A Macro to Avoid P21 FDAC036 Error Message for Regulatory Submission Datasets.

Prabhakara BURMA

PRINCIPAL SAS ANALYST

Acerta Pharma LLC

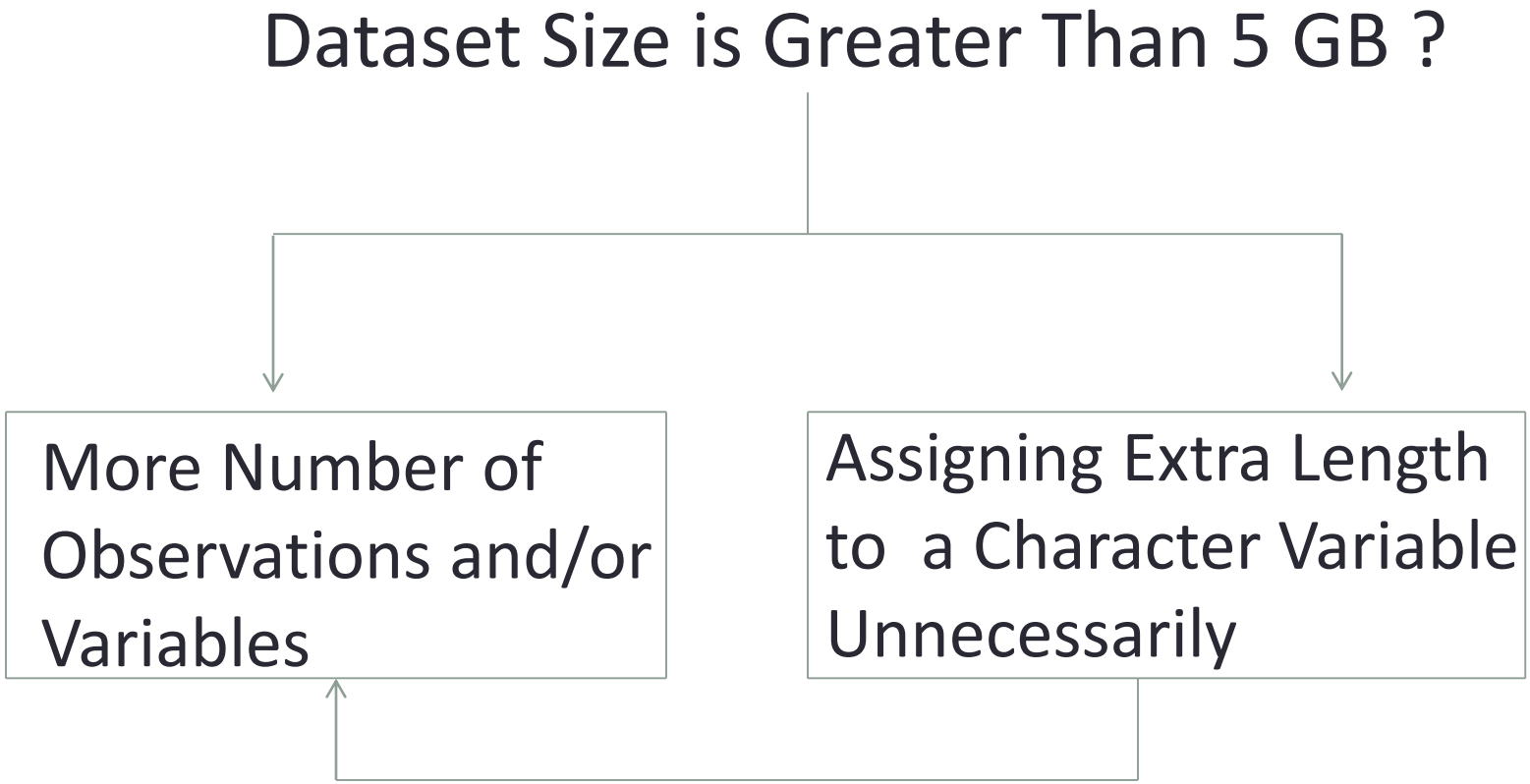
South San Francisco

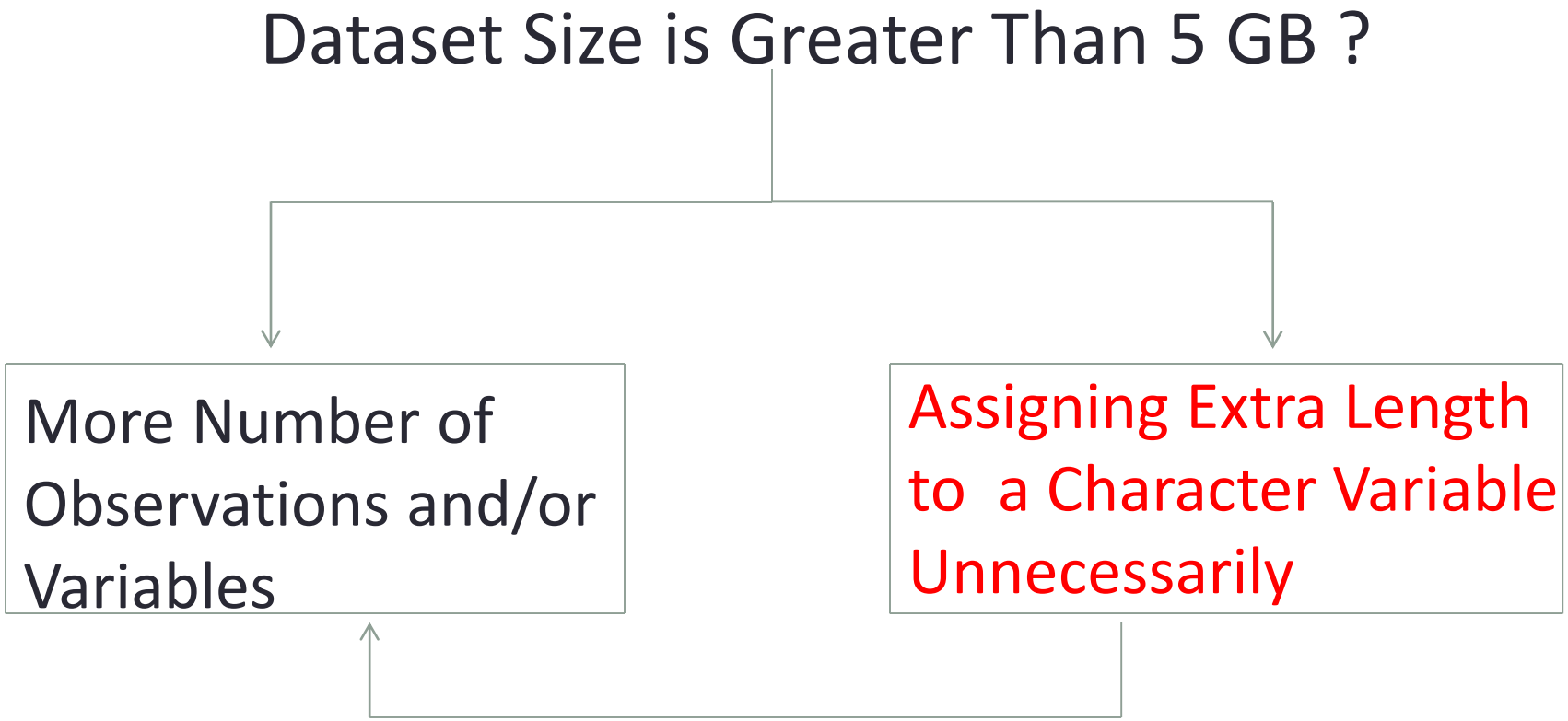
Bay Area SAS Users Group – 2018

Introduction

- As per the FDA requirements, the submission datasets must follow CDISC guidelines.
- Submission datasets must be in SAS V5 Transport file format.
- The overall size of the each individual file should not exceed 5 GB.
- Transport files to be validated for Data Conformance using Pinnacle 21 (P21) validator prior to submission.

- Pinnacle 21 validator will provide Error and Warning Messages if the transport file violates CDISC guidelines.
- One of the Error messages is **FDAC036** (Variable length is too long for actual data)
- All errors and warning messages given by the P21 validator should be addressed by the programming team.
- Addressing FDAC036 error message will significantly reduces the file size. (Note: Each .xpt file size <= 5GB)


Dataset Size is Greater Than 5 GB ?
More Number of Observations and/or Variables
Assigning Extra Length to a Character Variable Unnecessarily


Dataset Size is Greater Than 5 GB ?
More Number of Observations and/or Variables
Assigning Extra Length to a Character Variable Unnecessarily

Variable with Extra Length

Domain	Character Variable Name	Assigned Length in Spec	Actual Variable Value Max Length	Extra Length
AE	AETERM, AEDECOD etc..	200	60	140
DS	DSTERM	200	40	160
LB	LBCAT etc..	200	50	150

Remove Unnecessary Length for Character Variable

Reassign length for Character Variable with its maximum value length

Domain	Character Variable Name	Assigned Length in Spec	Actual Variable Value Length	Extra Length (remove)	New Length (reassign)
AE	AETERM	200	60	140	60
DS	DSTERM	200	40	160	40
LB	LBCAT	200	50	150	50

ADJLEN Macro

- **Advantages**:

- Significantly decrease the file sizes

- Avoid P21 FDAC036 (Variable length is too long for actual data) Error Message

Mechanism for ADJLEN

- %adjlen(dsn=,exc=)

Keyword Parameter	Parameter Value	Description
DSN	Input dataset eg: DM, LB, ...	%adjlen(dsn=dm,exc=)
EXC	Exclude the list of variables for which the length's should not be changed. eg: %str("DOMAIN", "STUDYID" ,"USUBJID")	%adjlen(dsn=dm,exc=%str("STUDYID"))

Macro Code

```
%macro adjlen (dsn=,exc=);
%if %sysfunc(exist(tempcont chk)) eq 1 or
%sysfunc(exist(max_count)) eq 1 or
%sysfunc(exist(tp)) eq 1 %then %do;
%put WARNING: Datasets name should not be identical with the
following names: tempcont chk, max_count, tp ;
%abort cancel;
%end;
proc contents data=&dsn out= tempcont (keep=name length type where=(TYPE eq 2 and
NAME not in (&exc))) noprint;
run;
proc sql noprint;
select 'max (length('||compress(name)||')) as '||compress(name) into: var
separated by ', ' from tempcont;
quit;
proc sql noprint; create table max_count as select &var from &dsn; quit;
proc transpose data= max_count out= tp;
run;
proc sql noprint;
select compress(_name_)||' character ('||strip(put(col1,best.))||')' into:
variable separated by ', ' from tp; alter table &dsn modify &variable;
quit;
proc datasets lib=work memtype=data nolist;
delete tempcont chk max_count tp ;
quit;
%mend;
```

ADJLEN Application

- **STEP-I**: Generating CARS dataset for ADJLEN macro illustration

```
data _1;
length make $30 model $80 type origin $60;
make="";
model="";
type="";
origin="";
run;

data cars;
set sashelp.cars;
keep make model type origin DriveTrain;
run;

data cars;
set _1 cars;
run;
```

STEP II: Original value length of each character variable in the dataset CARS.

Column Name	Type	Length
Aa make	Text	30
Aa model	Text	80
Aa type	Text	60
Aa origin	Text	60
Aa DriveTrain	Text	60

Actual value length of each character variable in the dataset CARS.

Column Name	Type	Length
Aa Make	Text	13
Aa Model	Text	40
Aa Type	Text	8
Aa Origin	Text	6
Aa DriveTrain	Text	5

Execute Macro ADJLEN

- **Step III**:

%***adjlen***(dsn=cars, exc=%str("DriveTrain"));

- Excluding DriveTrain variable from cars dataset.

"DriveTrain" Variable length in the cars dataset does not change.

After Execution of ADJLEN

- **STEP IV** : The length of each character variable has been reassigned with its maximum value length Except DriveTrain

Column Name	Type	Length
Aa make	Text	13
Aa model	Text	40
Aa type	Text	6
Aa origin	Text	6
Aa DriveTrain	Text	60

Results Comparison

- **STEP V**:

Before ADJLEN Execution

Column Name	Type	Length
make	Text	30
model	Text	80
type	Text	60
origin	Text	60
DriveTrain	Text	60

After ADJLEN Execution

Column Name	Type	Length
make	Text	13
model	Text	40
type	Text	6
origin	Text	6
DriveTrain	Text	60

Decrease the File Size

- Take sashelp.ZIPCODE dataset
- Execute ADJLEN Macro : %adjlen(dsn=ZIPCODE, exc=%str(""));
- Before Execution:

Member Name	Size of File
ZIPCODE	34.7MB

- After Execution:

Member Name	Size of File
ZIPCODE	32.7MB

Conclusion

- Using ADJLEN macro, maximum character variable value length can be reassigned to the corresponding variables in a dataset.

- Significantly decrease the size of the submission dataset.

- Avoid P21 FDAC036 Error Message

thank you!

Contact Information

- Prabhakara Rao Burma
- E-mail: prsas85@gmail.com

- Venkata N Madhira
- E-mail: venkatanmadhira@gmail.com

- **References**:
- https://www.pharmasug.org/proceedings/2015/SS/PharmaSUG-2015-SS08-SAS.pdf
- https://www.lexjansen.com/pharmasug/2017/SS/PharmaSUG-2017-SS05.pdf
- https://www.pinnacle21.com/validation-rules/sdtm
- https://www.fda.gov/downloads/forindustry/datastandards/studydatastandards/ucm312964.pdf

Questions?