

## Improving Performance.

You've probably read a number of documents about SAS Performance but here are a couple of tips that are simple to implement.

Please note that on small datasets you may not see any performance improvement. These options can be applied to small datasets but you will only really notice the difference when processing large volumes of data.

### 1.1 Use NOEQUALS option

If you do have to sort a large dataset use the NOEQUALS option. This option saves the sort process from ordering duplicate sort variables by the original order. It's not to be confused with NODUPKEY or NODUPPLICATES.

NOEQUALS can save upto 40% of sort time.

```
proc sort data=BIG_DATASET noequals;
  by variable1 variable2;
run;
```

Sort with EQUALS is SAS' default action this default can be set to NOEQUALS if required.

### 1.2 Don't sort datasets unnecessarily.

Obviously your data needs to be sorted to perform merges and to summarise using a BY. However if you can manage your data so that you only need to sort the data in the first instance you wouldn't need to subsequently sort it as the sort order will be implied.

For example:

```
proc sort data=TABLE1 noequals; by KEY1;
run;
```

```
proc sort data=TABLE2 noequals; by KEY1
KEY2; run;

data LARGE;
  merge TABLE1
        TABLE2;
  by KEY1;
run;
```

The dataset LARGE has an implied sortedby of KEY1 KEY2 and so doesn't need to be sorted if merging again by KEY1 of KEY1 KEY2

### 1.3 Use SAS Compression

Compressing/decompressing data has an overhead all of it's own but as the overhead is simply on CPU this overhead (Real Time) is usually less significant than disk I/O which is reduced on a compressed dataset.

```
options compress=yes;
```

### 1.4 Careful use of WHERE

Your aware that using a WHERE is more efficient than using IF. However the efficiency is lost if you use the where with a function.

So Don't do this:

```
data test;
  set YEARS;
  where START_DATE >=
intnx('month',today(),0);
run;
```

Do this instead:

```
%let
THIS_MTH="01%sysfunc(date(),monyy7.)"d;
%put &THIS_MTH;

data test;
  set YEARS;
  where START_DATE >= &THIS_MTH;
run;
```