

# How can I convert a SAS data file to an Mplus data file?

Authored by  
**stats writer**

July 1, 2024

## RECOMMENDED CITATION

stats writer (2024). *How can I convert a SAS data file to an Mplus data file?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=165013>

Converting a SAS data file to an Mplus data file involves the process of transforming the data from one statistical software format to another. This can be done by using the "save as" function in SAS to export the data as a text file, which can then be imported into Mplus. The data must be formatted correctly in order to be compatible with Mplus, which may require adjusting the variable names, variable types, and any missing data codes. Once the data is successfully transferred to Mplus, it can be used for statistical analysis and modeling using Mplus's unique features and capabilities. This conversion process allows for the seamless integration of data between SAS and Mplus, providing researchers with the ability to utilize the strengths of both software programs for their data analysis needs.

## How can I convert a SAS data file to an Mplus data file? | Mplus FAQ

Suppose that you have an SAS data file called <https://stats.idre.ucla.edu/wp-content/uploads/2016/02/sample.sas7bdat> that you would like to convert to Mplus for analyzing there. Here is a listing of the first 10 observations from this file.

```
proc print data="c:mplusfaqsample"(obs=10);  
run;
```

**Obs FEMALE RACE SES SCHTYP READ WRITE**

**1 . 3 1 1 34 35**

**2 0 A 2 1 44 41**

**3 0 4 B 1 55 39**  
**4 1 2 3 C 60 59**  
**5 0 4 1 1 D 37**  
**6 0 4 2 1 34 E**  
**7 0 3 2 1 34 37**  
**8 1 4 1 1 35 35**  
**9 0 4 3 1 44 33**  
**10 1 4 3 2 36 57**

**Note that female has at least one missing value (denoted by .).**

**Also, the other variables have user defined missing values defined as**

**follows -- race has missing values coded as .A and ses has**

**missing values coded as .B and sctyp has missing values as .C**

**and read has .D as missing and finally write has .E as a missing value. You can use the following steps to ease the process**

**of converting your file into an Mplus file. We do this in these steps.**

## 1. Get descriptive statistics for your current file

We can get the descriptive statistics for our file like this. Normally we would use proc means, however we want to get

descriptive statistics with listwise deletion (for comparability with

Mplus results) so we will use proc corr with the nomiss option and that will give us descriptive statistics with listwise deletion of missing data, as illustrated below.

```
proc corr data="c:mplusfaqsample" nomiss;  
run;
```

Here are our results.

### Simple Statistics

Variable N Mean Std Dev Sum Minimum Maximum

FEMALE 194 0.55155 0.49862 107.00000 0 1.00000

RACE 194 3.41237 1.05056 662.00000 1.00000 4.00000

SES 194 2.06186 0.72433 400.00000 1.00000 3.00000

SCHTYP 194 1.16495 0.37209 226.00000 1.00000 2.00000

```
READ 194 52.48454 10.14773 10182 28.00000 76.00000  
WRITE 194 53.14948 9.25423 10311 31.00000 67.00000
```

## 2. Convert all of the missing values to a single missing value code

This step converts all of the missing values (the system missings and the user missings) into a single code, -1234. You can pick any integer value you wish since that was a value that was easy to remember and not a valid value for any of our variables.

```
data sample2;  
set "c:mplusfaqsample";  
  
array allvars _numeric_ ;  
  
do over allvars;  
if missing(allvars) then allvars = -1234 ;  
end;  
run;  
  
proc print data=sample2(obs=10);
```

**run;**

**We show some of the cases below and see that it appears the missings were converted to -1234.**

**Obs FEMALE RACE SES SCHTYP READ WRITE**

**1 -1234 3 1 1 34 35**

**2 0 -1234 2 1 44 41**

**3 0 4 -1234 1 55 39**

**4 1 2 3 -1234 60 59**

**5 0 4 1 1 -1234 37**

**6 0 4 2 1 34 -1234**

**7 0 3 2 1 34 37**

**8 1 4 1 1 35 35**

**9 0 4 3 1 44 33**

**10 1 4 3 2 36 57**

**We also use proc means to examine our data and see that our N is now 200, indicating that SAS does not see any missing data (i.e. all of the missing values have been converted into a non missing value).**

```
options nolabel ;  
proc means data=sample2;  
run;
```

## The MEANS Procedure

```
Variable N Mean Std Dev Minimum Maximum  
-----  
FEMALE 200 -5.6300000 87.2967739 -1234.00 1.0000000  
RACE 200 -2.7750000 87.5044543 -1234.00 4.0000000  
SES 200 -4.1250000 87.4053077 -1234.00 3.0000000  
SCHTYP 200 -5.0150000 87.3398306 -1234.00 2.0000000  
READ 200 45.8750000 91.5252629 -1234.00 76.0000000  
WRITE 200 46.4400000 91.4773026 -1234.00 67.0000000  
-----
```

### 3. Write out the data to a file

In this next step, we write out the data to a file  
c:sample.dat

.

```
proc export data=sample2 outfile='c:sample.dat'  
dbms=dlm replace ;
```

**run;**

**Note that this file has the variable names on line 1 and then the data below that. Here are the first few lines of this file.**

**FEMALE RACE SES SCHTYP READ WRITE**

**-1234 3 1 1 34 35**

**0 -1234 2 1 44 41**

**0 4 -1234 1 55 39**

**1 2 3 -1234 60 59**

**0 4 1 1 -1234 37**

**0 4 2 1 34 -1234**

**0 3 2 1 34 37**

**1 4 1 1 35 35**

**0 4 3 1 44 33**

**1 4 3 2 36 57**

#### **4. Modify the data file and make your Mplus program**

**Now we will edit the file and at the same time create our Mplus program for reading the data. We want to create an Mplus program**

that looks like the one below, so we will edit the data file (c:sample.dat) and cut the variable names from line 1 and then paste them into the Mplus template program below. When editing the data file, you can use notepad to edit the file if your file is small, but if it is larger then you will need to use something like wordpad, but be careful to save the file as a text only file.

**Title:**

**Data:**

**File is c:sample.dat ;**

**Variable:**

**Names are**

**FEMALE RACE SES SHTYP READ WRITE ;**

**Missing = all(-1234) ;**

**Analysis:**

**Type = basic meanstructure ;**

**Output:**

**sampstat;**

## Note

that we use `Type = basic meanstructure ;` to get listwise deletion

of missing data to match the results of step 1. And here are excerpts of the results below. You can see that the means

correspond to the means from step 1, and the variances along the diagonal of

the covariance matrix match the variances (the standard deviations squared) from step 1. This would suggest

that the transfer was successful. You can now modify the Mplus program as you wish to run whatever analysis you like.

## Means

**FEMALE RACE SES SHTYP READ**

---

**0.552 3.412 2.062 1.165 52.485**

## WRITE

---

**53.149**

## Covariances

## FEMALE RACE SES SCHTYP READ

---

**FEMALE 0.249**

**RACE 0.020 1.104**

**SES -0.050 0.161 0.525**

**SCHTYP 0.002 0.046 0.036 0.138**

**READ -0.367 2.623 2.017 0.293 102.976**

**WRITE 1.104 2.301 1.239 0.395 54.357**

### Covariances

**WRITE**

---

**WRITE 85.641**

ARABPSYCHOLOGY.COM