

SAS Tips and Tricks

Leonard Landry, BLMA
Statistics Canada

SAS Tips and Tricks

Use of NOBS= option when subsetting with a WHERE statement.

Comparison of Concatenating datasets using a SET Statement and PROC DATASETS



Subsetting Using IF

```
23      data year99;  
24          set year8300;  
25          if year = 1999;  
26  
27      run;
```

NOTE: There were **40250204** observations read from the data set YEAR8300.

NOTE: The data set WORK.YEAR99 has **2413228** observations and 14 variables.

NOTE: DATA statement used (Total process time):

real time	5:11.07
cpu time	1:10.04

Subsetting Using WHERE

```
23 data year99;  
24     set year8300;  
25     where year = 1999;  
26  
27 run;
```

NOTE: There were **2413228** observations read from the data set YEAR8300.

NOTE: The data set WORK.YEAR99 has **2413228** observations and 14 variables.

NOTE: DATA statement used (Total process time):

real time	2:47.32
cpu time	45.80



NOBS= option

```
23 data year99;  
24     set year8300 nobs=obs end=end1;  
25     where year = 1999;  
26     if end1 then put obs=;  
27 run;
```

obs=40250204

NOTE: There were 2413228 observations read from the data set YEAR8300.

NOTE: The data set WORK.YEAR99 has 2413228 observations and 14 variables.

NOTE: DATA statement used (Total process time):

real time	2:58.87
cpu time	56.55



NOBS= option

```
SET<SAS-data-set(s) <(data-set-options(s) )>> <options>;
```

Syntax:

NOBS=variable

Creates and names a temporary variable whose value is usually the total number of observations in the input data set or data sets. If more than one data set is listed in the SET statement, NOBS= the total number of observations in the data sets that are listed. The number of observations includes those that are marked for deletion but are not yet deleted.

Caveats

- If more than one data set is listed in the SET statement, NOBS= the total number of observations in the data sets that are listed.
- The number of observations includes those that are marked for deletion but are not yet deleted.
- Do not use for stored compiled programs. At compilation time, SAS reads the descriptor portion of each data set and assigns the value of the NOBS= variable automatically.
- For certain SAS views, SAS cannot determine the number of observations. In these cases, SAS sets the value of the NOBS=variable to the largest positive integer value that is available in your operating environment.

Concatenating Datasets

```
data year8304;  
  set year8303 year04;  
run;
```

NOTE: There were 48639708 observations read from the data set WORK.YEAR8303.

NOTE: There were 2575250 observations read from the data set WORK.YEAR04.

NOTE: The data set WORK.YEAR8304 has 51214958 observations and 45 variables.

NOTE: DATA statement used (Total process time):

real time	1:06:44.73
cpu time	23:10.65

Concatenating Datasets

```
Proc datasets library=work nolist;  
  append base=year8303 data=year04;  
  delete year8304;  
  change year8303=year8304;  quit;
```

Concatenating Datasets

NOTE: Appending WORK.YEAR04 to WORK.YEAR8303.

NOTE: There were 2575250 observations read from the data set WORK.YEAR04.

NOTE: 2575250 observations added.

NOTE: The data set WORK.YEAR8303 has 51214958 observations and 45 variables.

NOTE: Deleting WORK.YEAR8304 (memtype=DATA).

NOTE: Changing the name WORK.YEAR8303 to WORK.YEAR8304 (memtype=DATA).

NOTE: PROCEDURE DATASETS used (Total process time):

real time	6:23.10
cpu time	4:19.64

SUMMARY

Comparison of Concatenating datasets using a SET Statement and PROC DATASETS

- With the SET Statement SAS reads all datasets completely and writes a new one
- With Proc Datasets SAS only reads and writes the dataset to be appended to the base dataset
- Works best when a smaller dataset is appended to a larger dataset



Questions / Comments



Statistics
Canada

Statistique
Canada

Leonard Landry

Database Manager

Business and Labour Market Analysis Division

R.H. Coats Building, 24th Floor, Section E

Ottawa, Ontario, Canada K1A 0T6

(613) 951-1284 Fax (613) 951-5403

Leonard.landry@statcan.ca

Canada

