

Session 7 - SPSS¹

In this class we will reorganize a data set - changing variables to cases. Using the 1995 Viet Nam data set, we will create a Child-Level data set. Instead of each woman being a case, each children will be a case. We will do this in two ways so that we can learn different tools. We will use Add Files, and write out raw data.

In this class, we will examine two ways of transforming data so that we change variables into cases. This is a surprisingly common problem. I will present two programs that deal with 14 variables below, but before I do that, I want to prepare by giving a simpler example, and explaining why this frequently needs to be done.

Suppose you wanted to examine criminal behavior. You have a data set that has detailed self-reported criminal behavior information during the past year. People report the offense, and if they were employed at the time. They can report this information for up to three crimes. Your research question is: Are people more likely to commit a crime if they are unemployed? Each respondent may have been employed for some offences, and not others. A way of dealing with this problem is to restructure the data set so that each crime is a case, rather than each person. (There are issues to consider like weighting and clustering that will be ignored here.) A graphical presentation of how the data needs to be transformed follows:

As person file:

Person ID	Age	Race	Crime 1	Empl 1	Crime 2	Empl 2	Crime 3	Empl 3
1								
2								

As crime file.

person1	age	Race	crime	employed
person1	age	Race	crime	employed
person1	age	Race	crime	employed
person2	age	Race	crime	employed
person2	age	Race	crime	employed
person2	age	Race	crime	employed

Now I will present two programs that show different ways of doing this. Besides learning how to transform data, we will learn some exciting tools.

```
*** PROGRAM BEGINS HERE *** .
* vn95u5.sps .
** Create child-level file from woman-level file
** by creating a file for each child, and then using Add Files.
title 'c:\all\SPSSclass\vn95\vn95u5.sps' .
* Documentation and data from following sites .
* http://csde.washington.edu/csde/vietnam/documents.html .
* http://csde.washington.edu/csde/vietnam/data.html .
** Read SPSS data set created by vn95.SPSS .
* Select only ever-married women, calculate age .
* Use "DO REPEAT" .

get file = 'c:\all\SPSSclass\vn95\data\vn95.sav'
/ keep = commid hhid I5C1B I5C5A I5Q1B I5E11
I5E4C1 I5E4C2 I5E4C3 I5E4C4 I5E4C5 I5E4C6 I5E4C7
I5E4C8 I5E4C9 I5E4C10 I5E4C11 I5E4C12 I5E4C13 I5E4C14
I5E5BC1 I5E5BC2 I5E5BC3 I5E5BC4 I5E5BC5 I5E5BC6 I5E5BC7
I5E5BC8 I5E5BC9 I5E5BC10 I5E5BC11 I5E5BC12 I5E5BC13 I5E5BC14
I5E6C1 I5E6C2 I5E6C3 I5E6C4 I5E6C5 I5E6C6 I5E6C7
I5E6C8 I5E6C9 I5E6C10 I5E6C11 I5E6C12 I5E6C13 I5E6C14
```

¹This document has been prepared by Patty Glynn, University of Washington.

```
I5E8c1  I5E8c2  I5E8c3  I5E8c4  I5E8c5  I5E8c6  I5E8c7
I5E8c8  I5E8c9  I5E8c10 I5E8c11 I5E8c12 I5E8c13 I5E8c14 .
```

```
***** Selection begins *****
* Select only ever-married women who are heads of households .
* or wives of heads of households .

* Select only ever-married women .
* women only .
select if I5Q1B = 2 .

* I5E11 How many children have you had? .

** select women who have had any children .
select if I5E11 ge 1 .

***** Variables pertaining to children that we want: ** .
* Sex of child          I5E4C1 - I5E4C14 .
* Year of birth         I5E5BC1 - I5E5BC14 .
* Child wanted         I5E6C1 - I5E6C14 .
* If died, age at death I5E8C1 - I5E8C14 .

** We will need to save a number of data sets during this process
* that we do not need at the end.
* Before you run this program, create a directory on a drive that
* you have write access to that is called spsstrsh (for spss trash).
* You can write files to this directory that you will delete later.

** Keep all cases - we will use this file to create the child files.
save out = 'c:\spsstrsh\all.sav' .

** Why have the following lined up? For ease of writing program
** and spotting errors.

*          child 1 .
get file = 'c:\spsstrsh\all.sav' / rename
          I5E4C1 = chsex
          I5E5BC1 = chyob
          I5E6C1 = chwant
          I5E8c1 = chdeath .
compute childnum = 1 .
** select only "valid" children .
select if
  chsex = 1 or
  chsex = 2 .
save outfile =
  'c:\spsstrsh\ch1.sav' .

*          child 2 .
get file = 'c:\spsstrsh\all.sav' / rename
          I5E4C2 = chsex
          I5E5BC2 = chyob
          I5E6C2 = chwant
          I5E8c2 = chdeath .
compute childnum = 2 .
** select only "valid" children .
select if
  chsex = 1 or
  chsex = 2 .
save outfile =
  'c:\spsstrsh\ch2.sav' .

*          child 3 .
get file = 'c:\spsstrsh\all.sav' / rename
          I5E4C3 = chsex
```

```

        I5E5BC3 = chyob
        I5E6C3 = chwant
        I5E8c3 = chdeath .
compute childnum = 3 .
** select only "valid" children .
select if
  chsex = 1 or
  chsex = 2 .
save outfile =
  'c:\spsstrsh\ch3.sav' .

*
  child 4 .
get file = 'c:\spsstrsh\all.sav' / rename
  I5E4C4 = chsex
  I5E5BC4 = chyob
  I5E6C4 = chwant
  I5E8c4 = chdeath .
compute childnum = 4 .
** select only "valid" children .
select if
  chsex = 1 or
  chsex = 2 .
save outfile =
  'c:\spsstrsh\ch4.sav' .

*
  child 5 .
get file = 'c:\spsstrsh\all.sav' / rename
  I5E4C5 = chsex
  I5E5BC5 = chyob
  I5E6C5 = chwant
  I5E8c5 = chdeath .
compute childnum = 5 .
** select only "valid" children .
select if
  chsex = 1 or
  chsex = 2 .
save outfile =
  'c:\spsstrsh\ch5.sav' .

*
  child 6 .
get file = 'c:\spsstrsh\all.sav' / rename
  I5E4C6 = chsex
  I5E5BC6 = chyob
  I5E6C6 = chwant
  I5E8c6 = chdeath .
compute childnum = 6 .
** select only "valid" children .
select if
  chsex = 1 or
  chsex = 2 .
save outfile =
  'c:\spsstrsh\ch6.sav' .

*
  child 7 .
get file = 'c:\spsstrsh\all.sav' / rename
  I5E4C7 = chsex
  I5E5BC7 = chyob
  I5E6C7 = chwant
  I5E8c7 = chdeath .
compute childnum = 7 .
** select only "valid" children .
select if
  chsex = 1 or
  chsex = 2 .
save outfile =
  'c:\spsstrsh\ch7.sav' .

```

```

*           child 8 .
get file = 'c:\spsstrsh\all.sav' / rename
           I5E4C8  = chsex
           I5E5BC8 = chyob
           I5E6C8  = chwant
           I5E8c8  = chdeath .
compute childnum = 8 .
** select only "valid" children .
select if
  chsex = 1 or
  chsex = 2 .
save outfile =
  'c:\spsstrsh\ch8.sav' .

*           child 9 .
get file = 'c:\spsstrsh\all.sav' / rename
           I5E4C9  = chsex
           I5E5BC9 = chyob
           I5E6C9  = chwant
           I5E8c9  = chdeath .
compute childnum = 9 .
** select only "valid" children .
select if
  chsex = 1 or
  chsex = 2 .
save outfile =
  'c:\spsstrsh\ch9.sav' .

*           child 10 .
get file = 'c:\spsstrsh\all.sav' / rename
           I5E4C10 = chsex
           I5E5BC10 = chyob
           I5E6C10 = chwant
           I5E8c10 = chdeath .
compute childnum = 10 .
** select only "valid" children .
select if
  chsex = 1 or
  chsex = 2 .
save outfile =
  'c:\spsstrsh\ch10.sav' .

*           child 11 .
get file = 'c:\spsstrsh\all.sav' / rename
           I5E4C11 = chsex
           I5E5BC11 = chyob
           I5E6C11 = chwant
           I5E8c11 = chdeath .
compute childnum = 11 .
** select only "valid" children .
select if
  chsex = 1 or
  chsex = 2 .
save outfile =
  'c:\spsstrsh\ch11.sav' .

*           child 12 .
get file = 'c:\spsstrsh\all.sav' / rename
           I5E4C12 = chsex
           I5E5BC12 = chyob
           I5E6C12 = chwant
           I5E8C12 = chdeath .
compute childnum = 12 .
** select only "valid" children .
select if
  chsex = 1 or

```

```

chsex = 2 .
save outfile =
  'c:\spsstrsh\ch12.sav' .

*          child 13 .
get file = 'c:\spsstrsh\all.sav' / rename
          I5E4C13 = chsex
          I5E5BC13 = chyob
          I5E6C13 = chwant
          I5E8C13 = chdeath .
compute childnum = 13 .
** select only "valid" children .
select if
  chsex = 1 or
  chsex = 2 .
save outfile =
  'c:\spsstrsh\ch13.sav' .

*          child 14 .
get file = 'c:\spsstrsh\all.sav' / rename
          I5E4C14 = chsex
          I5E5BC14 = chyob
          I5E6C14 = chwant
          I5E8C14 = chdeath .
compute childnum = 14 .
** select only "valid" children .
select if
  chsex = 1 or
  chsex = 2 .
save outfile =
  'c:\spsstrsh\ch14.sav' .

*** Put all of the files together .
*** Since we are adding cases - not adding variables - we use the
*** command ADD FILES .

ADD FILES
FILE= 'c:\spsstrsh\ch1.sav' /
FILE= 'c:\spsstrsh\ch2.sav' /
FILE= 'c:\spsstrsh\ch3.sav' /
FILE= 'c:\spsstrsh\ch4.sav' /
FILE= 'c:\spsstrsh\ch5.sav' /
FILE= 'c:\spsstrsh\ch6.sav' /
FILE= 'c:\spsstrsh\ch7.sav' /
FILE= 'c:\spsstrsh\ch8.sav' /
FILE= 'c:\spsstrsh\ch9.sav' /
FILE= 'c:\spsstrsh\ch10.sav' /
FILE= 'c:\spsstrsh\ch11.sav' /
FILE= 'c:\spsstrsh\ch12.sav' /
FILE= 'c:\spsstrsh\ch13.sav' /
FILE= 'c:\spsstrsh\ch14.sav' .

variable label
chsex   'I5E4C Sex of child'
chyob   'I5E4C Year of birth of Child'
chwant  'I5E6C Child Wanted?'
chdeath 'I5E8C If dead, how old?'.

** frequencies on chyob showed out-of-range values on year of birth .
missing values chyob (9998,9999) .
freq var = childnum.

save outfile = 'c:\spsstrsh\vn95u5' /
keep commid hhid I5C1B I5C5A I5Q1B I5E11
childnum chsex chyob chwant chdeath .

```

execute .

CHILDNUM

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1.00	1881	30.1	30.1	30.1
2.00	1630	26.0	26.0	56.1
3.00	1124	18.0	18.0	74.1
4.00	693	11.1	11.1	85.1
5.00	411	6.6	6.6	91.7
6.00	237	3.8	3.8	95.5
7.00	154	2.5	2.5	97.9
8.00	73	1.2	1.2	99.1
9.00	33	.5	.5	99.6
10.00	14	.2	.2	99.9
11.00	6	.1	.1	100.0
12.00	3	.0	.0	100.0
Total	6259	100.0	100.0	

Second option. Write out raw data and read it again.

```
*** PROGRAM BEGINS HERE *** .
* vn95u6.sps .
** Create child-level file from woman-level file
** by writing raw data and reading it again.
title "c:\all\SPSSclass\vn95\vn95u5.sps" .
* Documentation and data from following sites .
* http://csde.washington.edu/csde/vietnam/documents.html .
* http://csde.washington.edu/csde/vietnam/data.html .
** Read SPSS data set created by vn95.SPSS .
* Select only ever-married women, calculate age .
* Use "DO REPEAT" .

get file = "c:\all\SPSSclass\vn95\data\vn95.sav"
/ keep = commid hhid I5C1B I5C5A I5Q1B I5E11
I5E4C1 I5E4C2 I5E4C3 I5E4C4 I5E4C5 I5E4C6 I5E4C7
I5E4C8 I5E4C9 I5E4C10 I5E4C11 I5E4C12 I5E4C13 I5E4C14
I5E5BC1 I5E5BC2 I5E5BC3 I5E5BC4 I5E5BC5 I5E5BC6 I5E5BC7
I5E5BC8 I5E5BC9 I5E5BC10 I5E5BC11 I5E5BC12 I5E5BC13 I5E5BC14
I5E6C1 I5E6C2 I5E6C3 I5E6C4 I5E6C5 I5E6C6 I5E6C7
I5E6C8 I5E6C9 I5E6C10 I5E6C11 I5E6C12 I5E6C13 I5E6C14
I5E8c1 I5E8c2 I5E8c3 I5E8c4 I5E8c5 I5E8c6 I5E8c7
I5E8c8 I5E8c9 I5E8c10 I5E8c11 I5E8c12 I5E8c13 I5E8c14 .

***** Selection begins ***** .
* Select only ever-married women who are heads of households .
* or wives of heads of households .

* Select only ever-married women .
* women only .
select if I5Q1B = 2 .

* I5E11 How many children have you had? .

** select women who have had any children .
```

```
select if I5E11 ge 1 .
```

```
***** Variables pertaining to children that we want: ** .
```

```
* Sex of child          I5E4C1 - I5E4C14 .  
* Year of birth         I5E5BC1 - I5E5BC14 .  
* Child wanted         I5E6C1 - I5E6C14 .  
* If died, age at death I5E8C1 - I5E8C14 .
```

```
** We will need to save raw data during this process  
* that we do not need at the end.  
* Before you run this program, create a directory on a drive that  
* you have write access to that is called spsstrsh (for spss trash).  
* You can write files to this directory that you will delete later.
```

```
** the variable for Age at Death is a string variable, and missing value  
** may be blanks. This will cause problems if trying to read the data free feild.  
** Use a recode .
```

```
recode  
  I5E8c1  I5E8c2  I5E8c3  I5E8c4  I5E8c5  I5E8c6  I5E8c7  
  I5E8c8  I5E8c9  I5E8c10 I5E8c11 I5E8c12 I5E8c13 I5E8c14 (" " = "NA" ) .
```

```
** system missing cases are not written properly in write cases .  
** must change to 99 .
```

```
recode  
  I5E4C1  I5E4C2  I5E4C3  I5E4C4  I5E4C5  I5E4C6  I5E4C7  
  I5E4C8  I5E4C9  I5E4C10 I5E4C11 I5E4C12 I5E4C13 I5E4C14  
  I5E5BC1 I5E5BC2 I5E5BC3 I5E5BC4 I5E5BC5 I5E5BC6 I5E5BC7  
  I5E5BC8 I5E5BC9 I5E5BC10 I5E5BC11 I5E5BC12 I5E5BC13 I5E5BC14  
  I5E6C1  I5E6C2  I5E6C3  I5E6C4  I5E6C5  I5E6C6  I5E6C7  
  I5E6C8  I5E6C9  I5E6C10 I5E6C11 I5E6C12 I5E6C13 I5E6C14  
  (sysmis = 99) .
```

```
execute .
```

```
** write out the variables and and indicator of the child number .  
** Text can be enclosed in quotes.  
** Numbers are right justified and string variables are left justified.  
** Must put space between string and number.
```

```
WRITE OUTFILE="c:\spsstrsh\childfile.dat" records = 14 TABLE  
 / commid hhid I5C1B I5C5A I5Q1B I5E11 I5E4C1 I5E5BC1 I5E6C1 (9(f6.0))  
 ' ' I5E8c1 * " 1"  
 / commid hhid I5C1B I5C5A I5Q1B I5E11 I5E4C2 I5E5BC2 I5E6C2 (9(f6.0))  
 ' ' I5E8c2 * " 2"  
 / commid hhid I5C1B I5C5A I5Q1B I5E11 I5E4C3 I5E5BC3 I5E6C3 (9(f6.0))  
 ' ' I5E8c3 * " 3"  
 / commid hhid I5C1B I5C5A I5Q1B I5E11 I5E4C4 I5E5BC4 I5E6C4 (9(f6.0))  
 ' ' I5E8c4 * " 4"  
 / commid hhid I5C1B I5C5A I5Q1B I5E11 I5E4C5 I5E5BC5 I5E6C5 (9(f6.0))  
 ' ' I5E8c5 * " 5"  
 / commid hhid I5C1B I5C5A I5Q1B I5E11 I5E4C6 I5E5BC6 I5E6C6 (9(f6.0))  
 ' ' I5E8c6 * " 6"  
 / commid hhid I5C1B I5C5A I5Q1B I5E11 I5E4C7 I5E5BC7 I5E6C7 (9(f6.0))  
 ' ' I5E8c7 * " 7"  
 / commid hhid I5C1B I5C5A I5Q1B I5E11 I5E4C8 I5E5BC8 I5E6C8 (9(f6.0))  
 ' ' I5E8c8 * " 8"  
 / commid hhid I5C1B I5C5A I5Q1B I5E11 I5E4C9 I5E5BC9 I5E6C9 (9(f6.0))  
 ' ' I5E8c9 * " 9"  
 / commid hhid I5C1B I5C5A I5Q1B I5E11 I5E4C10 I5E5BC10 I5E6C10 (9(f6.0))  
 ' ' I5E8c10 * " 10"  
 / commid hhid I5C1B I5C5A I5Q1B I5E11 I5E4C11 I5E5BC11 I5E6C11 (9(f6.0))  
 ' ' I5E8c11 * " 11"  
 / commid hhid I5C1B I5C5A I5Q1B I5E11 I5E4C12 I5E5BC12 I5E6C12 (9(f6.0))  
 ' ' I5E8c12 * " 12"  
 / commid hhid I5C1B I5C5A I5Q1B I5E11 I5E4C13 I5E5BC13 I5E6C13 (9(f6.0))  
 ' ' I5E8c13 * " 13"  
 / commid hhid I5C1B I5C5A I5Q1B I5E11 I5E4C14 I5E5BC14 I5E6C14 (9(f6.0))  
 ' ' I5E8c14 * " 14"
```

execute .

```
DATA LIST file = "c:\spsstrsh\childfile.dat"  
  / commid hhid I5C1B I5C5A I5Q1B I5E11 chsex chyob chwant (9(f6.0))  
  chdeath (a4) childnum (f3.0) .
```

execute .

variable label

```
chsex      'I5E4C Sex of child'  
chyob      'I5E4C Year of birth of Child'  
chwant     'I5E6C Child Wanted?'  
chdeath    'I5E8C If dead, how old?'  
HHID       'Survey Household Number'  
I5C1B      'respondent year of birth'  
I5C5A      'highest grade of regular education completed'  
I5Q1B      'Sex of respondent'  
I5E11      'How many children have you had?' .
```

```
** Select only "REAL" children .  
select if chsex = 1 or chsex = 2 .
```

```
missing values chsex chyob chwant (99,9998,9999) .  
freq var = chsex chyob chwant .
```

```
freq var = childnum .
```

CHILDNUM

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1881	30.1	30.1	30.1
	2	1630	26.0	26.0	56.1
	3	1124	18.0	18.0	74.1
	4	693	11.1	11.1	85.1
	5	411	6.6	6.6	91.7
	6	237	3.8	3.8	95.5
	7	154	2.5	2.5	97.9
	8	73	1.2	1.2	99.1
	9	33	.5	.5	99.6
	10	14	.2	.2	99.9
	11	6	.1	.1	100.0
	12	3	.0	.0	100.0
	Total	6259	100.0	100.0	

HOMEWORK: MORE ABOUT WRITING RAW DATA .

Write programs to practice what you learn about writing raw data. Refer to the section about "WRITE" in the SPSS documentation available at:

<https://depts.washington.edu/csde1/spss/spssbase.pdf>

If you can't access this file, it came with SPSS 10 on the CD.