

When PROC COMPARE is not enough - the fast check of variables discrepancy through thousands of observations

ROMAN CHEKHUN

Different number of observations – a bad scenario for the QC validation process

- Standard PROC COMPARE is too cumbersome.
- First of all, aim for equal number of observations in this case.

Macro “subj_check”

Details the following:

- Number of records for each subject;
- Showing subjects and number of records if it's different;
- --TESTCD variable for the Findings Observations Class;
- Showing subjects and number of records for each non-consistent --TESTCD;
- Difference in number of records for each value of QNAM in Supplemental Qualifiers;
- Difference in number of records for each subject with specific QNAM in Supplemental Qualifiers;

w_datacutcomp	subj_check
Designed for data transfer process.	Designed for validation process with standard domains.
Designed to check the difference between versions of raw data.	Designed to check the difference between the same SDTM/ADaM domain datasets with the different numbers of records.
Works with file system and directories	Works with datasets in temporary folder Work
Works with the scope of raw dataset different versions	Works with two datasets of the same SDTM domain
Shows new variables that are not in old raw data	Shows difference in the number of records and values for specific variables
Reports result in PDF files.	Reports result in HTML in Results Viewer

Differences
between
another macro
with similar
functions

Advantages over PROC COMPARE

Compact table display of irrelevant data.

Ability to save results of comparing in datasets with subsequent use.

Case 1: Difference in number of observations for each subject.

```
%macro subj_check(domain, opt=, qnam=);
```

Example: two SDTM datasets AE, QCAE have different number of observations.



```
%subj_check(AE, opt=sdm);
```



USUBJID	QC_num	PROD_num
CDISC01.100008	22	12
CDISC01.100014	17	12
CDISC01.200001	22	12
CDISC01.200002	22	12

Case 2: Difference in number of observations for --TESTCD for each subject.*

```
%macro subj_check(domain, opt=, qnam=);
```

Example: two SDTM datasets LB, QCLB have different number of observations.



```
%subj_check(LB, opt=sdm);
```



*The same way it can be compared ADaM dataset with PARAMCD variable.

```
%subj_check(ADLB, opt=adam);
```

LBTESTCD	USUBJID	QC_test_num	PROD_test_num
BILI	CDISC01.100008	2	1
BILI	CDISC01.100014	2	1
BILI	CDISC01.200001	2	1
BILI	CDISC01.200002	2	1
BUN	CDISC01.100008	2	1
BUN	CDISC01.100014	2	1
BUN	CDISC01.200001	2	1
BUN	CDISC01.200002	2	1
GLUC	CDISC01.100008	6	3
GLUC	CDISC01.200001	6	3
GLUC	CDISC01.200002	6	3
HCT	CDISC01.100008	2	1
HCT	CDISC01.100014	2	1
HCT	CDISC01.200001	2	1
HCT	CDISC01.200002	2	1

Case 3: Difference in number of observations for each subject with specific QNAM in Supplemental Qualifiers.

```
%macro subj_check(domain, opt=, qnam=);
```

Example: two Supplemental Qualifiers datasets SUPPAE, SUPQCAE have different number of observations.



```
%subj_check(AE, opt=supp,  
qnam="SUPPVAR1" "SUPPVAR4");
```



USUBJID	QNAM	QC_num	PROD_num
820101	SUPPVAR1	29	9
820101	SUPPVAR4	25	14
820104	SUPPVAR1	11	6
820104	SUPPVAR4	16	11
820105	SUPPVAR1	7	
820105	SUPPVAR4	3	
820106	SUPPVAR1	16	
820106	SUPPVAR4	17	
820107	SUPPVAR1	4	
820107	SUPPVAR4	5	
820201	SUPPVAR1	14	
820201	SUPPVAR4	1	

Case 4: Difference in number of observations for each value of the same number of QNAM parameters in Supplemental Qualifiers

```
%macro subj_check(domain, opt=, qnam=);
```

Example: two Supplemental Qualifiers datasets SUPPAE, SUPPQCAE have different number of observations.



```
%subj_check(AE, opt=supp);
```



QNAM	QC_num	PROD_num
SUPPVAR1	208	37
SUPPVAR2	208	37
SUPPVAR4	112	34
SUPPVAR5	7	4
SUPPVAR6	221	36
SUPPVAR7	104	34
SUPPVAR8	135	32
SUPPVAR9	225	35

Case 5: Difference in number of observations for each value of the different number of QNAM parameters in Supplemental Qualifiers

```
%macro subj_check(domain, opt=, qnam=);
```

Example: two Supplemental Qualifiers datasets SUPPAE, SUPPQCAE have different number of observations.



```
%subj_check(AE, opt=supp);
```



QNAM	QC_num	PROD_num
SUPP11	112	
SUPP44	104	
SUPPVAR1	208	37
SUPPVAR2	208	37
SUPPVAR4	0	34
SUPPVAR5	7	4
SUPPVAR6	221	36
SUPPVAR7	0	34
SUPPVAR8	135	32
SUPPVAR9	225	35

Thank You

Name: Roman Chekhun

Organization: Quartesian

E-mail: roman.chekhun@quartesian.com

Web: <https://www.quartesian.com>

