

Using the BCP for Creating SAS Datasets and Codebooks

International Blaise Users Conference, April 2015

Lilia Filippenko, Mai Nguyen, and Chris Carson



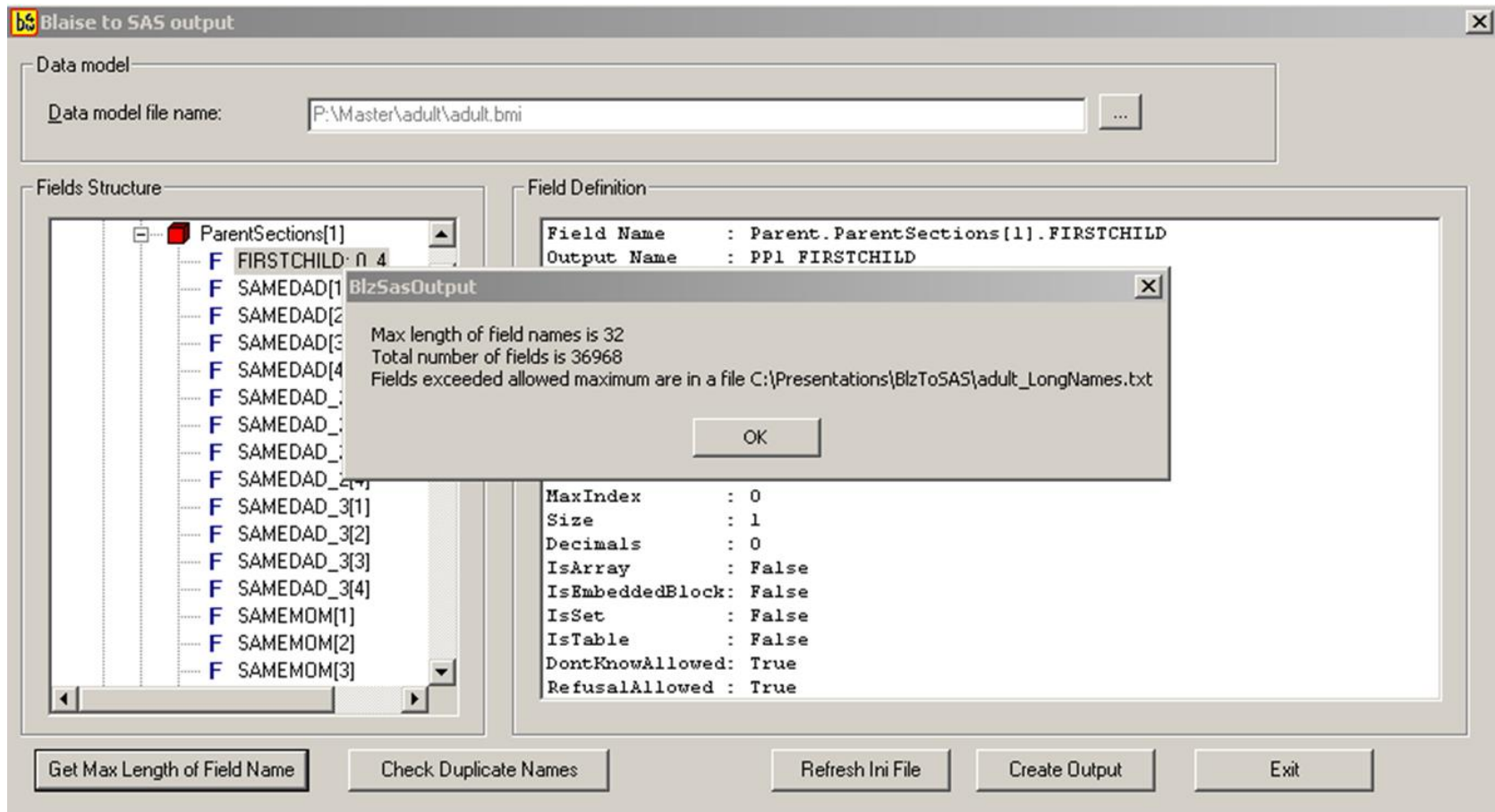
Outline

- Background
- Description of BlzToSAS application
 - Instrument Development
 - Blaise Instrument Requirements
 - Define Initialization File
 - Rename Variables
 - Define and Review Labels
 - Examine Blaise Instrument
 - Create SAS programs
 - Data Delivery
 - Final SAS Dataset
 - Codebook
- Conclusion



Background

- Cameleon with large Blaise Instruments
- BlzToSAS application uses Blaise Component Package (BCP)



Instrument Development – Blaise Instrument Requirements

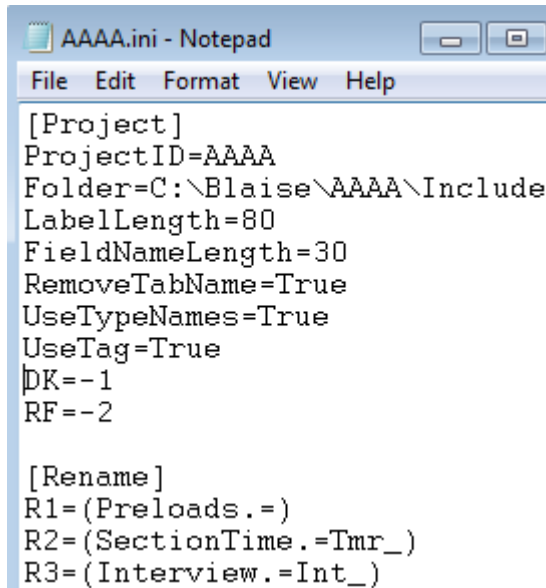


- Field Description
 - Label in a SAS dataset (“Label”)
 - Status of the field in delivery file – “NoDeliv”
- Field Tag
 - Variable name in a SAS dataset (“Analytic”)
 - An index is added to the variable name for arrays or set variables

```
QD08 (QDRELATED)
ENUS
"HAND R SHOWCARD QD08.@/@/How are you related to ^prld.c_fname?"
/"Adult relation to youth"
: TREL
```

Instrument Development - Initialization File

- Example of simple initialization file



```
AAAA.ini - Notepad
File Edit Format View Help
[Project]
ProjectID=AAAA
Folder=C:\Blaise\AAAA\Include
LabelLength=80
FieldNameLength=30
RemoveTabName=True
UseTypeNames=True
UseTag=True
DK=-1
RF=-2

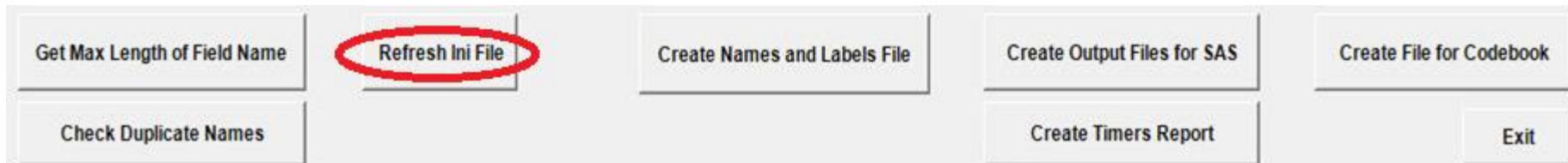
[Rename]
R1=(Preloads.=)
R2=(SectionTime.=Tmr_)
R3=(Interview.=Int_)
```

- Project ID to use custom code
- Folder name for output files
- Max size for default label
- Max size for variable name
- Use table name in a variable name
- Use Blaise defined type names
- Use a tag as an output field name
- Value for “Don’t know”
- Value for “Refusal”
- Initial rename statements

- Additional options:

- FileRenames - file name where rename statements are defined
- FileLabels – file name with defined labels to use in delivery file

Instrument Development - Rename Variables



- Field Tag – specified during instrument development

Fields Structure

- FR8: BFR8
 - FR8_1: TYesNo
 - FR8_2: TYesNo
 - FR8_3: TYesNo
 - FR8_4: TYesNo
 - FR8_5: TYesNo
 - FR8_5a: tblFR8_5a
 - FR8_5a[1]
 - FR8_5a[2]
 - FR8_5a: TYesNo
 - FR8_5a[3]

Field Definition

Field Name : FR.FR8.FR8_5a.FR8_5a[2].FR8_5a
Output Name : FR8_5a_2
Output Type : Element #2 of ARRAY[1..49] OF FR8_5a
Field type : blftEnumeration
DisplayDataType: TYesNo
ElementDataType: ENUM(2)
Local Name : FR8_5a
Field Label : You helped # 2
Field Tag : FR8_5a
BlockText :
MinValue : 1
MaxValue : 2

- Rename statements – wildcard instructions

```
AddHPS_renames.txt - Notepad
File Edit Format View Help
(FR6_9_13.FR6_9_13=CH)
(FR6_3sp.=)
(FR6_3_ERROR.FR6_3_ERROR=CHS)
(FR6_5sp.FR6_5sp=CHS)
(FR6_5.FR6_5=CH)
(FR6_6_8.FR6_6_8=CH)
(FR6_3.=)
(FR6_3_13.FR6_3_13=CH)
(FR6_9_13.FR6_9_13=CH)
(FR6_3sp.=)
(FR6_3_ERROR.FR6_3_ERROR=CHS)
(FR6_5sp.FR6_5sp=CHS)
```

Field Definition

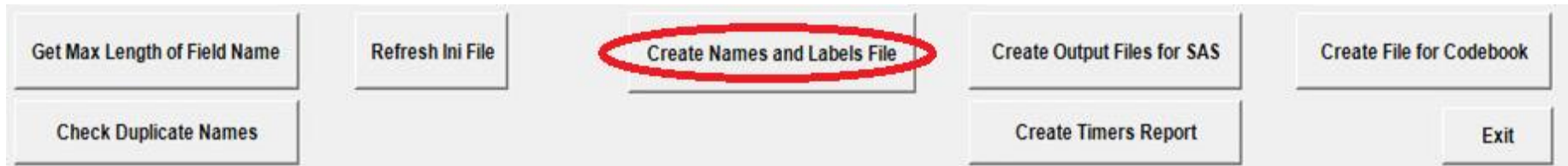
Field Name : FR.FR6.FR6_9_13.FR6_9_13[1].FR6_9
Output Name : CH_1_FR6_9
Output Type : Element #1 of ARRAY[1..49] OF FR6_9_13
Field type : blftEnumeration
DisplayDataType: TFR6_9s
ElementDataType: ENUM(3)
Local Name : FR6_9
Field Label : Relationship # 1
Field Tag :

Instrument Development - Examine Blaise Instrument

The screenshot displays the 'Blaise to SAS' application window. At the top, the 'Data model' section shows the file path: 'L:\RCD\Blaise\CAPIS\Sleep\Sleep.bmi'. Below this, the 'Fields Structure' pane on the left lists various fields, including 'PRC3a: TPRtmas' which is highlighted with a question mark. The 'Field Definition' pane on the right shows details for 'PRC.PRCLoop[1].PRC3a', including its output name 'PRC3a_1' and type 'blftEnumeration'. A central dialog box titled 'BlzToSAS' contains the following text: 'Max length of field names is 31', 'Total number of fields is 35413', and 'Incorrect names of fields or types are in a file C:\Blaise\BlzSASoutput\BlzToSAS\bin\NISVSS_IncorrectNames.txt'. The dialog has an 'OK' button. At the bottom of the application, several buttons are visible: 'Get Max Length of Field Name' (circled in red), 'Check Duplicate Names' (circled in red), 'Refresh Ini File' (circled in blue), 'Create Names and Labels File', 'Create Output Files for SAS', 'Create File for Codebook', 'Create Timers Report', and 'Exit'.

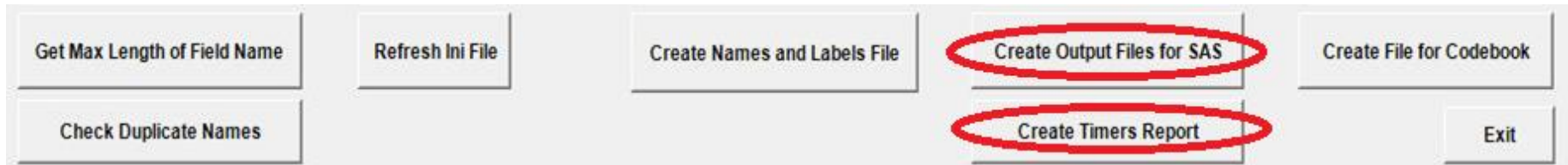
A few iterations of revisions to initialization file and the Blaise instrument might be needed to have variable and type names acceptable for SAS.

Instrument Development – Labels Definition



- Excel output file:
 - Module name
 - Blaise full name
 - Variable name as in specifications
 - SAS variable name
 - Default label (field description or part of a question text).
 - Question text in default language
- Excel input file:
 - SAS variable name
 - New label text

Instrument Development - Create SAS programs



- Manipula setup
- SAS program with include files:
 - SAS proc formats for enumerated types
 - Input SAS statements
 - SAS variable labels
 - SAS formats for all variables
 - Recodes for multiple choice variables
 - SAS “drop” statements
 - Statements to recode DK and RF values
- SAS program to produce timers report

Name	Size	Type
AddHPS.man	173 KB	Blaise Mani
AddHPS_dk_rf.inc	381 KB	INC File
AddHPS_drop.inc	1 KB	INC File
AddHPS_formats.inc	116 KB	INC File
AddHPS_input.inc	463 KB	INC File
AddHPS_labels.inc	396 KB	INC File
AddHPS_procfmt.inc	21 KB	INC File
AddHPS_toggles.inc	21 KB	INC File
AddHPS_Names.csv	977 KB	Microsoft Ex
AddHPS.sas	2 KB	SAS System
AddHPSTimersReport.sas	3 KB	SAS System

Instrument Development - Create SAS dataset (1)

- Manipula Setup

```
INPUTFILE InpFile: Meta1 ('AddHPS', BLAISE)
SETTINGS
  ACCESS = SHARED

OUTPUTFILE OutFile: Meta1 ('AddHPS.asc', ASCII)

MANIPULATE
  IF InpFile.Preloads.SPART_FIRST = DK THEN OutFile.Preloads.SPART_FIRST := '-1' ENDIF
  IF InpFile.Preloads.SPART_FIRST = RF THEN OutFile.Preloads.SPART_FIRST := '-2' ENDIF
  IF InpFile.Preloads.SPART_LAST = DK THEN OutFile.Preloads.SPART_LAST := '-1' ENDIF
  IF InpFile.Preloads.SPART_LAST = RF THEN OutFile.Preloads.SPART_LAST := '-2' ENDIF
```

- SAS driver program

```
AddHPS.sas *
libname library '\\...\BlzToSAS\Include';
libname out '\\...\BlzToSAS\Include';

TITLE 'AddHPS';
%include '\\...\BlzToSAS\Include\AddHPS_procfmt.inc';

DATA out.AddHPS;
INFILE '\\...\BlzToSAS\Include\AddHPS.asc' LRECL = 34179 TRUNCOVER;
%include '\\...\BlzToSAS\Include\AddHPS_input.inc';
%include '\\...\BlzToSAS\Include\AddHPS_toggles.inc';
%include '\\...\BlzToSAS\Include\AddHPS_labels.inc';
%include '\\...\BlzToSAS\Include\AddHPS_formats.inc';
%include '\\...\BlzToSAS\Include\AddHPS_dk_rf.inc';
%include '\\...\BlzToSAS\Include\AddHPS_drop.inc';

RUN;
```

Instrument Development - Create SAS dataset (2)

- Include file with formats for enumerated types:

```
PROC FORMAT library=library;  
  
VALUE TCont  
  1='Continue'  
 -1='Don''t know'  
 -2='Refusal'  
;  
  
VALUE TYNconst  
  1='Yes'  
  2='No'  
 -1='Don''t know'  
 -2='Refusal'  
;
```

- Include file with input statements

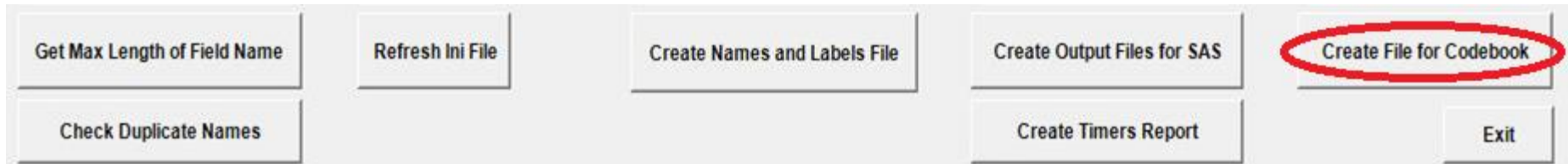
```
SV7_17_89          324279 - 324279 /* SV7_17[89] */  
SV7_17_90          324280 - 324280 /* SV7_17[90] */  
SV7_17_91          324281 - 324281 /* SV7_17[91] */  
SV7_17_92          324282 - 324282 /* SV7_17[92] */  
SV7_17_93          324283 - 324283 /* SV7_17[93] */  
SV7_17_94          324284 - 324284 /* SV7_17[94] */  
RAge               324285 - 324287 /* RAge */  
BreakS             324288 - 324288 /* BreakS */  
BreakSV            324289 - 324289 /* BreakSV */  
BreakPA            324290 - 324290 /* BreakPA */
```

```
;
```

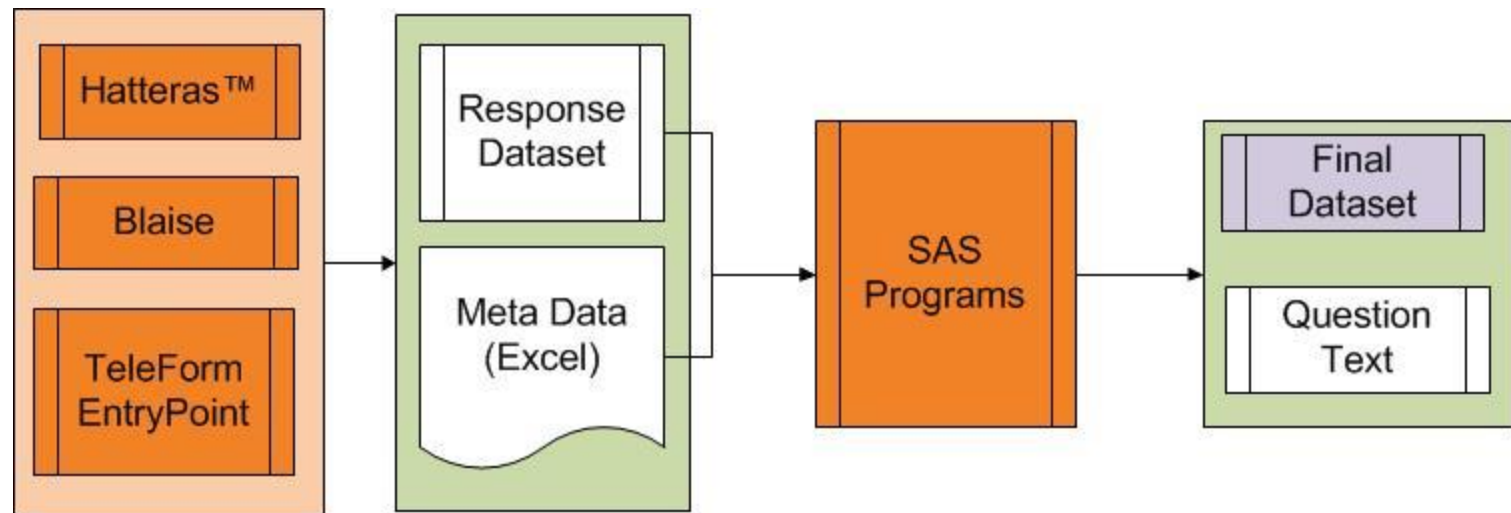
Ln 35415,

- Blaise instrument with 35,413 variables – 165,021 lines of code

Data Delivery – Final SAS Dataset

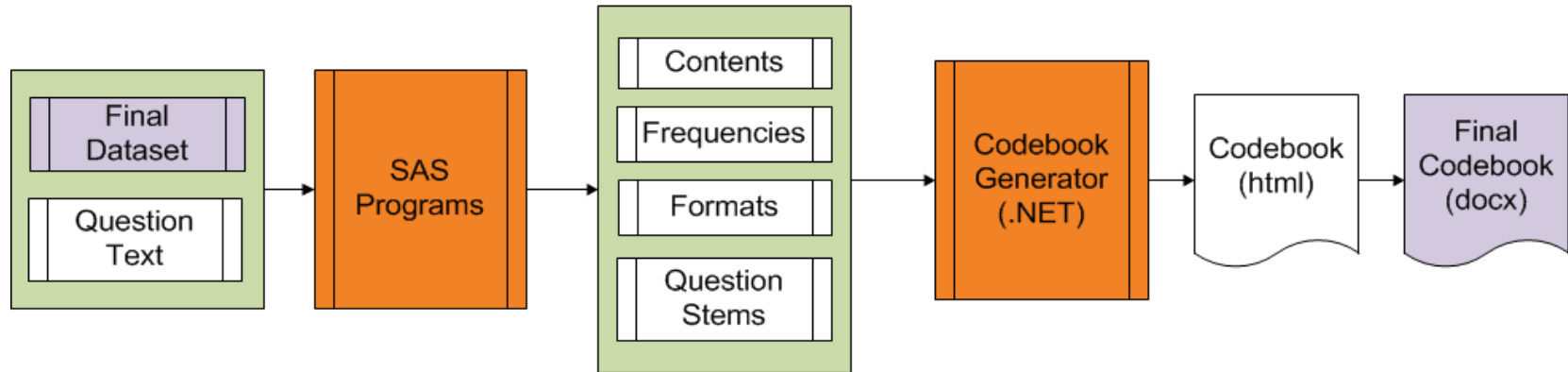


- Metadata file for codebook - Excel spreadsheet:
 - List of variables
 - Labels and question texts
 - Format names
- Generic SAS based Codebook Generation system is used to prepare final SAS dataset



Data Delivery - Codebook

- Program-generated codebook



- Final codebook

- Word/PDF document
- 2-column format
- Pagination
- Header/footer

Screener Codebook
September 9, 2013

Variable: Person_1_SCO_11					
Description: Is HH member 1 eligible					
Stem: HH MEMBER					
Notes:					
Label	Value	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Legitimate skip	.	3	1.52%	3	1.52%
YES	1	195	98.48%	198	100.00%
DELETE	4	0	0.00%	198	100.00%
Dontknow	-1	0	0.00%	198	100.00%
Refusal	-2	0	0.00%	198	100.00%

Variable: Person_2_ageR					
Description: Age of person 2					
Stem: How old is "firstname" ENTER 0 FOR LESS THAN 1 YEAR					
Notes:					
Label	Value	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Legitimate skip	.	58	29.29%	58	29.29%
Data present		137	69.19%	195	98.48%
Dontknow	-1	2	1.01%	197	99.49%
Refused	-2	1	0.51%	198	100.00%

Conclusion

- Projects producing extremely large data models or data models with complex nested structures and/or relational structures benefit from using BlzToSAS application for generating SAS datasets.
- RTI's adoption of this system has paid off in efficiency gains and improvements in the quality and reliability of file deliveries.
- BlzToSAS application has been well integrated into RTI's codebook generation system, so that a visually appealing codebook document can accompany our file deliveries.

More Information

Lilia Filippenko

lfilippenko@rti.org

Mai Nguyen

mnguyen@rti.org

Chris Carson

cpc@rti.org

