



ArcelorMittal

# SAS Programming – TIP for TODAY

GHSUG- October 21, 2016

*ArcelorMittal*  
*Lesley Harschnitz*  
*lesley.harschnitz@arcelormittal.com*



# Data Set Options

- Data set options are conditions that can be applied to a dataset on read and/or write processing.
- Depending on the option, the use may be restricted to certain situations.
- Data Set options can be used when a data set is specified by name.

- Example 1: In a data step

```
Data sample(keep=var1 var2);  
  merge part1(in=A) part2(in=B);  
  by var1;  
  if A and B;  
Run;
```

IN= option: use only with SET, MERGE, MODIFY, UPDATE

- Example 2: In a proc step

```
Proc PRINT data=sample(obs=20);  
Run;
```

Data step options can be used to increase programming efficiency, and make coding shorter and easier to understand.



# Favorites

Option	What it does	Restricted to
IN= <i>varname</i>	Creates a flag variable that is set to 1 when the dataset contributes an observation to the current record	SET, MERGE, MODIFY, UPDATE
FIRSTOBS= <i>num</i>	Specifies the observation to start dataset processing (defaults to 1)	Valid for READ processing only
OBS= <i>num</i>	Specifies the number of observations to process (defaults to all)	
RENAME= <i>varlist</i>	Renames a variable during processing	
KEEP= <i>varlist</i> DROP= <i>varlist</i>	Keeps or drops variables with respect to processing	

A reference: <https://support.sas.com/documentation/cdl/en/ledsoptsref/63326/PDF/default/ledsoptsref.pdf>



# IN =

- Attaches a flag to the dataset that is set to 1 when the dataset contributes an observation to the record

First

Obs	num	valueA
1	1	4.8210
2	2	13.3231
3	3	5.1984
4	4	35.0768
5	5	16.7751
6	6	4.6644
7	7	38.7240
8	8	7.5737
9	9	22.2425
10	10	31.8497

Second

Obs	num	valueB	valueC	valueD
1	1	5.1391	8.2247	3.01031
2	3	32.3213	4.5305	0.11987
3	5	5.3493	4.5760	2.86865
4	7	8.9066	10.2812	1.51691
5	9	2.9205	0.2421	1.44510

Third

Obs	num	valueA	valueB	valueC	valueD
1	1	26.0014	5.1391	8.2247	3.01031
2	2	14.0782	.	.	.
3	3	15.4969	32.3213	4.5305	0.11987
4	4	29.2875	.	.	.
5	5	5.3539	5.3493	4.5760	2.86865
6	6	28.2194	.	.	.
7	7	9.6418	8.9066	10.2812	1.51691
8	8	20.2618	.	.	.
9	9	16.8062	2.9205	0.2421	1.44510
10	10	8.0355	.	.	.

```
data third;  
  merge first(in=A) second(in=B);  
  by num;  
  if A;  
run;
```

*“The information contained on this slide is fictitious and was created for illustrative purposes only. As such, this information should not be used for any other purpose.”*



# FIRSTOBS = and OBS=

- Firstobs= start with this observation
- Obs= end with this observation

Third

Obs	num	valueA	valueB	valueC	valueD
1	1	6.3589	17.0475	2.4840	0.40637
2	2	34.9603	.	.	.
3	3	20.1276	32.1499	3.9372	2.63318
4	4	26.3913	.	.	.
5	5	1.4909	21.3644	27.6637	3.06954
6	6	0.0373	.	.	.
7	7	21.7658	18.17	.	.
8	8	17.2886	.	.	.
9	9	8.4841	0.62	.	.
10	10	6.3025	.	.	.

Can also be used on a data step to test first x records

```

proc print data=third(firstobs=4);
run;

proc print data=third(obs=4);
run;

proc print data=third(firstobs=2 obs=4);
run;

```

Firstobs=4

Obs	num	valueA	valueB	valueC	valueD
4	4	26.3913	.	.	.
5	5	1.4909	21.3644	27.6637	3.06954
6	6	0.0373	.	.	.
7					
8					
9					
10					

Obs=4

Obs	num	valueA	valueB	valueC	valueD
1	1	6.3589	17.0475	2.48402	0.40637
2	2	34.9603	.	.	.
3	3	20.1276	32.1499	3.93719	2.63318
4	4	26.3913	.	.	.

Firstobs=2 Obs=4

Obs	num	valueA	valueB	valueC	valueD
2	2	34.9603	.	.	.
3	3	20.1276	32.1499	3.93719	2.63318
4	4	26.3913	.	.	.

*“The information contained on this slide is fictitious and was created for illustrative purposes only. As such, this information should not be used for any other purpose.”*

# Rename=

- Rename=(var1=new1 var2=new2.....)

Third

Obs	num	valueA	valueB	valueC	valueD
1	1	6.3589	17.0475	2.4840	0.40637
2	2	34.9603	.	.	.
3	3	20.1276	32.1499	3.9372	2.63318
4	4	26.3913	.	.	.
5	5	1.4909	21.3644	27.6637	3.06954
6	6	0.0373	.	.	.
7	7	21.7658	18.1771	1.9556	2.33847
8	8	17.2886	.	.	.
9	9	8.4841	0.6276	18.6162	2.16863
10	10	6.3025	.	.	.

Fourth

Obs	num	Test1	Test2	Test3	valueD
1	1	6.3589	17.0475	2.4840	0.40637
2	2	34.9603	.	.	.
3	3	20.1276	32.1499	3.9372	2.63318
4	4	26.3913	.	.	.
5	5	1.4909	21.3644	27.6637	3.06954
6	6	0.0373	.	.	.
7	7	21.7658	18.1771	1.9556	2.33847
8	8	17.2886	.	.	.
9	9	8.4841	0.6276	18.6162	2.16863
10	10	6.3025	.	.	.

```
data fourth;
  set third(rename=(valueA=Test1 valueB=Test2 valueC=Test3));
run;
```

Where you use this option is important....

*“The information contained on this slide is fictitious and was created for illustrative purposes only. As such, this information should not be used for any other purpose.”*



# DROP= // KEEP=

- (Drop= varlist) <<OR>> (Keep= varlist)

Fourth

Obs	num	Test1	Test2	Test3	valueD
1	1	6.3589	17.0475	2.4840	0.40637
2	2	34.9603	.	.	.
3	3	20.1276	32.1499	3.9372	2.63318
4	4	26.3913	.	.	.
5	5	1.4909	21.3644	27.6637	3.06954
6	6	0.0373	.	.	.
7	7	21.7658	18.1771	1.9556	2.33847
8	8	17.2886	.	.	.
9	9	8.4841	0.6276	18.6162	2.16863
10	10	6.3025	.	.	.

Fifth

Obs	num	Test3	valueD
1	1	2.4840	0.40637
2	2	.	.
3	3	3.9372	2.63318
4	4	.	.
5	5	27.6637	3.06954
6	6	.	.
7	7	1.9556	2.33847
8	8	.	.
9	9	18.6162	2.16863
10	10	.	.

Sixth

Obs	num	Test1	Test2	Test3
1	1	6.3589	17.0475	2.4840
2	2	34.9603	.	.
3	3	20.1276	32.1499	3.9372
4	4	26.3913	.	.
5	5	1.4909	21.3644	27.6637
6	6	0.0373	.	.
7	7	21.7658	18.1771	1.9556
8	8	17.2886	.	.
9	9	8.4841	0.6276	18.6162
10	10	6.3025	.	.

```
data fifth(keep=num ValueD Test3) sixth(keep= num Test1--Test3);  
  set fourth;  
run;
```

Select drop or keep depending on the length of the varlist....

“The information contained on this slide is fictitious and was created for illustrative purposes only. As such, this information should not be used for any other purpose.”

