

Fundamental Advice to Save Time in SAS® Programming

October 12, 2016

Mickael Borne



Barcelona 2016

Fast Track to Approval: Speed and Efficiency

9th–12th October Princesa Sofia Gran Hotel, Barcelona

How can we, SAS users, be actors in the
“Fast Track to Approval”?

Introduction

□ Agenda:

- Speed Up the Coding
 - Avenues to Explore
 - SAS Interlude
 - Advice
- Speed Up the Processing
 - Avenues to Explore
 - SAS Interlude
 - Advice

SPEED UP THE CODING – Avenues to explore

☐ Good Programming Practices

- ☐ Increase the robustness and reusability of the programs and reduce logical and syntax errors
- ☐ save maintenance time and you can facilitate the transfer of code

☐ Training, Mentoring, Web articles,...

- ☐ Stay abreast of evolving technology

☐ Programming Automation

- ☐ Saves development time,
- ☐ Ensures compliance with SOPs

SPEED UP THE CODING – SAS interlude

❑ SAS Function CoMPiler (FCMP procedure)

➔ write your own user-defined functions or CALL routines

```
PROC FCMP OUTLIB=LIBNAME.DATASET.PACKAGE ;
  FUNCTION NameOfFunction (Argument1 Argument2 /*...*/) ;
    /* ... Program Statements ... */
    RETURN (Expression) ;
  ENDSUB ;

  SUBROUTINE NameOfRoutine (Argument1 Argument2 /*...*/) ;
    OUTARGS OutArgument1 OutArgument2 /*...*/ ;
    RC=/**/ ;
    RETURN ;
  ENDSUB ;

RUN ;

** Created functions have to be accessible ** ;

OPTIONS CMPLIB=LIBNAME.DATASET ;
```

SPEED UP THE CODING – Programming Advice

“Take time to plan” → Think, Write and then Code

“Execute your program step-by-step” → Open intermediate datasets

“You are not alone” → Use Web resources

“Think simplification” → Extract complex part of code

“Think modularity” → Structure logically your code

SPEED UP THE PROCESSING – Avenues to explore

☐ SAS I/O Processing

- ☐ The most time consuming part of your program
- ☐ I/O optimization is done at the SAS Setinit level
- ☐ Use of PROC SORT options or PROC DATASETS

☐ Handling Large Datasets

- ☐ Very time consuming
- ☐ Use an Index
- ☐ When possible execute code one time instead of many times

☐ Join datasets

- ☐ Compare JOIN techniques (MERGE, SET, HASH Object, PROC SQL)

SPEED UP THE PROCESSING – SAS interlude

❑ HASH Object

➔ Fastest technique to join large datasets

```
data WORK.RESULT ;*** Join 2 datasets using the HASH object *** ;
  set WORK.DATASET1 ;
  * Variables coming from HASH object have to be declared * ;
  length VAR1FROMDS2 $ 8 VAR2FROMDS2 $ 40 VAR3FROMDS2 8. ;
  call missing(PARAMN,PARAMCD,PARAM,PARDECNB) ;
  * CALL MISSING used to avoid Uninitialized notes in the log file * ;
  if (_N_=1) then do ;
    declare hash dico(dataset:"WORK.DATASET2") ;* HASH OBJECT DECLARATION * ;
    dico.defineKey('KEYVAR') ;* KEY DEFINITION * ;
    dico.defineData('VAR1FROMDS2','VAR2FROMDS2','VAR3FROMDS2') ; * DATA * ;
    dico.defineDone() ;
  end ;
  rc=dico.find(key:KEYVAR) ;* join * ;
  if (not rc) then do ;* Matching records * ;
    /* ... */
  end ;
  else do ;* Non-matching records * ;
    /* ... */
  end ;
  drop rc ;
run ;
```

SPEED UP THE PROCESSING – Programming Advice

“Datasets on a diet” → Filter, drop and length

“SAS raw dataset importation in the WORK library”

“Fast processing is not an end in itself” → Quality most important

“Decrease the data step number” → Use functions, Procedures

“What else?” → Don't forget the ELSE

Questions?



CONTACT INFORMATION

Mickael Borne

Work Phone: +33 1 84 19 34 59

Email: mborne@4clinics.com