



Presenter

Jinquan Wu - Pfizer

Works as a Sr. Statistical Programmer at Pfizer and supports Inflammation and Immunology Therapeutic Area.



SAS Macro for Converting Variable Attributes from XLSX Specifications to SDTM/ADAM SAS datasets

Introduction

- SDTM/ADaM are created strictly based on the specs.
- Specs usually come in multi-tabbed excel worksheets with XLSX ext.
- Must move the attributes from the spec to SAS program .
- Manually copying and pasting is **tedious, time-consuming and error-prone.**



Objective

automatically converts variable attributes from
XLSX specs to SAS datasets

Macro: util_spec_attrib.sas

```
util_spec_attrib(specpath      =      ,  
                  filename     =      ,  
                  worksheet    =      ,  
                  range        =      ,  
                  adsl_core_var =      ,  
                  adslpath     =      ,  
                  indsn        =      ,  
                  outdsn       =      );
```

SDTM/ADAM Specs Suitable to Read Attributes from

Snapshot 1. SDTM Specification Sample 1

Dataset	Variable	Label	Data Type	Length	Significant Digits	Format	Codelist	Origin
DM	STUDYID	Study Identifier	text	200				Protocol
DM	DOMAIN	Domain Abbreviation	text	2			DOMAIN	Assigned
DM	USUBJID	Unique Subject Identifier	text	60				Derived
DM	SUBJID	Subject Identifier for the Study	text	20				CRF
DM	RFSTDTC	Subject Reference Start Date/Time	text	19			ISO 8601	Derived
DM	RFENDTC	Subject Reference End Date/Time	text	19			ISO 8601	Derived
DM	RFXSTDTC	Date/Time of First Study Treatment	text	19			ISO 8601	Derived
DM	RFXENDTC	Date/Time of Last Study Treatment	text	19			ISO 8601	Derived

Study

Datasets

Variables

ValueLevel

WhereClauses

Methods

Comme ...

+

:

◀

▶

SDTM/ADAM Specs Suitable to Read Attributes from

Snapshot 2. SDTM Specification Sample 2

Name	Label	Sequenc	Length	CodeListRef	Type	Core	Origin	Inform Origin
STUDYID	Study Identifier	1	20		Character	Req	Protocol	
DOMAIN	Domain Abbreviation	2	2	DOMAIN	Character	Req	Assigned	
USUBJID	Unique Subject Identifier	3	60		Character	Req	Derived	
DSSEQ	Sequence Number	4	8		Number	Req	Derived	
DSREFID	Reference ID	5	20		Character	Perm	CRF	DSREFID
DSTERM	Reported Term for the Disposition Event	6	100	DSUNBL	Character	Req	CR, Assigned	DSTERM, DSDECOD
DSDECOD	Standardized Disposition Term	7	50	NCOMPLT	Character	Req	CR, Assigned	DSDECOD
DSCAT	Category for Disposition Event	8	20	DSCAT	Character	Exp	Assigned	
EPOCH	Epoch	9	50	EPOCH	Character	Exp	Derived	
DSDTC	Date/Time of Collection	10	19		Character	Perm	Assigned	
DSSTDTC	Start Date/Time of Disposition Event	11	19		Character	Exp	CRF	DSSTDAT_DTS
DSDY	Study Day of Visit/Collection/Exam	12	8		Number	Perm	Derived	
DSSTDY	Study Day of Start of Disposition Event	13	8		Number	Perm	Derived	
SUPPDS=DSPHASE	Disposition Phase	999	50	EPOCH	Character	Perm	CRF	EPOCH
SUPPDS=DSRCTYP	Re-consent Type(s)	999	25	DSRCTYP	Character	Exp	CRF	DSRCTYP
SUPPDS=DSCSOBT	Consent Was	999	12	DSOBT	Character	Exp	CRF	DSCSOBT

◀ ▶ ...

READ ME

Domain mapping

DS

Change Log

⊕

:

◀

—SDTM/ADAM Specs Suitable to Read Attributes from

Snapshot 3. SDTM Specification Sample 3

	A	B	C	D	E	F	G	H
1	T-DOMAIN		DS (One record per disposition status or protocol milestone per subject)					
2	T-DESCRIPTION		Disposition					
3	T-MODEL		CaPS SDTM 3.2					
4	T- Class:		EVENTS					
5	PDS Source:		FINAL,STC,ADVERSE, RANDOM					
6	Sort Order / Key:		STUDYID USUBJID DSDECODE DSSTDTC					
7								
8	Variable	Label	Type	Length	Core	CT or Format	Origin	Derivation/Comment
9	STUDYID	Study Identifier	Character	8	Req		Assigned	FINAL.PROTNO,STC.PROTNO,ADVERSE.PROTNO
10	DOMAIN	Domain Abbreviation	Character	2	Req	DOMAIN	Assigned	Default to "DS"
11	USUBJID	Unique Subject Identifier	Character	22	Req		Assigned	FINAL.PID,STC.PID,ADVERSE.PID, RANDOM.PID
12	DSSEQ	Sequence Number	Number	8	Req		Derived	Sequence Number given to ensure uniqueness of
13	DSREFID	Reference ID	Character	5	Perm		RANDOM.PTRAND	equal to RANDOM.PTRAND
14	DSTERM	Reported Term for the Disposition Event	Character	198	Req		FINAL, STC, ADVERSE, RANDOM, Assigned	a. If FINAL.WDRLPVR is not missing or FINSTAT= then set DSTERM = WDRLPVR concatenated with
15	DSDECODE	Standardized Disposition Term	Character	38	Perm	NCOMPLT.	RANDOM, Assigned	a. If FINAL.WDRLPVR is not missing or FINSTAT=
16	DSCAT	Category for Disposition Event	Character	18	Exp	DSCAT	Assigned	if FINAL.FINSTAT ne . or SBSMNA='NOT APPLIC then DSCAT = 'DISPOSITION EVENT'
17	EPOCH	Epoch	Character	9	Perm	EPOCH	Derived	DSDTC or DSSTDTC is greater than date portion
18	DSDTC	Date/Time of Collection	Character	10	Perm	ISO 8601	STC.COLLEDATE,	dsdtc equal to
19	DSSTDTC	Start Date/Time of Disposition Event	Character	10	Exp	ISO 8601.	FINAL.DIEDDATEF, FINAL.WDRLDAT,	a. If DSDECODE = 'DEATH' then DSSTDTC = FINAL.DIEDDATE
20	DSDY	Study Day of Visit/Collection/Exam	Number	8	Perm		Derived	when DSDTC is on or after DM.RFSTDTC, else DSDY = (DSDTC) - (date portion of
21	DSSTDY	Study Day of Start of Disposition Event	Number	8	Perm		Derived	DSSTDY= (DSSTDTC) - (date portion of DM.RFSTDTC + 1 when DSSTDTC on or after DM.RFSTDTC,

—SDTM/ADAM Specs Suitable to Read Attributes from

Snapshot 4. ADaM Specification Sample 1

Dataset	Variable	Label	Data Type	Length	Significant Digits	Format	Codelist	Origin
ADSL	STUDYID	Study Identifier	text	20				Predecessor
ADSL	USUBJID	Unique Subject Identifier	text	60				Predecessor
ADSL	SUBJID	Subject Identifier for the Study	text	10				Predecessor
ADSL	SITEID	Study Site Identifier	text	12				Predecessor
ADSL	RANDNO	Randomization Number	text	20				Derived
ADSL	AGE	Age	integer	8				Predecessor
ADSL	AGEU	Age Units	text	6			AGEU	Predecessor
ADSL	AAGE	Analysis Age	integer	8				Derived
ADSL	AAGEU	Analysis Age Unit	text	6			AGEU	Assigned
ADSL	SEX	Sex	text	1			SEX	Predecessor

◀ ▶ ...

Datasets

Variables

ValueLevel

WhereClauses

Codelists

Dictio ...

⊕

:

◀

▶

—SDTM/ADAM Specs Suitable to Read Attributes from

Snapshot 5. ADaM Specification Sample 2

Var in SDTM (DV or SUPPDV)	Type in SDTM	Core in SDTM	Label in SDTM	CT in SDTM	Var in ADDV	Type in ADDV	Core in ADDV	Label in ADDV	CT in ADAM
STUDYID	Char	Req	Study Identifier		STUDYID	Char	Req	Study Identifier	
USUBJID	Char	Req	Unique Subject		USUBJID	Char	Req	Unique Subject Identifier	
DOMAIN	Char	Req	Domain Abbreviation	DOMAIN	DOMAIN	Char	Req	Domain Abbreviation	
					TRTP	Char	Req	Planned Treatment	
					TRTA	Char	Cond	Actual Treatment	
					TRTPN	Num	Perm	Planned Treatment (N)	
					TRTAN	Num	Perm	Actual Treatment (N)	
					APERSDT	Num		Period Start Date	date3.
DVSEQ	Char	Req	Sequence Number		DVSEQ	Num	Req	Sequence Number	
					DVSPID	Char	Perm	Sponsor-Defined Identifier	
					DVTERM	Char	Req	Protocol Deviation Term	
					DVDECOD	Char	Perm	Protocol Deviation Coded Term	
					DVCAT	Char	Perm	Category for Protocol Deviation	
					DVSCAT	Char	Perm	Subcategory for Protocol Deviation	
					DVSTDTC	Char	Perm	Start Date/Time of Deviation	
					CAPE	Char	Perm	Analysis Population Exclusion	
					ASTDT	Num	Cond	Analysis Start Date	
					ASTDTM	Num	Cond	Analysis Start Date/Time	
					ASTDTF	Char	Cond	Analysis Start Date Imputation Flag	(DATEFL)
					ASTTM	Num	Perm	Analysis Start Time	TIME8.
					ASTTMF	Char	Cond	Analysis Start Time Imputation Flag	
					EPOCH	Char		Epoch	
					ACTSITE	Char		Actual Site of Deviation Occurrence	
					DESGTOR	Char	Perm	Visit Designator	

◀ ▶
ADDV Desc
Specifications
ADaM Definition
Data & Parameter Files
+
:
◀ ▶

Application

- 1) For most SDTM/ADAM Spec (eg. Snapshot 1, 2 and 4), only five parameters are needed to call this macro , which reads entire worksheet and gets column name from first row, eg.

```
%util_spec_attrb( specpath= C:\Bxxx1015\spec,  
                  filename= Bxxx1015_dm_mapping_spec.xlsx,  
                  worksheet=DM,  
                  indsn=dm_final,  
                  outdsn=DM);
```

Application

2) Sometime, spec (eg. Snapshot 3) may contain many more columns or rows than what we need. Use **range** to select columns and rows, eg.

```
%util_spec_attrib(specpath= C:\Axxx1056\spec,  
                  filename=Axxx1056_pds_ds.xlsx,  
                  worksheet=DS,  
                  range=A8:H21,  
                  indsn=DS8,  
                  outdsn=DS);
```

Application

3) For ADAM (eg. Snapshot 5) that need to copy core variables directly from ADSL, two more parameters needed, eg.

```
%util_spec_attrib(specpath= C:\Bxxx1015\spec,  
                  filename=Bxxx1015_addv_spec.xlsx,  
                  worksheet= Specifications,  
                  adsl_core_var= subjid siteid aage aageu sex race ...,  
                  adslpath= C:\Bxxx1015\adam,  
                  indsn=addv4,  
                  outdsn=addv);
```

under Pfizer system, can skip adslpath because it defaulted to DATVPROT



Conclusions

- Automatically convert var attributes from specs to SAS datasets.
- Improves programming efficiency and accuracy.
- Portable with minimal programmer intervention.
- Helpful for different levels of SAS programmers.

Reference

- DelGobbo, Vince. 2019. “Integrating SAS® and Microsoft Excel: Exploring the Many Available to You”. Proceedings of PharmaSUG 2019. Philadelphia, PA. Available at <https://www.lexjansen.com/pharmasug/2019/HT/PharmaSUG-2019-HT-067.pdf>
- Zhong, Wayne. 2012. “Efficiently Trim Character Variable Lengths to Fit Data, Reduce Dataset Size”. Proceedings of PharmaSUG 2012. San Francisco, CA. Available at <https://www.pharmasug.org/proceedings/2012/CC/PharmaSUG-2012-CC17.pdf>



Acknowledgement

Thank Gabriela Piasecki, Gori Maya Gurung and Ajay Krishna Shrestha for testing this macro.

Thank LiFeng Good for reviewing the paper and providing additional insight.



Contact Information

Jinquan Wu

Pfizer, Inc.

500 Arcola Rd, Collegeville, PA 19426

jinquan.wu6@gmail.com

<https://www.pfizer.com/>



Questions?