

How can the DATA Step be used in SAS?

Authored by
stats writer

June 25, 2024

RECOMMENDED CITATION

stats writer (2024). *How can the DATA Step be used in SAS?*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=151635>

The DATA Step is a fundamental component in SAS software that allows users to read, manipulate, and create data sets. It is a powerful tool that can be used to perform data transformation, data cleaning, and data analysis tasks. The DATA Step uses a programming language to create custom data sets, allowing for flexibility and customization in data management. It is a versatile and efficient method for processing large amounts of data and can be used to merge, sort, and filter data sets. With its wide range of functions and capabilities, the DATA Step is an essential feature for data management and analysis in SAS.

Use the DATA Step in SAS (With Examples)

You can use the DATA step in SAS to create datasets.

There are two common ways to use the DATA step:

1. Create a dataset from scratch.
2. Create a dataset from an existing dataset.

The following examples show how to use each method in practice.

Example 1: Use DATA Step to Create Dataset from Scratch

The following syntax shows how to use the DATA step to create a dataset with three variables:

```
/*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

Here is exactly what we did in this example:

First, we used data to name the dataset.

Then, we used input to specify the variable names (\$ specifies a character variable).

Then, we used datalines to tell SAS that the upcoming lines represented values in the dataset.

Example 2: Use DATA Step to Create Dataset from Existing Dataset

We can use the data step along with the set statement to create a dataset from another dataset that already exists.

For example, we can use the following syntax to create a new dataset called `new_data` that uses the variables from the dataset called `my_data` but drops the 'returns' variable:

```
/*create new dataset that drops returns from my_data*/  
data new_data;  
set my_data;  
drop returns;  
run;  
  
/*view dataset*/  
proc printdata=new_data;
```

Obs	store	sales
1	A	10
2	A	7
3	A	7
4	A	8
5	A	6
6	B	10
7	B	14
8	B	13
9	B	9
10	B	5
11	C	12
12	C	10
13	C	10
14	C	12
15	C	9

Here is exactly what we did in this example:

First, we used data to name the new dataset.

Then, we used set to specify the existing dataset to create the new dataset from.

Then, we used drop to drop the 'returns' variable from the new dataset.

The end result is a new dataset contains the exact same variables from the original dataset except the 'returns' variable has been dropped.

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

ARABPSYCHOLOGY.COM