

How to find out and use the number of observations in a given SAS data set

Fortune Walla

August 16, 2015

Created on Sunday, March 16th, 2014 at 5:43 am

If one is in a situation where they have to know the number of observations in a particular SAS data set or use it for further calculations, SAS allows many ways of doing so. Let us look at a few quick ways using the DATA step. SAS documentation also includes a SAS Macro approach which will not be replicated here.

Method 1: The most straightforward way is to load the set in a `_null_` DATA step.

```
data _null_;  
set dataset3obs;  
run;
```

The output is given by

```
NOTE: There were 3 observations read from the data set WORK.DATASET3OBS.  
NOTE: DATA statement used (Total process time):  
real time 0.00 seconds  
cpu time 0.00 seconds
```

Method 2: Another way to extract & print is to output the `NOBS=` value, but since the DATA step loops to read all observations, the `put` statement is executed at every iteration.

```
data _null_;  
set dataset3obs nobs=nobs;  
put nobs;  
run;
```

The output is given by

```
3  
3  
3  
NOTE: There were 3 observations read from the data set WORK.DATASET3OBS.  
NOTE: DATA statement used (Total process time):  
real time 0.00 seconds  
cpu time 0.00 seconds
```

Method 3: One way to overcome this & print a single value is to output the number of observations either at the beginning or at the end of DATA step.

```
data _null_;  
set dataset3obs nobs=nobs;  
if _n_=1 then put nobs;  
run;
```

The output is given by

```
3
NOTE: There were 3 observations read from the data set WORK.DATASET3OBS.
NOTE: DATA statement used (Total process time):
      real time 0.00 seconds
      cpu time 0.00 seconds
```

Method 4: Output at the end of the data loop.

```
data _null_;
set dataset3obs nobs=nobs end=last;
if last then put nobs;
run;
```

The output is given by

```
3
NOTE: There were 3 observations read from the data set WORK.DATASET3OBS.
NOTE: DATA statement used (Total process time):
      real time 0.00 seconds
      cpu time 0.00 seconds
```

Method 5: One can even extract the value before the SET statement, as SAS loads it in the descriptor portion during the compilation phase.

```
data _null_;
set dataset3obs nobs=nobs;
if _n_=1 then put nobs;
run;
```

The output is given by

```
3
NOTE: There were 3 observations read from the data set WORK.DATASET3OBS.
NOTE: DATA statement used (Total process time):
      real time 0.00 seconds
      cpu time 0.00 seconds
```

There are many other methods but these are simple for just a quick lookup.