
| 4 | A | 21 |
| 5 | A | 22 |
| 6 | A | 25 |
| 7 | A | 29 |
| 8 | A | 31 |
| 9 | B | 16 |
| 10 | B | 22 |
| 11 | B | 25 |
| 12 | B | 29 |
| 13 | B | 30 |
| 14 | B | 31 |
| 15 | B | 33 |
| 16 | B | 38 |

Example 1: Calculate One Specific Percentile Value

The following code shows how to calculate the 70th percentile for the points variable:

/*calculate 70th percentile value for points*/

```
proc univariate data=original_data;  
var points;  
output out=percentile_data  
pctlpts = 70 pctlpre = P_;  
run;
```

```
/*view results*/proc print data=percentile_data;
```

| Obs | P_70 |
|-----|------|
| 1 | 30 |

The value at the 70th percentile turns out to be 30.

Example 2: Calculate Multiple Specific Percentile Values

```
/*calculate 70th, 80th, and 90th percentile value for  
points*/  
proc univariate data=original_data;  
var points;  
output out=percentile_data  
pctlpts = 70 80 90 pctlpre = P_;  
run;
```

| Obs | P_70 | P_80 | P_90 |
|-----|------|------|------|
| 1 | 30 | 31 | 33 |

Here's how to interpret the output:

The value at the 70th percentile is 30. The value at the 80th percentile is 31. The value at the 90th percentile is 33.

Example 3: Calculate Percentiles by Group

The following code shows how to calculate the values at the 70th, 80th, 90th, and 95th percentile for the points variable, grouped by the team variable:

```
/*sort original data by team*/  
proc sortdata=original_data;  
by team;  
run;
```

```
/*calculate percentiles for points grouped by team*/  
proc univariate data=original_data;  
var points;  
by team;  
output out=percentile_data  
pctlpts = 70, 80, 90 95 pctlpre = P_;
```

run;

| Obs | team | P_70 | P_80 | P_90 | P_95 |
|-----|------|------|------|------|------|
| 1 | A | 25 | 29 | 31 | 31 |
| 2 | B | 31 | 33 | 38 | 38 |

The output table shows the values for the 70th, 80th, 90th, and 95th percentile for the points variable for both teams A and B.

Additional Resources

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