

How can I use `_N_` in SAS?

Authored by
stats writer

June 23, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I use `_N_` in SAS?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=148087>

SAS is a powerful statistical software that allows users to manipulate, analyze, and visualize data. One of the key features of SAS is its ability to handle large datasets efficiently. To make the most out of SAS, users can utilize the variable `_N_`, which represents the current observation number in a dataset. This variable can be used in various ways, such as creating conditional statements, tracking the number of observations, and creating unique identifiers for each observation. By utilizing `_N_`, users can efficiently manage and analyze their data in SAS.

Use `_N_` in SAS (3 Examples)

The `_N_` automatic variable in SAS is used to count the number of loops of the data step.

Here are the most common ways to use `_N_` in practice:

Method 1: Use `_N_` to Select First Row in Dataset

```
data new_data;  
set original_data;  
if _N_ = 1 then output;  
run;
```

Method 2: Use `_N_` to Select First *N* Rows in Dataset

```
data new_data;  
set original_data;  
if _N_ <= 5 then output; /*select first 5 rows*/run;
```

Method 3: Use `_N_` to Add Row Numbers to Dataset

```
data new_data;  
set original_data;  
row_number = _N_;  
run;
```

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

```
/*create dataset*/  
data original_data;  
input team $ points rebounds;  
datalines;  
Warriors 25 8  
Wizards 18 12  
Rockets 22 6  
Celtics 24 11  
Thunder 27 14  
Spurs 33 19  
Nets 31 20  
Mavericks 34 10  
Kings 22 11  
Pelicans 39 23  
;  
run;
```

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

Obs	team	points	rebounds
1	Warriors	25	8
2	Wizards	18	12
3	Rockets	22	6
4	Celtics	24	11
5	Thunder	27	14
6	Spurs	33	19
7	Nets	31	20
8	Maverick	34	10
9	Kings	22	11
10	Pelicans	39	23

Example 1: Use `_N_` to Select First Row in Dataset

The following code shows how to use `_N_` to select just the first row of the dataset:

```
/*create new dataset that contains only the first row*/  
data new_data;  
set original_data;  
if _N_ = 1 then output;  
run;
```

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

Obs	team	points	rebounds
1	Warriors	25	8

We can see that the new dataset contains only the first row of the original dataset.

Example 2: Use `_N_` to Select First N Rows in Dataset

`/*create new dataset that contains first 5 rows of original dataset*/`

`data new_data;`

`set original_data;`

`if _N_ <= 5 then output;`

`run;`

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

Obs	team	points	rebounds
1	Warriors	25	8
2	Wizards	18	12
3	Rockets	22	6
4	Celtics	24	11
5	Thunder	27	14

We can see that the new dataset contains only the first five rows of the original dataset.

Example 3: Use _N_ to Add Row Numbers to Dataset

The following code shows how to use _N_ to add a column that displays row numbers for each row in the dataset:

```
/*create new dataset that contains column with row numbers*/  
data new_data;  
set original_data;  
row_number = _N_;run;  
  
/*view new dataset*/proc printdata=new_data;
```

Obs	team	points	rebounds	row_number
1	Warriors	25	8	1
2	Wizards	18	12	2
3	Rockets	22	6	3
4	Celtics	24	11	4
5	Thunder	27	14	5
6	Spurs	33	19	6
7	Nets	31	20	7
8	Maverick	34	10	8
9	Kings	22	11	9
10	Pelicans	39	23	10

We can see that the new dataset has a column called

row_number that contains the row number for each row in the dataset.

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

ARABPSYCHOLOGY.COM