

Using SAS System and Dataset Options to Control Your Data

David Franklin, Independent Consultant, New Hampshire, USA

The Original Data

```
data trial.crew (alter=Dads write=Army label='Wilmington-on-Sea');
infile cards;
input name $ 1-17 rank $ 19-36 sn $ 39-43;
cards;
George Mainwaring Captain 61885
Arthur Wilson Sergeant 50796
Jack Jones Lance-Corporal 96822
Joe Walker Private 52925
Frank Pike Private 57467
James Frazer Chief Petty Officer 07973
Charles Godfrey Private 45322
;
run;
```

The Data After Modifications

Obs	name	rank	sn
1	George Mainwaring	Captain	61885
2	Arthur Wilson	Sergeant	50796
3	Jack Jones	Lance-Corporal	96822
5	Frank Pike	Private	57467
6	James Frazer	Private	07973
7	Charles Godfrey	Private	45322
8	Cheeseman	Private, W.C.	98072

Printing the Audit Trail

```
proc print data=trial.crew (type=audit);
title 'Audit Trail After Modifying CREW Dataset';
run;
```

Audit Trail After Modifying CREW Dataset

Obs	name	rank	sn	reason
1	James Frazer	Chief Petty Officer	07973	
2	James Frazer	Private	07973	Corrected rank, ref. memo 2008-10-13
3	Joe Walker	Private	52925	
4	Joe Walker	Private	52925	Disappeared, ref. memo 2008-10-13
5	Joe Walker	Private	52925	
6	Cheeseman	Private, W.C.	98072	New member, ref. memo 2008-10-13

Obs	_ATDATETIME_	_ATOBSNO_	_ATRETURNCODE_	_ATUSERID_	_ATOPCODE_	_ATMESSAGE_
1	13OCT2008:15:12:30	6	.	dfrankli	DR	
2	13OCT2008:15:12:30	6	.	dfrankli	DW	
3	13OCT2008:15:12:30	4	.	dfrankli	DR	
4	13OCT2008:15:12:30	4	.	dfrankli	DW	
5	13OCT2008:15:12:30	4	.	dfrankli	DD	
6	13OCT2008:15:12:31	8	.	dfrankli	DA	

Modification code:
 DA=Record added
 DD=Record deleted
 DR=Copy of record before update
 DW=Copy of record after update
 EA=Record added failure
 ED=Record deleted failure
 EW=Record update failure.

Controlling an Audit Trail

It is possible to suspend, resume and terminate the audit trail for a dataset using the following code examples:

```
proc datasets lib=trial;
*Suspend an audit trail;
audit crew (alter=Dads); suspend;
*Resume an audit trail;
audit crew (alter=Dads); resume;
*Terminate an audit trail;
audit crew (alter=Dads); terminate;
quit;
run;
```

While it is possible to do these actions on an audit trail within a SAS dataset these are not encouraged since it will invalidate the integrity of the dataset.

TIME TO MODIFY SOME DATA WHEN AN AUDIT TRAIL IS ACTIVE

If a SAS Dataset is rebuilt, i.e. use of SET statement, then audit trail information pertaining to that dataset will be lost so some particular strategies have to be adopted when editing data in the SAS dataset.

The three tasks that are generally done in a SAS dataset are:

- modify an existing observation
- add a new observation
- delete an existing observation

The SQL procedure with UPDATE, INSERT and DELETE statements is one way of modifying the data while keeping the audit trail.

```
proc sql;
*Correct record;
update trial.crew (write=Army)
set rank='Private',
reason='Corrected rank, ref. memo 2008-10-13'
where sn='07973';
*Delete record - update first to add reason;
update trial.crew (write=Army)
set reason='Disappeared, ref. memo 2008-10-13'
where sn='52925';
delete
from trial.crew (write=Army) where sn='52925';
*Add new record;
insert into trial.crew (write=Army)
set name='Cheeseman',
rank='Private, W.C.',
sn='98072',
reason='New member, ref. memo 2008-10-13';
quit;
```

Audit Trails

Getting The Audit Trail Started In a SAS Dataset

An audit trail is useful if the dataset is in a state where few if any changes are needed, or it is a dataset that is like a list that is constantly changing. It is not good for the case where the dataset will be rebuilt many times over the life of the dataset. To start an audit trail use the AUDIT statement in the DATASETS procedure. The following initiates the audit trail to the dataset CREW with a user defined variable REASON (so the reason for a change is recorded):

```
proc datasets lib=trial;
audit crew (alter=Dads);
initiate;
user_var reason $100;
quit;
run;
```

Note that the ALTER password option was called as prior to this step the dataset was created with an ALTER password.

SAS "Versioning"

Copy the Data

One way is to use the SAS DATASETS procedure with the COPY statement and place a copy into another directory or placing it on another media.

SAS Versions

Another way is to create versions using the GENMAX= option (the following code shows how this is invoked):

```
data trial.crew (alter=Dads write=Army
label='Wilmington-on-Sea'
genmax=2);
infile cards;
input name $ 1-17 rank $ 19-36 sn $ 39-43;
cards;
George Mainwaring Captain 61885
Arthur Wilson Sergeant 50796
Jack Jones Lance-Corporal 96822
Joe Walker Private 52925
Frank Pike Private 57467
James Frazer Chief Petty Officer 07973
Charles Godfrey Private 45322
;
run;
```

The GENMAX=2 option allows for three versions of the dataset to exist at any one time with number 0 as the current version, 2 as the most recent version, and 1 as the oldest version. A maximum of 999 versions, excluding the latest version, can be set for a dataset with the oldest version "dropping off" if the new version is created and the number of versions exceed that allowable.

Working with Versions

To access a particular version of a dataset, the option GENNUM= is used and can be used both in a direct and relative reference, as the following examples show:

```
*Print current version of dataset;
proc print data=trial.crew;
run;
*Print previous version of dataset;
proc print data=trial.crew (gennum=-1);
run;
```

To compare the current data against the previous version the following code can be used:

```
proc compare base=trial.crew data=trial.crew (gennum=-1);
run;
```

Versions of a SAS dataset can be deleted using the DELETE statement in the DATASETS procedure, as the following example shows:

```
proc datasets library=trial;
*Deletes all versions except current version;
delete crew (gennum=hist);
*Deletes current version and moves previous version to current;
delete crew;
*Deletes all versions of dataset;
delete crew (gennum=all);
quit;
run;
```

Controlling Access

Controlling Access to SAS Datasets through SAS Options

There are four SAS Options that are very good for controlling access to a SAS dataset:

- ALTER= password to alter structure
- READ= password to read
- WRITE= password to modify the data
- ENCRYPT= password to encrypt the data

Conclusion

Method	Pro(s)	Con(s)
Audit Trails	Can take notes at an individual record as to why a record was modified, added or deleted.	Not easy to report on changes to a dataset.
SAS Versioning	Changes to data can be quickly seen using tools like the COMPARE procedure.	No note feature is available as to why a version was needed or any Record as to why an individual observation was changed.

Contact Information

David Franklin
 Independent SAS Consultant
 16 Roberts Road
 Litchfield, NH 03052
 Phone: 603-275-6809
 Email: dfranklinuk@compuserve.com
ourworld.compuserve.com/homepages/dfranklinuk