

A SAS macro to check SDTM domains against controlled terminology

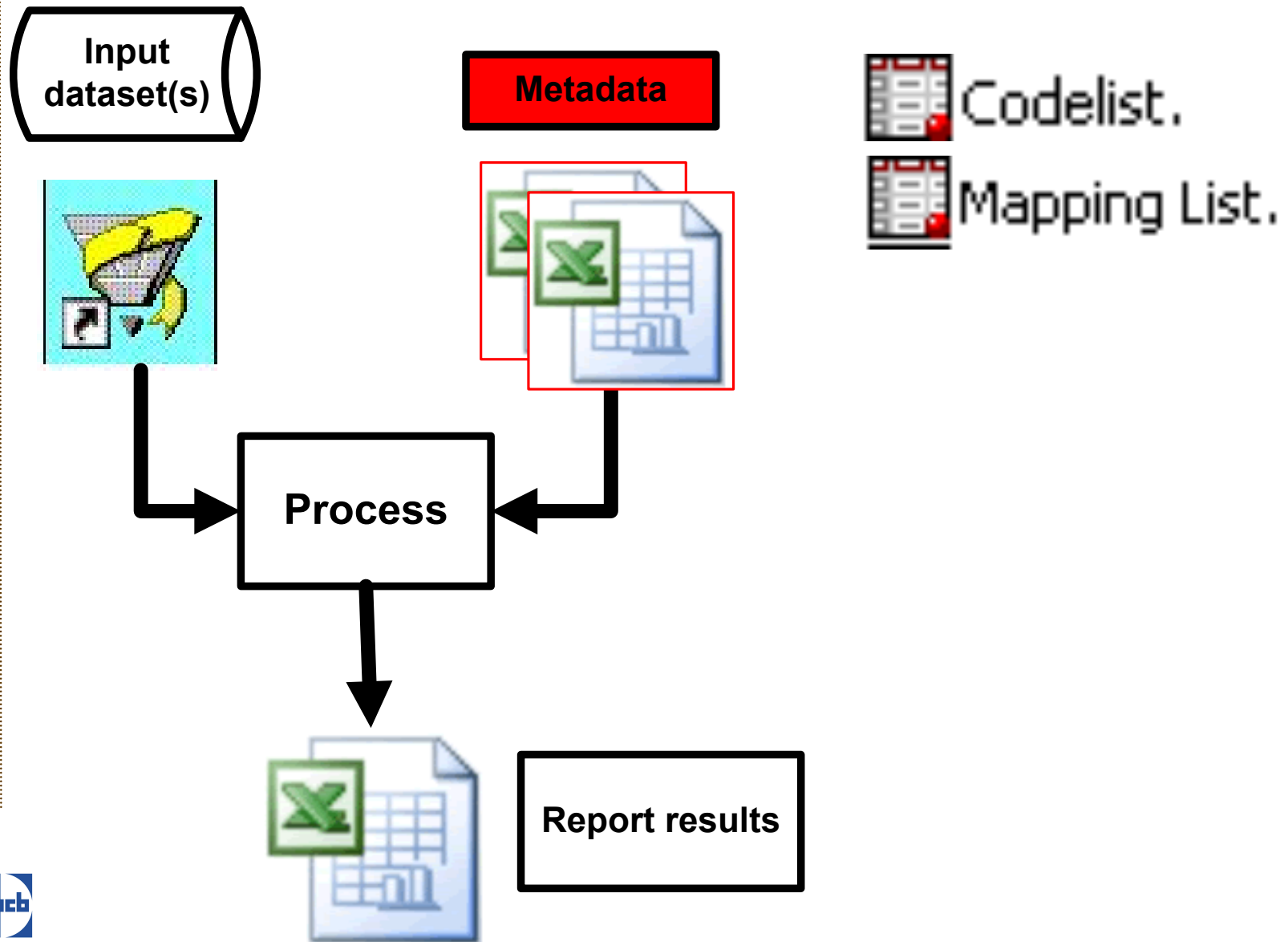


Christer, living with Parkinson's disease

When do we need this?

- During SAS development of SDTM domains
- Checking overall compliance to Controlled terminology (CT) for a study.
- Checking CT of one domain across multiple studies e.g. for pooling.

Process Flow: The metadata



Controlled terminology (CDISC)

Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	CDISC Submission Value	CDISC Synonym(s)	C
C66767		No	Action Taken with Study Treatment	ACN	Action Taken with Study Treatment	A
C49503	C66767		Action Taken with Study Treatment	DOSE INCREASED		A
C49504	C66767		Action Taken with Study Treatment	DOSE NOT CHANGED		A
C49505	C66767		Action Taken with Study Treatment	DOSE REDUCED		A
C49501	C66767		Action Taken with Study Treatment	DRUG INTERRUPTED		A
C49502	C66767		Action Taken with Study Treatment	DRUG WITHDRAWN		A
C48660	C66767		Action Taken with Study Treatment	NOT APPLICABLE	NA	D
C17998	C66767		Action Taken with Study Treatment	UNKNOWN	U; Unknown	N
C66768		No	Outcome of Event	OUT	Outcome of Event	A
C48275	C66768		Outcome of Event	FATAL	Grade 5; 5	TI
C49494	C66768		Outcome of Event	NOT RECOVERED/NOT		O

	NAME	EXTENSIB	Codelist Name	CDISC Submission Value	CDIS
1	ACN	No	Action Taken with Study Treatment	DOSE INCREASED	
2	ACN	No	Action Taken with Study Treatment	DOSE NOT CHANGED	
3	ACN	No	Action Taken with Study Treatment	DOSE REDUCED	
4	ACN	No	Action Taken with Study Treatment	DRUG INTERRUPTED	
5	ACN	No	Action Taken with Study Treatment	DRUG WITHDRAWN	
6	ACN	No	Action Taken with Study Treatment	NOT APPLICABLE	NA
7	ACN	No	Action Taken with Study Treatment	UNKNOWN	U; Unknown
8	OUT	No	Outcome of Event	FATAL	Grade 5; 5
9	OUT	No	Outcome of Event	NOT RECOVERED/NOT RESOLVED	
10	OUT	No	Outcome of Event		
11	OUT	No	Outcome of Event		
12	OUT	No	Outcome of Event		
13	OUT	No	Outcome of Event	UNKNOWN	U; Unknown
14	AGESPAN	Yes	Age Span	ADOLESCENT (12-17 YEARS)	



Codelist.

CDISC SDTM 3.1.2

Validation Rules for controlled terminology (CT)

CDISC SDTM 3.1.2 Validation Rules

The following is a listing of CDISC SDTM version 3.1.2 validation rules supported by OpenCDISC Validator. The list includes all Janus validation rules as defined in the [Janus SDTM Validation Specification v.1](#), CDISC controlled terminology rules, and additional rules defined by the OpenCDISC community.

Domain

Category

<Any>

Type

<Any>

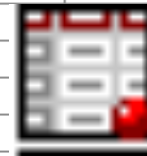
WebSDM ID

APPLY

Rule ID	Message	Description of Rule	Domain	Category	Type	WebSDM ID
CT0001	Value for --ACN not found in ACN controlled terminology codelist	Variable values should be populated with terms found in 'Action Taken with Study Treatment' (C66767) CDISC controlled terminology codelist	Events	Terminology	Warning	R5108
CT0002	Value for AESEV not found in AESEV controlled terminology codelist	Variable values should be populated with terms found in 'Severity/Intensity Scale for Adverse Events' (C66769) CDISC controlled terminology codelist	AE	Terminology	Warning	R5110

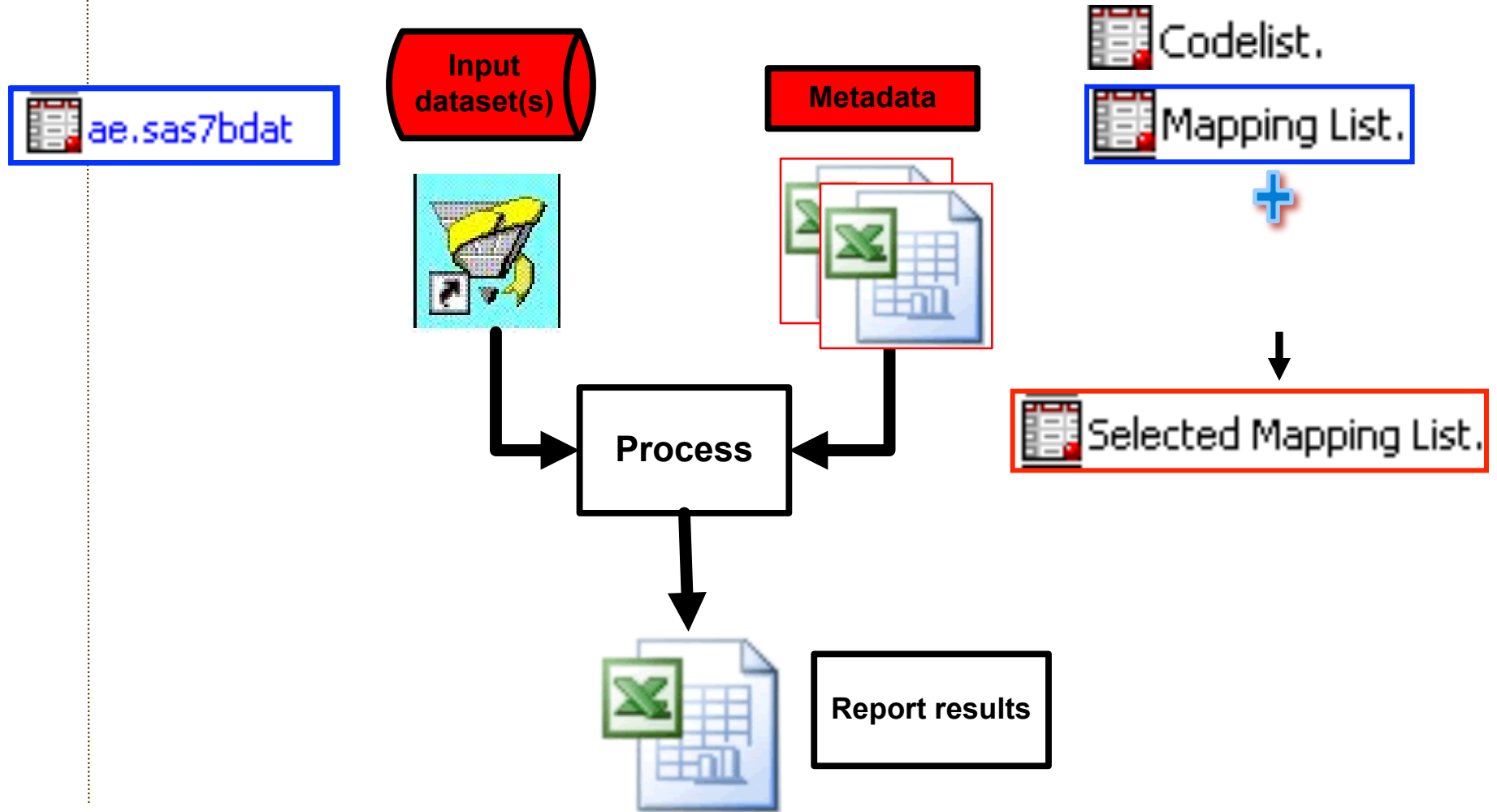
■ ■ ■

	RULEID	VARIABLE	CODELIST	EXTENSIB	WHERE	DOMAIN	TYPE	SEVERITY	WEBSID
1	CT0001	--ACN	ACN	N		Events	Warning	Medium	R5108
2	CT0002	AESEV	AESEV	N		AE	Warning	Medium	R5110
3	CT0003	TSVAL	AGESPAN	N	TSPARMCD= 'AGESPAN'	TS	Warning	Medium	
4	CT0004	AGEU	AGEU	N		DM	Error	High	R5062
5	CT0005	TSVAL	AGEU	N	TSPARMCD= 'AGEU'	TS	Warning	Medium	
6	CT0006	COUNTRY	COUNTRY	N		DM	Warning	Medium	R5008
7	CT0007	DATEST	DATEST	N		DA	Warning	Medium	
8	CT0008	DATESTCD	DATESTCD	N		DA	Warning	Medium	
9	CT0009	DOMAIN	DOMAIN	Y					
10	CT0010	DSCAT	DSCAT	Y					R5122
11	CT0011	EGMETHOD	EGMETHOD	Y					R5125
12	CT0012	EGSTRESC	EGSTRESC	Y					R5129
13	CT0013	EGTEST	EGTEST	Y		EG	Warning	Medium	R5123
14	CT0014	EGTESTCD	EGTESTCD	Y		EG	Warning	Medium	R5124
15	CT0015	ETHNIC	ETHNIC	Y		DM	Warning	Medium	R5113
16	CT0016	--EVAL	EVAL	Y		Findings	Warning	Medium	
17	CT0017	QEVAl	EVAL	Y		SUPPQUAL	Warning	Medium	



Mapping List.

Derive variables that have to comply to CT

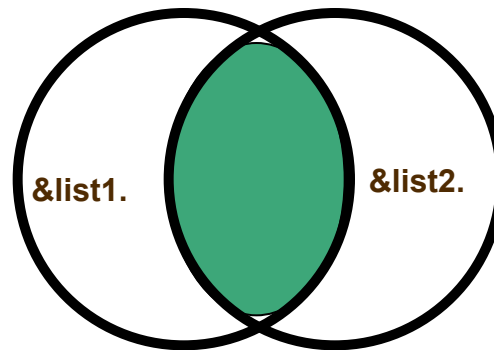


Intersection of two lists of values



ae.sas7bdat

#	Variable	Type	Len
1	STUDYID	Char	200
2	DOMAIN	Char	200
3	USUBJID	Char	200
4	AESEQ	Num	8
5	AESPID	Char	200
6	AETERM	Char	200
7	AEMODIFY	Char	200
8	AEDECOD	Char	200
9	AEBODSYS	Char	200
10	AESEV	Char	200
11	AESER	Char	200
12	AEACN	Char	200
13	AEACNOTH	Char	200
14	AEREL	Char	200
15	AEPATT	Char	200
16	AEOUT	Char	200
17	AESTDTC	Char	200
18	AEENDTC	Char	200



Mapping List.

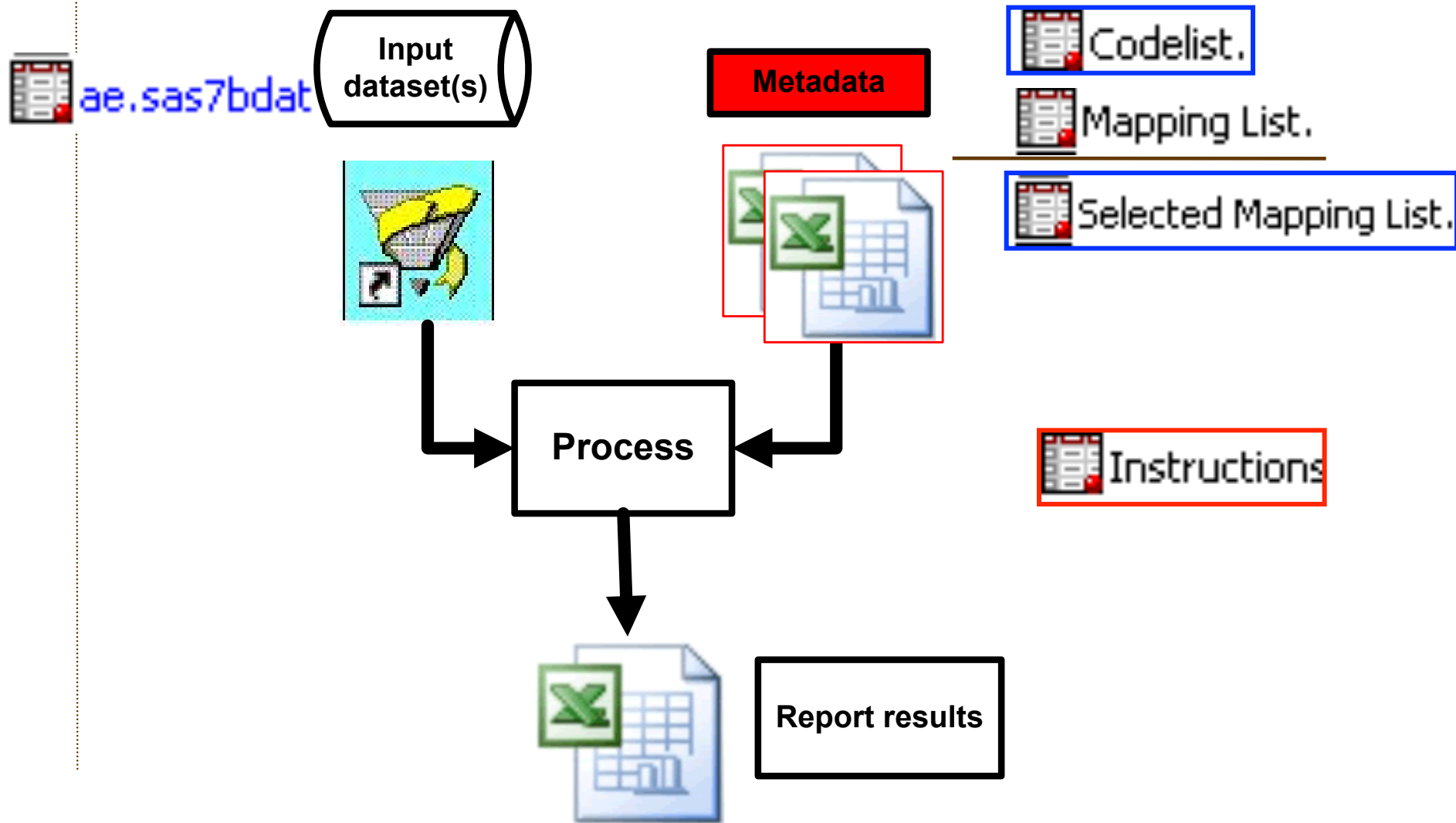
	NR	VARIABLE	CODELIST	EXTENSIB	RULEID	WHERE
1	1	AEACN	ACN	N	CT0001	
2	2	AEBODSYS	SOC	N	CT0037	
3	3	AEENRF	STENRF	N	CT0039	
4	4	AEOUT	OUT	N	CT0027	
5	5	AESCONG	NY	N	CT0065	
6	6	AESDISAB	NY	N	CT0066	
7	7	AESDTH	NY	N	CT0067	
8	8	AESER	NY	N	CT0064	
9	9	AESEV	AESEV	N	CT0002	
10	10	AESHOSP	NY	N	CT0068	
11	11	AESLIFE	NY	N	CT0069	
12	12	AESMIE	NY	N	CT0072	



Selected Mapping List.

	RuleID	Variable	Codelist	where
1	CT0001	AEACN	ACN	
2	CT0002	AESEV	AESEV	
3	CT0009	DOMAIN	DOMAIN	
4	CT0027	AEOUT	OUT	
5	CT0037	AEBODSYS	SOC	
6	CT0064	AESER	NY	

Join metadata information (and create instructions)



Build values list



Codelist.

RuleID	Variable	Codelist	where			
			NAME	EXTENSIB	LONGNAME	
1	AEACN	ACN	1	ACN	No	Action Taken with Study Treatment
2	AESEV	AESEV	2	ACN	No	Action Taken with Study Treatment
3	DOMAIN	DOMAIN	3	ACN	No	Action Taken with Study Treatment
4	AEOUT	OUT				
5	AEBODSYS	SOC				
6	AESEV	NY				



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



Codelist.



C

Creating a list iteratively



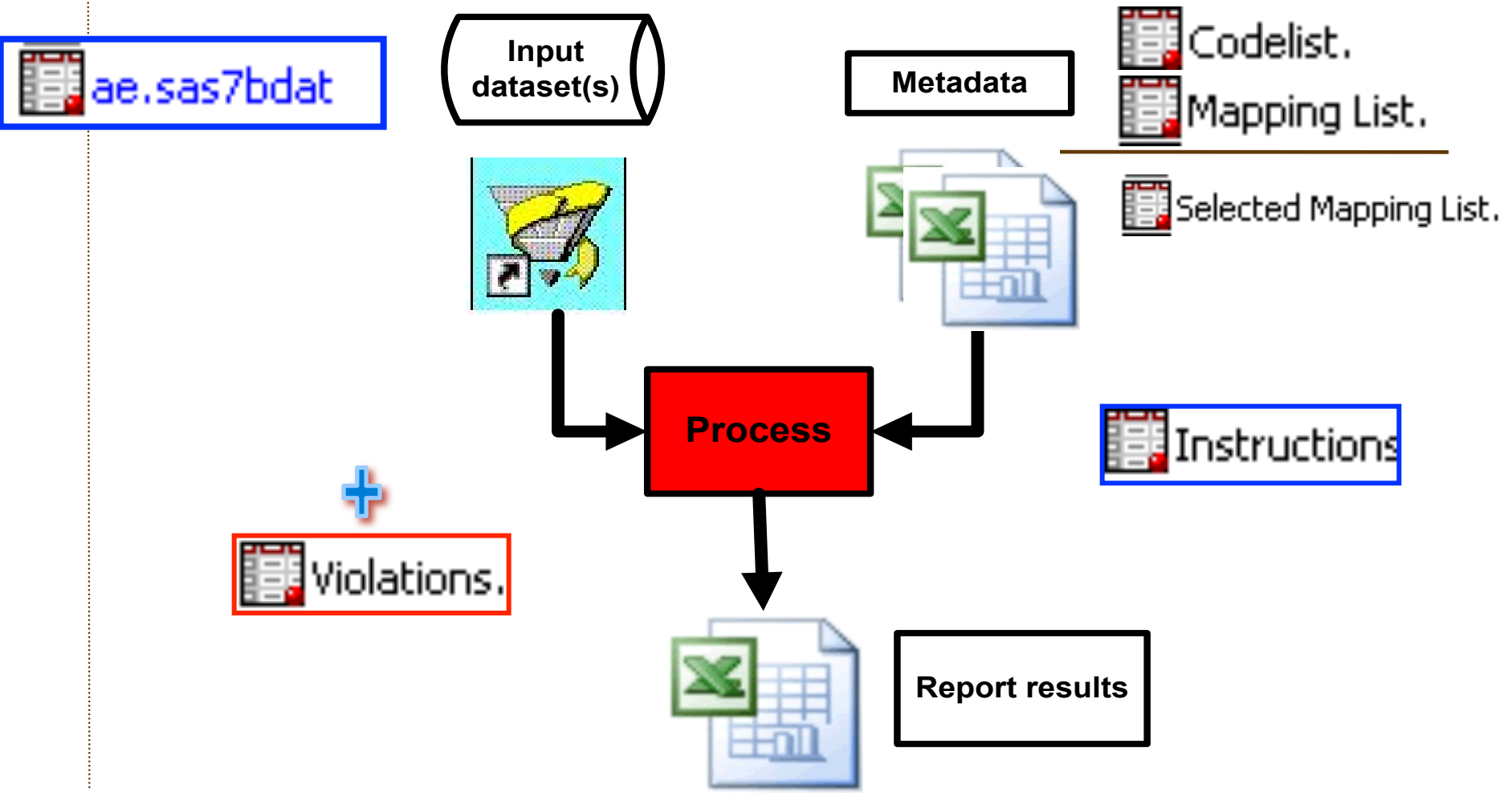
Codelist.

```
data ctlist2;
  retain no values ;
  length values $32767;
  set Codelist.
  by name;
  if first.name then values="";
  values=catx( ' ' ',strip(values),strip(list));
  if last.name then do;
    values=' ' ||strip(values)|| ' ' || ' ' ;
    no+1;
    output;
  end;
  drop list;
run;
```

NAME	EXTENSIB	LONGNAME	LIST
ACN	No	Action Taken with Study Treatment	DOSE INCREASED
ACN	No	Action Taken with Study Treatment	DOSE NOT CHANGED
ACN	No	Action Taken with Study Treatment	DOSE REDUCED
ACN	No	Action Taken with Study Treatment	DRUG INTERRUPTED
ACN	No	Action Taken with Study Treatment	DRUG WITHDRAWN
ACN	No	Action Taken with Study Treatment	NOT APPLICABLE
ACN	No	Action Taken with Study Treatment	UNKNOWN



NAME	EXTENSIB	Codelist Name	VALUES
ACN	No	Action Taken with Study Treatment	"DOSE INCREASED" "DOSE NOT CHANGED" "DOSE REDUCED" "DRUG INTERRUPTED" "DRUG WITHDRAWN" "NOT APPLICABLE" "UNKNOWN"



Creating the violations dataset

USUBJID	AES	AEACN	AEOUT	AEREL	AESEV
RA0074-002-02101	1	NOT APPLICABLE	RECOVERED/RESOLVED	NOT RELATED	MILD
RA0074-002-02103	1	NOT APPLICABLE	RECOVERED/RESOLVED	NOT RELATED	MILD
RA0074-002-02104	1	NOT APPLICABLE	RECOVERED/RESOLVED	NOT RELATED	MILD
RA0074-002-02107	1	NOT APPLICABLE	RECOVERED/RESOLVED	NOT RELATED	MILD
RA0074-002-02108	1	NOT APPLICABLE	UNKNOWN	NOT RELATED	MILD
RA0074-002-02109	1	NOT APPLICABLE	UNKNOWN	NOT RELATED	MILD
RA0074-002-02112	1	NOT APPLICABLE	RECOVERED/RESOLVED	RELATED	MILD
RA0074-002-02118	1	NOT APPLICABLE	RECOVERED/RESOLVED	RELATED	MILD
RA0074-002-02122	1	NOT APPLICABLE	RECOVERED/RESOLVED	NOT RELATED	MILD
RA0074-002-02124	1	NOT APPLICABLE	RECOVERED/RESOLVED	NOT RELATED	MILD
RA0074-002-02127	1	NOT APPLICABLE	RECOVERED/RESOLVED	RELATED	MILD
RA0074-002-02128	1	NOT APPLICABLE	RECOVERED/RESOLVED	RELATED	MILD
RA0074-002-02130	1	NOT APPLICABLE	RECOVERED/RESOLVED	RELATED	MILD
RA0074-002-02137	1	NOT APPLICABLE	RECOVERED/RESOLVED	RELATED	MILD



ae.sas7bdat



Violations.

NO	RuleID	Variable	Codelist	VALUES	where
1	CT0001	AEACN	ACN	"DOSE REDUCED""DOSE NOT CHANGED""DOSE INCREASED""DOSE INTERRUPTED""DRUG WITHDRAWN""NOT APPLICABLE""UNKNOWN""DOSE INCREASED"""	
2	CT0037	AEBODSYS	SOC	"IMMUNE SYSTEM DISORDERS""EAR AND LABYRINTH DISORDERS""HEPATOBIILIARY DISORDERS""NEOPLASMS BENIGN, MALIGNANT AND UNSPECIFIED (INCL CYSTS AND POLYPS)" "CARDIAC DISORDERS""REPRODUCTIVE SYSTEM AND BREAST DISORDERS""INFECTIONS AND INFESTATIONS""PREGNANCY, PUERPERIUM AND PERINATAL CONDITIONS" "MUSCULOSKELETAL AND CONNECTIVE TISSUE DISORDERS""ENDOCRINE DISORDERS" "METABOLISM AND NUTRITION DISORDERS""GASTROINTESTINAL DISORDERS""RESPIRATORY, THORACIC AND MEDIASTINAL DISORDERS""SURGICAL AND MEDICAL PROCEDURES" "INVESTIGATIONS""PSYCHIATRIC DISORDERS""GENERAL DISORDERS AND ADMINISTRATION SITE CONDITIONS""SKIN AND SUBCUTANEOUS TISSUE DISORDERS""BLOOD AND LYMPHATIC SYSTEM DISORDERS""CONGENITAL, FAMILIAL AND GENETIC DISORDERS""VASCULAR DISORDERS""NERVOUS SYSTEM DISORDERS""SCALD, BURN, Frostbite, and Thermal Injury""URINARY DISORDERS""INJURY, POISONING AND FATAL DISORDERS"""	
3	CT0027	AEOUT	OUT	"UNKNOWN""RECOVERING/RESOLVING""RECOVERED/RESOLVED""FATAL""NOT RECOVERED"""	
4	CT0064	AESER	NY	"Y""NA""U""N"""	
5	CT0002	AESEV	AESEV	"MILD""SEVERE""MODERATE"""	
6	CT0009	DOMAIN	DOMAIN	"DA""AD""EE""AE""TI""SU""PE""SG""BR""PC""TV""LB""HU""BM""EX""EG""MH""MB""TS""SL""SK""TE""SC""TA""DV""CM""QS""SV""AU""PG""MS""PP""IE""FH""VS""DS""ST""OM""DC""DM""CE""ML""FA""SE""CO"""	



Instructions

```

if aeacn not in
  ("DOSE INCREASED","DOSE NOT CHANGED","DOSE REDUCED","DRUG INTERRUPTED","DRUG WITHDRAWN","NOT APPLICABLE","UNKNOWN","")
then do;
  ...
output;
end;

```

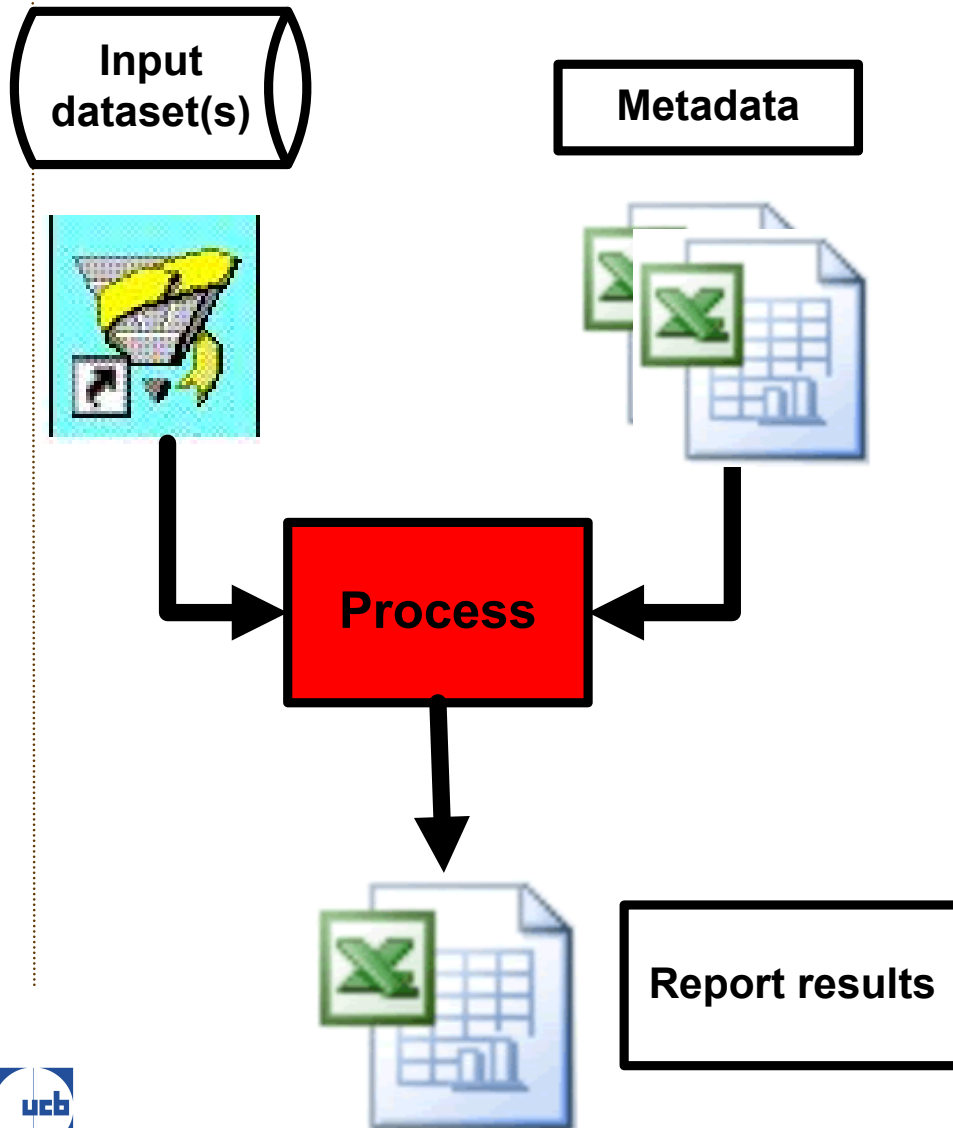
Violations dataset: One OBS per violation

```
data  Violations.(keep=keyinfo source domain ruleid variable value);
length source variable domain $8 ruleid $40 value $200 keyinfo $1000;
set ae ;
```

NO	RuleID	Variable	Code	VALUES	where
1	CT0001	AEACN	ACN	"DOSE REDUCED" "DOSE NOT CHANGED" "DRUG INTERRUPTED" "DRUG WITHDRAWN" "NOT APPLICABLE" "UNKNOWN" "DOSE INCREASED" ""	
2	CT0037	AEBDDSYS	SOC	"IMMUNE SYSTEM DISORDERS" "EAR AND LABYRINTH DISORDERS" "HEPATOBIILIARY DISORDERS" "NEOPLASMS BENIGN, MALIGNANT AND UNSPECIFIED (INCL CYSTS AND POLYPS)" "CARDIAC DISORDERS" "REPRODUCTIVE SYSTEM AND BREAST DISORDERS" "INFECTIONS AND INFESTATIONS" "PREGNANCY, PUERPERIUM AND PERINATAL CONDITIONS" "MUSCULOSKELETAL AND CONNECTIVE TISSUE DISORDERS" "ENDOCRINE DISORDERS" "METABOLISM AND NUTRITION DISORDERS" "GASTROINTESTINAL DISORDERS" "RESPIRATORY, THORACIC AND MEDIASTINAL DISORDERS" "SURGICAL AND MEDICAL PROCEDURES" "INVESTIGATIONS" "PSYCHIATRIC DISORDERS" "GENERAL DISORDERS AND ADMINISTRATION SITE CONDITIONS" "SKIN AND SUBCUTANEOUS TISSUE DISORDERS" "BLOOD AND LYMPHATIC SYSTEM DISORDERS" "CONGENITAL, FAMILIAL AND GENETIC DISORDERS" "VASCULAR DISORDERS" "NERVOUS SYSTEM DISORDERS" "SOCIAL CIRCUMSTANCES" "RENAL AND URINARY DISORDERS" "INJURY, POISONING AND PROCEDURAL COMPLICATIONS" "EYE DISORDERS" ""	
3	CT0027	AEOUT	OUT	"UNKNOWN" "RECOVERED" "RECOVERED/RESOLVED" "FATAL" "L" "N" ""	AE
4	CT0064	AESER	NY	"MILD" "SEVERE" ""	
5	CT0002	AEESEV	AEESEV	"DA" "AD" "IE" "AE" "AC" "TA" "DV" "CM" "DS" "SV" "AU" "FG" "MS" "PP" "TE" "FH" "VS" "DS" "ST" "OM" "DC" "DM" "CE" "ML" "FA" "SE" "CD" ""	

rw;

Extension: Multiple studies and multiple domains



Extension: Multiple

- domains/datasets
- Studies (libnames)

LIBNAMES = Index_J

List of Libnames

DSETLIST = Index_I

List of domains/datasets

Extentsion: Multiple domains/datasets

&DSETLIST.= ae eg vs ...

```
%let _dscount=1;  
%do %until (%qscan(%quote(&DSETLIST.),&_dscount.) = %str( ) );  
  %let _dset&_dscount.= %upcase(%qscan(%quote(&dsetlist.),&_dscount.));  
  %let _dscount=%eval(&_dscount.+1);  
%end;
```

e.g.:

&_dset1. = ae

&_dset2. = eg

&_dset3. = vs

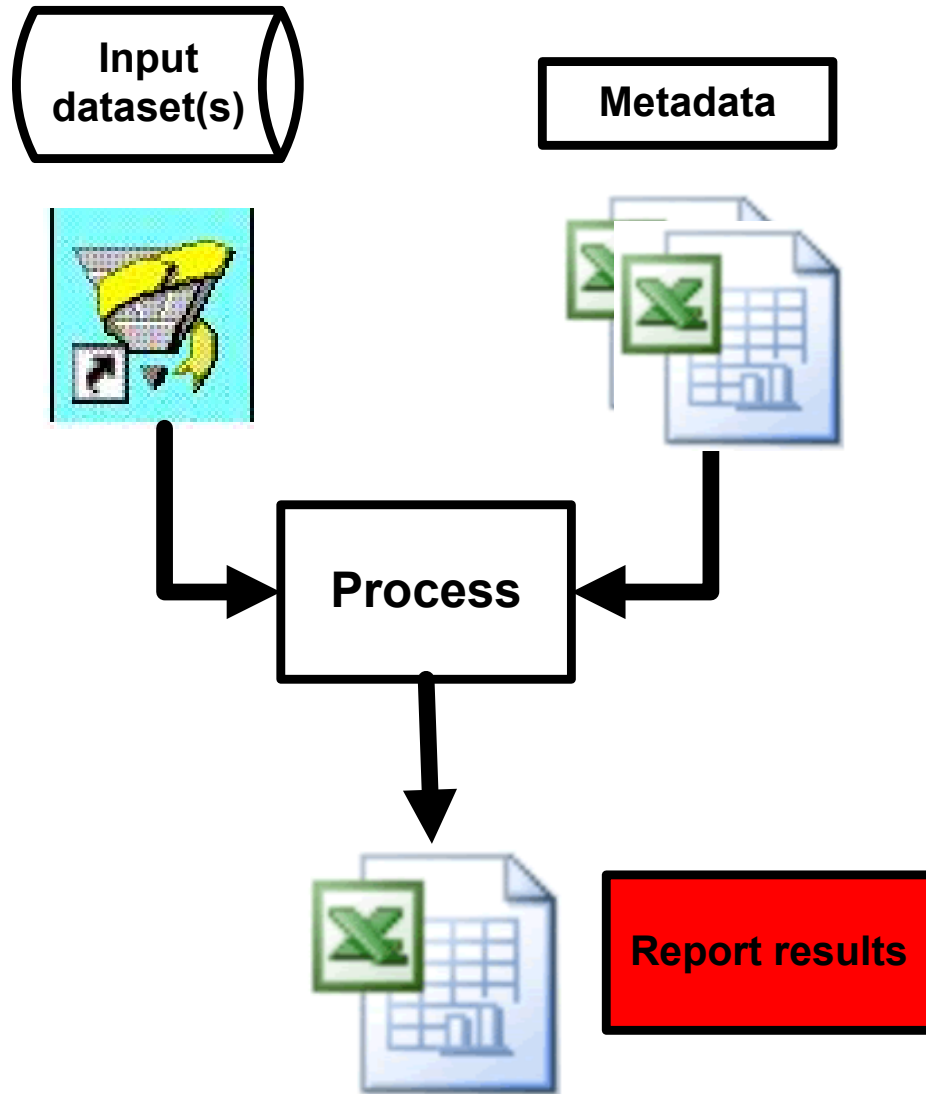
etc.

Extension: Multiple libnames

```
%do _i=1 %to &_dscount.-1;  
...  
%end;
```

```
proc sql noprint ;  
  select distinct(compress(upcase(name)||'|' $'||strip(pu||max(length),best.)))  
  into :_lengthv separated by ' '  
  from sashelp.vcolumn  
  where libname in (quoted list of libnames) and  
         upcase(memname)=upcase("&&_dset&_i..") and lowercase(type)="char"  
  group by name;  
quit;
```

Reporting and tabulation



Xml-Output (in Excel)

```
ods tagsets.ExcelXP file="&_file." path="&_outdir." style=UCB2;  
ods tagsets.ExcelXP options (frozen_headers="yes" frozen_rowheaders="2" row_repeat="1"  
    absolute_column_width="8,8,8,22,5,60 &c4." pages_fitwidth="1" pages_fitheight="200" orientation="landscape"  
    autofilter="Yes" sheet_name="Controlled term. Violations" fudge_factor="0.5" formulas="no"  
    autofit_height="Yes" center_vertical="Yes" center_horizontal="Yes" gridlines="Yes");  
proc report nowd data=chk_all1 ;  
    label  
        source ="Libname"  
        ...  
  
        ruleid  ="Rule ID (OPEN CDISC)" ;  
    columns source ... ruleid;  
run;  
  
ods tagsets.ExcelXP close;
```

http://support.sas.com/rnd/base/ods/odsmarkup/excelxp_help.html

XML Output

	A	B	C	D	E	F	G	H	I
	Libname	SAS Variable	Controlled term	Incorrect value	Extensible?	Full comment	Rule ID (OPEN CDISC)	Type	Severity
1									
2	C87088	AESEV	AESEV	UNKNOWN	N	Value UNKNOWN of variable AESEV not found in codelist AESEV. Key values: STUDYID= C87088, USUBJID= C87085-041-0282, AESEQ= 1	CT0002	Warning	Medium
3	C87088	AESEV	AESEV	UNKNOWN	N	Value UNKNOWN of variable AESEV not found in codelist AESEV. Key values: STUDYID= C87088, USUBJID= C87085-051-0353, AESEQ= 6	CT0002	Warning	Medium
4	C87088	AESEV	AESEV	UNKNOWN	N	Value UNKNOWN of variable AESEV not found in codelist AESEV. Key values: STUDYID= C87088, USUBJID= C87085-051-0435, AESEQ= 5	CT0002	Warning	Medium
5	C87088	AESEV	AESEV	UNKNOWN	N	Value UNKNOWN of variable AESEV not found in codelist AESEV. Key values: STUDYID= C87088, USUBJID= C87085-055-0356, AESEQ= 6	CT0002	Warning	Medium

Controlled term. Violations / Controlled term. Overview

	A	B	C	D	E	F
	Controlled term	Extensible	Long description	Possible values	Value synonyms	CT (non-grouped)
1						
2	AESEV	No	Severity/Intensity Scale for Adverse Events	MILD	Grade 1; 1	AESEV
3				MODERATE	Grade 2; 2	AESEV
4				SEVERE	Grade 3; 3	AESEV

Controlled term. Violations / Controlled term. Overview

Conclusion

- Can be used within SAS programs
- ADaM or SDTM
- Multiple studies can be checked simultaneously
- Users can control and modify input metadata (sponsor or study specific codelists)
- Scope is addressed via macro parameters
 - All domains of current study
 - Specific domains across various studies etc.

Questions?

