

Keeping Up with Updating Dictionaries

Divyaja Padamati, Eliassen Group Inc., North Carolina.

ABSTRACT

Often in Pharmaceutical Industries we work with data for poster presentations, manuscripts or for additional exploratory analyses after the CSR submission. Here comes the important question. How often do we check for updated dictionary versions and what do we do when this occurs? In this paper we shall discuss about; (1) Different checks that are required when the data is using updated dictionary terms. (2) If updated dictionary terms are not used, then how should we proceed? The skill level needed will be intermediate to advanced SAS programmer.

INTRODUCTION

For SAS programmers, first thing that comes to our mind when we hear the word “dictionary” is mapping team. We mostly give importance to the preferred term from dictionary that we use for any CSR reports/TLFs. It is one of the major responsibilities of a programmer to ensure that appropriate data has been used.

As I mentioned earlier, if the study is still ongoing and we are working on CSR submission then we mostly have data which uses the current dictionary. We cannot assume the same when we are working in a team where the manuscripts or presentations are being done way after the submission is done. It is vital to know about the dictionary versions in this case.

Let's see how we can programmatically check for the version of dictionary used in the data. Based on the existence of the version in the data there will be two scenarios we come across:

- Data with version variable.
- Data without version variable.

In the first case we shall compare the versions from data and dictionary directly. In the second case where we do not have version variable, then we check the data creation date and compare it with the latest dictionary date.

Based on the output of the above steps the team can decide whether to use the updated dictionary or use the same version used in the previous submissions made using this data.

MACRO INPUT PARAMETERS

INDT	Input dataset
DICT	Dictionary location
STUDY	Data location
DICTIONARY	Current Dictionary

```
%macro dictionary_check (study=,dict=,indt=,dictionary=) ;

libname study "&study." access=readonly ;
libname diction "&dict." access=readonly ;

/* Getting all the variable names into a MV */
proc sql noprint ;
  select name into :var_name separated by ' ' from dictionary.columns
  where strip(lowercase(memname))="adcm" and strip(lowercase(libname))="study" ;
quit ;

/* Creating a MV version based on the existence of version variable */
data _null_ ;
  if find("&var_name", 'version', 'i') then call symputx ('ver',1) ;
  else call symputx ('ver',0) ;
run ;
%put &ver ;
```

```
%if &ver. = 1 %then %do ;

  proc sql noprint ;
    select version into :dat_ver from study.&indt. ;
    select version into :dict_ver from diction.&dictionary. ;
  quit ;

/* Comparing the dictionary versions */

  %if "&dat_ver" ne "&dict_ver" %then %do ;
    dm "log;clear;" ;
    %put ***** ;
    %put *** DATA DICTIONARY HAS BEEN UPDATED. DICTIONARY VERSION: &DICT_VER., DATA
VERSION: &DAT_VER ;
    %put ***** ;
  %end ;

%end ;

%else %do ;

/* need to compare the dates as version variable does not exist */
  proc contents data = study.&indt. out=data_time noprint ;
  run ;
  proc contents data = diction.&dictionary out=dict_time noprint ;
  run ;

  proc sql ;
    select distinct (datepart(crdate)) into :data_crdate from data_time ;
    select distinct (datepart(crdate)) into :dict_crdate from dict_time ;
  quit ;

  %if &data_crdate lt &dict_crdate %then %do ;
    dm "log;clear;" ;
    %put ***** ;
    %put *** DATA DICTIONARY HAS BEEN UPDATED. DICTIONARY DATE: &dict_crdate,
DATA DATE: &data_crdate. ;
    %put ***** ;
  %end ;
%end ;
%mend dictionary_check ;
```

SAMPLE CALL:

```
%dictionary_check(study=%str(/study/data), dict=%str(/submission/dictionary),
indt=adcm, dictionary=whodrug );
```

CONCLUSION

Working with a data involving dictionaries which keep updating on regular basis is itself a challenge for SAS programmers. This paper gives an idea about how to identify the dictionaries used in the data. Once the version has been identified, further action will be taken based on the requirements.

REFERENCES

SAS® 9.2 Macro Language: Reference. Cary, NC: SAS Institute Inc.
SAS® 9.2 Language Reference: Dictionary. Cary, NC: SAS Institute Inc.

ACKNOWLEDGMENTS

Eliassen Group Biometrics and Data Solutions
1005 Slater Rd #100,
Durham, NC 27703

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Divyaja R Padamati
Enterprise: Eliassen Group Biometrics and Data Solutions
Address: 1005 Slater Rd #100
City, State ZIP: Durham, NC 27703
E-mail: dpadamati@eliassen.com
Web: egbiometrics.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. © indicates USA registration.

Other brand and product names are trademarks of their respective companies.