



Programming and Use of Blaise to SAS Transformation Tool in the PSID

April Beaulé, Mohammad Mushtaq
The 14th International Blaise Conference
London, UK
April 2012

About the PSID

- Nationally representative sample of 8,900 U.S. families
- Panel survey conducted since 1968
- Survey instrument programmed in Blaise and C#.net

Background

- Main survey instrument reprogrammed in Blaise in 2001
- Size and complexity of instrument required a new extraction tool
- DETL (Data Extraction Transformation & Loading)
- Takes advantage of all the metadata from the Blaise tables

Organizing Data for Editors

- Organize variables by person and by content area
- More questions asked about Head & Wife
- Question series are looped but need to subset for processing
- Basic extraction is at the block level but we need or organize data differently for editing

User Definitions

The screenshot displays the PSID Application Suite interface. The title bar reads "PSID Application Suite: - [Data: Extraction, Transformation and Loading]". The menu bar includes "Action", "Edit", "Query", "Block", "Record", "Field", "Window", and "Help". The toolbar contains various icons for file operations and navigation. The "Wave:" dropdown is set to "2011-PSID".

On the left side, there are two sections of radio buttons:

- IW Data**
 - Blaise Meta
 - SAS Setup
 - Table Meta
 - Setup Mgmt
 - Review IF
- Cross Wave**
 - Var Change
 - Copy Forw'd

The main area is titled "Setup" and contains a text box with "USER = ABEAULE". Below this is a table with columns: Domain, Name, Path, Notes, and a "Copy" button.

| Domain | Name | Path | Notes | Copy |
|----------|----------|--|---|------|
| ALLYEARS | BLZ_HOME | C:\Blaise\Blaise48\Bin\ | Blaise software home | Copy |
| ALLYEARS | SAS_HOME | C:\Program Files\SAS\SASFoundation\9.2\ | SAS software home (SAS.EXE) directory. | Copy |
| 2011PSID | SAS_DATA | X:\PROCESS\Data11\Edsys\BLmeta\ | Output (SAS) data files (.SAS7BDAT) location. | Copy |
| 2011PSID | SAS_FILE | X:\PROCESS\Data11\Edsys\BLmeta\pgmslogs\ | Program files (.SAS) location. | Copy |
| 2011PSID | SAS_LOG | X:\PROCESS\Data11\Edsys\BLmeta\pgmslogs\ | Log files (.LOG) location. | Copy |
| | | | | Copy |

Display Blaise Meta

Wave: 2011-PSID BMI: PSID11

Click on the output table SocSec - block defined by the Blaise programmer

IW Data

- Blaise Meta
- SAS Setup
- Table Meta
- Setup Mgmt
- Review IF

Crds Wave

- Var Change
- Copy Forw'd

| varId | Name | Label | Type | Width | Dec | Loops | Min | Max | DK | RF | Frame | Universe |
|-------|------------------|--|------|-------|-----|-------|-----|-----|----|----|-------|----------|
| 058 | BG7_G11 | Business Series, G7 - G11 | | 3 | | 5 | | | | | | |
| 059 | B5ocSec | Social Security Questions, G33 - G35 | | 3 | | 24 | | | | | | |
| 060 | BG73_G90 | G73 through G90 Question Series | | 3 | | 24 | | | | | | |
| 061 | BG77_G81 | Job Description Series | | 4 | | 4 | | | | | | |
| 062 | BG90d_G97 | | | 3 | | 24 | | | | | | |
| 063 | BG105 | List of People who received money | | 3 | | 24 | | | | | | |
| 064 | BSection_W | Section W | | 2 | | 1 | | | | | | |
| 065 | BW124_W128 | Inheritance Questions | | 3 | | 3 | | | | | | |
| 066 | BSection_P | Section P - Head's and Wife's Pensions | | 2 | | 1 | | | | | | |
| 067 | BSection_PDetail | Section P - Head's and Wife's Pensions | | 3 | | 2 | | | | | | |

Subset by:

varType <NULL>
Width <NULL>
nDec <NULL>

Show Vars

Raw Freqs

Parameter

- Application
- User Setup

Variables

| varId | Name | Label | Type | Width | Dec | Loops | Min | Max | DK | RF | Frame | Universe |
|-------|-----------------|-------|------|-------|-----|-------|-----|--------|----|----|-------|----------|
| 001 | BlaiseKey | | char | 7 | | 1 | | | N | N | | |
| 002 | B5ocSecInstance | | num | 5 | 0 | 1 | | | N | N | | |
| 003 | G33a_1 | | num | 1 | 0 | 6 | 1 | 7 | Y | Y | | |
| 004 | G33a_2 | | num | 1 | 0 | 6 | 1 | 7 | Y | Y | | |
| 005 | G33a_3 | | num | 1 | 0 | 6 | 1 | 7 | Y | Y | | |
| 006 | G33a_4 | | num | 1 | 0 | 6 | 1 | 7 | Y | Y | | |
| 007 | G33a_5 | | num | 1 | 0 | 6 | 1 | 7 | Y | Y | | |
| 008 | G33a_6 | | num | 1 | 0 | 6 | 1 | 7 | Y | Y | | |
| 009 | G33a5pec | | char | 100 | | 1 | | | Y | Y | | |
| 010 | G34 | | num | 6 | 0 | 1 | 1 | 999997 | Y | Y | | |
| 011 | G34Per | | num | 1 | 0 | 1 | 5 | 7 | Y | Y | | |
| 012 | G34PerSpec | | char | 100 | | 1 | | | Y | Y | | |
| 013 | G35_01 | | num | 2 | 0 | 13 | 1 | 13 | Y | Y | | |

G34 [qText] | Frames | Freqs | Universe

Blaise Variable: G34 Loops:

Question Text

[F1] - HELP How much was the total amount from Social Security? <xG34> w ENTER amount here, then ENTER unit of time on next screen (Month, Year)

Click to view question text

Meta information - question text appears here

Display Frequencies

The image shows the SPSS 'Variables' dialog box and a resulting frequency table. A callout box on the left provides helpful information about the variable table columns: width, loops, and DK/RF allowed.

Helpful information displayed including width, loops, whether DK/RF allowed

| varId | Name | Label | Type | Width | Dec | Loops | Min - Max | DK | RF |
|-------|------------|-------|------|-------|-----|-------|------------|----|----|
| | | | char | 7 | | 1 | | N | N |
| | | | num | 5 | 0 | 1 | | N | N |
| | | | num | 1 | 0 | 6 | 1 - 7 | Y | Y |
| 005 | G33a_3 | | num | 1 | 0 | 6 | 1 - 7 | Y | Y |
| 006 | G33a_4 | | num | 1 | 0 | 6 | 1 - 7 | Y | Y |
| 007 | G33a_5 | | num | 1 | 0 | 6 | 1 - 7 | Y | Y |
| 008 | G33a_6 | | num | 1 | 0 | 6 | 1 - 7 | Y | Y |
| 009 | G33aSpec | | char | 100 | | 1 | | Y | Y |
| 010 | G34 | | num | 6 | 0 | 1 | 1 - 999997 | Y | Y |
| 011 | G34Per | | num | 1 | 0 | 1 | 5 - 7 | Y | Y |
| 012 | G34PerSpec | | char | 100 | | 1 | | Y | Y |
| 013 | G35_01 | | num | 2 | 0 | 13 | 1 - 13 | Y | Y |

The frequency table for variable G34 is shown on the right:

| Value | Count | Percent |
|------------|-------|---------|
| 1 - 150000 | 2430 | 93.21 |
| 999998 | 90 | 3.45 |
| 999999 | 87 | 3.34 |

Annotations in the image include:

- A callout box pointing to the variable table columns: **Helpful information displayed including width, loops, whether DK/RF allowed**.
- An arrow pointing to the 'Freqs' checkbox in the 'Variables' dialog for G34: **User clicks here to see the freqs**.
- An arrow pointing to the frequency table: **Displays the frequencies**.

Define SAS Setups

Example Transformation from Wide to Long

| VIEWTABLE: Btio.Bdeceased | | | | | | | | | |
|---------------------------|-----------|-------------------|----|-------|-------|-------|-------|-------|---|
| | BlaiseKey | BDECEASEDInstance | D1 | D2_01 | D2_02 | D2_03 | D2_04 | D2_05 | |
| 1 | 0004211 | | 1 | 97 | 1 | . | . | . | . |
| 2 | 0041134 | | 1 | 1 | 1 | . | . | . | . |
| 3 | 0079001 | | 1 | 97 | 1 | . | . | . | . |
| 4 | 0088003 | | 1 | 97 | 1 | 2 | . | . | . |
| 5 | 0117001 | | 1 | 97 | 1 | . | . | . | . |
| 6 | 0129004 | | 1 | 1 | 1 | . | . | . | . |
| 7 | 0165002 | | 1 | 97 | 1 | . | . | . | . |
| 8 | 0173002 | | 1 | 97 | 1 | . | . | . | . |
| 9 | 0297131 | | 1 | 1 | 1 | . | . | . | . |

Collected in a wide format- with lots of empty space

| VIEWTABLE: BI11.D1d3 | | | | | |
|----------------------|---------|---------|----|----|-------|
| | BKEY | DECINST | D1 | i1 | AQSNN |
| 1 | 0004211 | 1 | 97 | 1 | 1 |
| 25 | 0041134 | 1 | 1 | 1 | 1 |
| 49 | 0079001 | 1 | 97 | 1 | 1 |
| 73 | 0088003 | 1 | 97 | 1 | 1 |
| 74 | 0088003 | 1 | 97 | 2 | 2 |
| 97 | 0117001 | 1 | 97 | 1 | 1 |
| 121 | 0129004 | 1 | 1 | 1 | 1 |
| 145 | 0165002 | 1 | 97 | 1 | 1 |
| 169 | 0173002 | 1 | 97 | 1 | 1 |
| 193 | 0297131 | 1 | 1 | 1 | 1 |
| 217 | 0301003 | 1 | 2 | 1 | 1 |
| 241 | 0393001 | 1 | 1 | 1 | 1 |
| 265 | 0418001 | 1 | 97 | 1 | 1 |
| 289 | 0419002 | 1 | 1 | 1 | 1 |

Transformed to a long format- only keep records filled that have data

Define SAS Setups

Merge Type

File Type: 0. <NULL> 1. Long 2. Wide 3. Merge (recode)

Action: Query Files Add New File

Select a File: BC21ATRAN BC41 BC51TRAN BC61BC67 BC63 BCAQSN

Data IF: CYFuHuN IN(1,4) and CYAQRTHN IN(101,102)

Long/Wide ID Var:

SAS Process: Code Data

Tab ID Change

Add Vars: Blaise SAS (PP)

Sub-set using an IF statement

Click radio button to add Blaise tables or choose SAS tables you have already generated

Display Input Variables and Edit Specs

Sort by: tabID varID Name

Keep vars: All None

| Tabid | Varid | Name | Keep | Type | Zero Fill | Loops | Multi? | Which MultiType | Width | New | Key | New Name | New Var |
|-------|-------|---------------------|-------------------------------------|------|-------------------------------------|-------|--------------------------|-----------------|-------|-----|-------------------------------------|---------------|--------------------------|
| 0006 | 001 | BlaiseKey | <input checked="" type="checkbox"/> | char | <input type="checkbox"/> | 1 | <input type="checkbox"/> | | 7 | 7 | <input checked="" type="checkbox"/> | BKEY | <input type="checkbox"/> |
| 0006 | 002 | BCYPersInfoInstance | <input type="checkbox"/> | num | <input checked="" type="checkbox"/> | 1 | <input type="checkbox"/> | | 5 | 0 | <input type="checkbox"/> | CYPERINFOINST | <input type="checkbox"/> |
| 0006 | 003 | AQSN | <input checked="" type="checkbox"/> | num | <input checked="" type="checkbox"/> | 1 | <input type="checkbox"/> | | 2 | 0 | <input type="checkbox"/> | AQSN | <input type="checkbox"/> |
| 0006 | 004 | CYFuHu | <input checked="" type="checkbox"/> | num | <input checked="" type="checkbox"/> | 1 | <input type="checkbox"/> | | 2 | 0 | <input type="checkbox"/> | CYFuHuN | <input type="checkbox"/> |
| 0006 | 005 | CYAQRTH | <input checked="" type="checkbox"/> | num | <input checked="" type="checkbox"/> | 1 | <input type="checkbox"/> | | 3 | 0 | <input type="checkbox"/> | CYAQRTHN | <input type="checkbox"/> |

Input Manager

File/variables Delete

Table ID: [Dropdown]

VarID From: [Dropdown]

VarID To: [Dropdown]

Rename variables [Dropdown]

Delete

Handling Multi-mentions

During which months of 2010 did you get this income?

- ♦ ENTER all that apply
- ♦ For multiple response, use space bar or dash to separate responses

- | | | |
|-----------------------------------|----------------------------------|---|
| <input type="checkbox"/> 1. Jan | <input type="checkbox"/> 7. July | <input checked="" type="checkbox"/> 13. All |
| <input type="checkbox"/> 2. Feb | <input type="checkbox"/> 8. Aug | |
| <input type="checkbox"/> 3. March | <input type="checkbox"/> 9. Sept | |
| <input type="checkbox"/> 4. April | <input type="checkbox"/> 10. Oct | |
| <input type="checkbox"/> 5. May | <input type="checkbox"/> 11. Nov | |
| <input type="checkbox"/> 6. June | <input type="checkbox"/> 12. Dec | |

Months in 2010 Rec'd SS Ben

13

G35_M

```
111111111111
111111111111
111111111111
111111111111
111111111111
111111111111
111111111111
000001100000
111111111111
111111111111
000011111111
111111111111
111111111111
111111111111
111111111111
111111111111
111111111111
000111111111
111111111111
```

Output type=
Month String

Resulting SAS
output

Pull Forward

Variable Change Report from 2009 to 2011

PSID Blaise Variable Name Change Report
(PY = 2009, CY = 2011)
Count = 580

| Table | Bmi | Year | Variable | CY/PY Status |
|---------|--------|------|-------------|----------------|
| BA37FOR | PSID09 | 2009 | A37FOR2Mo | Dropped in PY. |
| | | | A37FOR2Yr | Dropped in PY. |
| | | | A37FOR3 | Dropped in PY. |
| | | | A37FOR4 | Dropped in PY. |
| | | | A37FOR5 | Dropped in PY. |
| | | | A37FOR5SPEC | Dropped in PY. |
| | | | A37FOR6 | Dropped in PY. |
| BA37FOR | PSID11 | 2011 | A37B | New in CY. |
| | | | A37C | New in CY. |
| | | | A37D | New in CY. |
| | | | A37E | New in CY. |
| | | | A37F | New in CY. |
| | | | A37FSPEC | New in CY. |
| | | | A37G | New in CY. |

Handling Changes between Pretest and Production

1. Long
2. Wide
3. Merge (recode)

Add New File

HEHCSTOMACH
INDR2HW
INDS
INSTADDR
IO26

Tab ID Change

Add Vars:
 Blaise SAS (PP)

SAS Pro
 Code

Display Input Variables and Edit Specs

Sort by:
 tabID varID Name

Keep vars:
 All None

Zero Fill Multi? Which MultiType Width New Width Key Var? New Name Copy New Var

| Tabid | Varid | Name | Keep | Type | Loops | Width | New Width | Key Var? | New Name | Copy New Var |
|-------|-------|----------|------|------|-------|-------|-----------|-------------------------------------|----------|--------------------------|
| 0002 | | BKEY | | | | 7 | 7 | <input checked="" type="checkbox"/> | | <input type="checkbox"/> |
| 0002 | | PYINST | | | | 5 | 5 | <input type="checkbox"/> | | <input type="checkbox"/> |
| 0002 | | AQSNN | | | | 2 | 2 | <input checked="" type="checkbox"/> | | <input type="checkbox"/> |
| 0002 | | PYAQRTHN | | | | 3 | 3 | <input type="checkbox"/> | | <input type="checkbox"/> |
| 0002 | | PYFUHUN | | | | 2 | 2 | <input type="checkbox"/> | | <input type="checkbox"/> |
| 0002 | | FSN | | | | 1 | 1 | <input type="checkbox"/> | | <input type="checkbox"/> |
| 0002 | | WHYFOLLN | | | | 2 | 2 | <input type="checkbox"/> | | <input type="checkbox"/> |
| 0002 | | PNN | | | | 3 | 3 | <input type="checkbox"/> | | <input type="checkbox"/> |
| 0002 | | ID68C | | | | 4 | 4 | <input type="checkbox"/> | | <input type="checkbox"/> |
| 0002 | | DADFUN | | | | 1 | 1 | <input type="checkbox"/> | | <input type="checkbox"/> |

Table ID Change

| Tablename | Tabid | New Tabid |
|-----------|-------|-----------|
| INDS | 2 | |
| INDS | 6 | |
| INDS | 8 | |
| INDS | 9 | |
| INDS | 10 | |

Close Change

File Management

Wave:
2011-PSID

File Management:

Select Type: Long Wide Merge

| Name | Type | SAS Code | SAS Data | Created By | Notes | |
|-------------|---------------|--------------------------|--------------------------|------------|----------------------------|--------|
| A37_B | Merge(recode) | <input type="checkbox"/> | <input type="checkbox"/> | ABEAULE | step b merge wide with and | Delete |
| BC14A14F | Merge(recode) | <input type="checkbox"/> | <input type="checkbox"/> | ABEAULE | | Delete |
| BC14CTRAN | Merge(recode) | <input type="checkbox"/> | <input type="checkbox"/> | ABEAULE | | Delete |
| BC14FTRAN | Merge(recode) | <input type="checkbox"/> | <input type="checkbox"/> | ABEAULE | | Delete |
| BC18JOBTYPE | Merge(recode) | <input type="checkbox"/> | <input type="checkbox"/> | ABEAULE | | Delete |
| BC21ATRAN | Merge(recode) | <input type="checkbox"/> | <input type="checkbox"/> | ABEAULE | | Delete |
| BC41 | Merge(recode) | <input type="checkbox"/> | <input type="checkbox"/> | ABEAULE | | Delete |
| BC51TRAN | Merge(recode) | <input type="checkbox"/> | <input type="checkbox"/> | ABEAULE | | Delete |
| BC61BC67 | Merge(recode) | <input type="checkbox"/> | <input type="checkbox"/> | ABEAULE | | Delete |
| BC63 | Merge(recode) | <input type="checkbox"/> | <input type="checkbox"/> | ABEAULE | | Delete |

Copy or Rename:

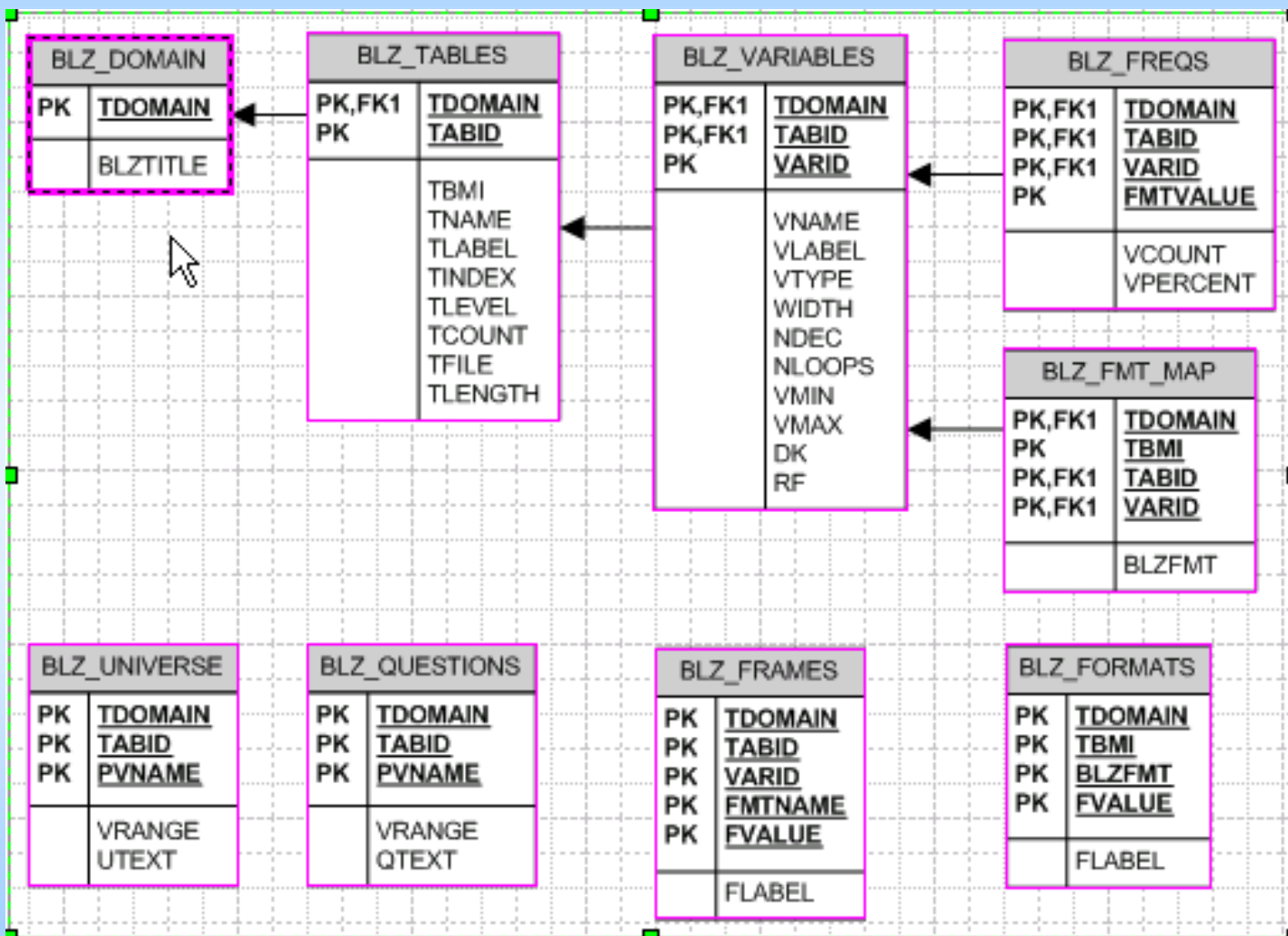
Copy Rename

Input Filename:
<NULL>

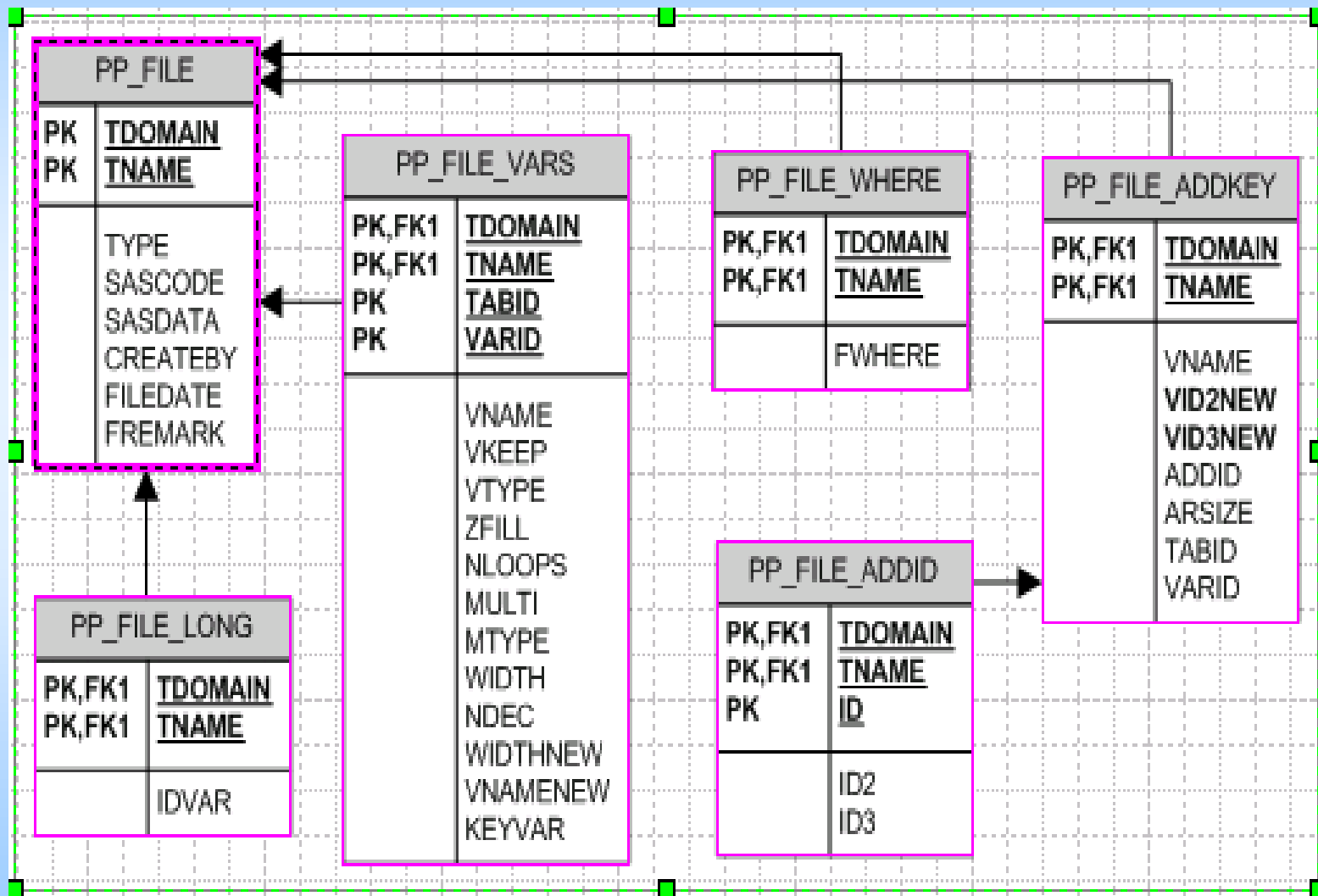
Output Filename:

Finish

Data model 1



Data model 2



Conclusion

- DETL provides reports that allow processing staff to methodically correct and adjust for every variable change
- Informs the user about the metadata as setups are created
- Quick and easy transformations of multi-mention variables
- SAS code and log can be run and saved all with one click
- Assists the user in making decisions about how to organize content and at what level

Thank You

April Beaulé

abeaule@isr.umich.edu

Mohammad Mushtaq

mushtaq@isr.umich.edu