

# How do you perform an Inner Join in SAS? Can you provide an example?

Authored by  
**stats writer**

July 1, 2024

## RECOMMENDED CITATION

stats writer (2024). *How do you perform an Inner Join in SAS? Can you provide an example?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=164294>

An Inner Join in SAS is a data manipulation technique used to combine data from two or more tables based on a common key or variable. This type of join only includes rows from the tables that have matching values in the specified key or variable. To perform an Inner Join in SAS, one must use the MERGE statement and specify the tables to be joined, as well as the common key or variable. An example of this would be merging a customer table with an order table based on the customer ID, resulting in a table with only the customers who have placed orders.

## Perform an Inner Join in SAS (With Example)

You can use the following basic syntax to perform an inner join with two datasets in SAS:

```
proc sql;  
create table final_table asselect * from data1 as x join  
data2 as y  
on x.ID = y.ID;  
quit;
```

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

**Related:**

**Example: Inner Join in SAS**

**Suppose we have the following two datasets in SAS:**

```
/*create datasets*/
```

```
data data1;  
input team $ points;  
datalines;  
Mavs 99  
Spurs 93  
Rockets 88  
Thunder 91  
Warriors 104  
Cavs 93  
Nets 90  
Hawks 91  
;  
run;
```

  

```
data data2;  
input team $ rebounds;  
datalines;  
Mavs 21  
Spurs 18  
Warriors 27  
Hawks 29  
Knicks 40  
Raptors 30  
;
```

**run;**

**/\*view datasets\*/**

**proc printdata=data1;**

**proc printdata=data2;**

Obs	team	points
1	Mavs	99
2	Spurs	93
3	Rockets	88
4	Thunder	91
5	Warriors	104
6	Cavs	93
7	Nets	90
8	Hawks	91

Obs	team	rebounds
1	Mavs	21
2	Spurs	18
3	Warriors	27
4	Hawks	29
5	Knicks	40
6	Raptors	30

**Notice that the two datasets share one variable in common: team.**

**We will use the following syntax to perform an inner**

join and create a new dataset that contains only the rows in which the team variable shows up in both datasets:

```
/*perform inner join*/  
proc sql;  
create table final_table asselect * from data1 as x join  
data2 as y  
on x.team = y.team;  
quit;
```

```
/*view results of inner join*/  
proc printdata=final_table;
```

Obs	team	points	rebounds
1	Mavs	99	21
2	Spurs	93	18
3	Warriors	104	27
4	Hawks	91	29

The resulting dataset contains only the rows in which the team variable appeared in both datasets.

If you refer to the two datasets from earlier, you'll notice that there are only four teams that appear in both

**datasets: Mavs, Spurs, Warriors, and Hawks.**

**Since we chose to join the two datasets on the team variable, these are the four teams that also appear in the final dataset.**

### **Additional Resources**

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

ARABPSYCHOLOGY.COM