

# How do I merge datasets using the SAS method, based on two variables?

Authored by  
**stats writer**

June 23, 2024

## RECOMMENDED CITATION

stats writer (2024). *How do I merge datasets using the SAS method, based on two variables?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=149449>

The process of merging datasets using the SAS method involves combining two datasets based on two variables that are present in both datasets. This method allows for the efficient integration of data from multiple sources, utilizing the matching variables to create a single, comprehensive dataset. By merging datasets using the SAS method, users can easily access and analyze data from different sources, facilitating accurate and insightful data analysis. This method is commonly used in data manipulation and management tasks, making it an essential tool for data professionals.

## **SAS: Merge Datasets Based on Two Variables**

**You can use the following basic syntax to merge two datasets in SAS based on two variables matching:**

```
data final_data;  
merge data1 (in = a) data2 (in = b);  
by ID Store;  
if a and b;  
run;
```

**This particular example merges the datasets called data1 and data2 based on the variables called ID and Store and only returns the rows where a value exists in both datasets.**

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

## Example: Merge Datasets in SAS Based on Two Variables

Suppose we have the following dataset in SAS that contains information about sales associates at some company:

```
/*create first dataset*/
```

```
data data1;
```

```
input ID Store $;
```

```
datalines;
```

```
1 A
```

```
1 B
```

```
1 C
```

```
2 A
```

```
2 C
```

```
3 A
```

```
3 B
```

```
;
```

```
run;
```

```
/*view first dataset*/
```

```
title "data1";
```

```
proc printdata = data1;
```

**data1**

Obs	ID	Store
1	1	A
2	1	B
3	1	C
4	2	A
5	2	C
6	3	A
7	3	B

And suppose we have another dataset that contains information about the sales made at various stores by each associate:

```
/*create second dataset*/
```

```
data data2;
```

```
input ID Store $ Sales;
```

```
datalines;
```

```
1 A 22
```

```
1 B 25
```

```
2 A 40
```

```
2 B 24
```

```
2 C 29
```

```
3 A 12
```

```
3 B 15
```

```
;
```

```
run;
```

```
/*view second dataset*/
```

```
title "data2";
```

```
proc printdata = data2;
```

Obs	ID	Store	Sales
1	1	A	22
2	1	B	25
3	2	A	40
4	2	B	24
5	2	C	29
6	3	A	12
7	3	B	15

We can use the following merge statement to merge the two datasets based on matching values in the `ID` and `Store` columns, then only return rows where a value exists in both columns:

```
/*perform merge*/data final_data;
```

```
merge data1 (in = a) data2 (in = b);
```

```
by ID Store;
```

```
if a and b;
```

```
run;
```

```
/*view results*/  
title "final_data";  
proc printdata=final_data;
```

**final\_data**

Obs	ID	Store	Sales
1	1	A	22
2	1	B	25
3	2	A	40
4	2	C	29
5	3	A	12
6	3	B	15

The resulting dataset returns the rows where the values in the ID and Store columns both match.

**Note:** You can find the complete documentation for the SAS merge statement .