

```
-----  
name: <unnamed>  
log: D:\Jason\workshop\long format data\reshape and merge2.log  
log type: text  
opened on: 6 Nov 2023, 11:07:24
```

```
. *****  
. * Change the data from wide to long  
. *****  
. use "D:\Jason\workshop\long format data\reshape.dta", clear
```

```
. list, sepby(family_id)
```

```
+-----+  
| family~d  sex1  sex2  sex3  sex4  sex5  sex6  sex7  sex8 |  
+-----+  
1. |          9    2    2    .    .    .    .    . |  
+-----+  
2. |       8212    1    2    1    1    2    2    1    2 |  
+-----+
```

```
. reshape long sex, i(family_id) j(pernum)  
(j = 1 2 3 4 5 6 7 8)
```

```
Data                               Wide  ->  Long  
-----  
Number of observations              2  ->  16  
Number of variables                  9  ->  3  
j variable (8 values)                ->  pernum  
xij variables:                       sex1 sex2 ... sex8 ->  sex  
-----
```

```
. sort family_id pernum
```

```
. list, sepby(family_id)
```

```
+-----+  
| family~d  pernum  sex |  
+-----+  
1. |          9        1    2 |  
2. |          9        2    2 |  
3. |          9        3    . |  
4. |          9        4    . |  
5. |          9        5    . |  
6. |          9        6    . |  
7. |          9        7    . |  
8. |          9        8    . |  
+-----+  
9. |       8212        1    1 |  
10. |       8212        2    2 |  
11. |       8212        3    1 |  
12. |       8212        4    1 |  
13. |       8212        5    2 |  
14. |       8212        6    2 |  
15. |       8212        7    1 |  
16. |       8212        8    2 |  
+-----+
```

```
. *remove records with missing values  
. drop if sex ==.  
(6 observations deleted)
```

```
. list, sepby(family_id)
```

```
+-----+  
| family~d  pernum  sex |  
+-----+  
1. |          9        1    2 |  
2. |          9        2    2 |  
+-----+  
3. |       8212        1    1 |  
4. |       8212        2    2 |  
5. |       8212        3    1 |  
6. |       8212        4    1 |  
7. |       8212        5    2 |  
8. |       8212        6    2 |  
9. |       8212        7    1 |  
10. |       8212        8    2 |
```

```

+-----+
.
. save "D:\Jason\workshop\long format data\long format.dta", replace
file D:\Jason\workshop\long format data\long format.dta saved
.
. *****
. * Change the data from long to wide
. *****
.
. reshape wide sex, i(family_id) j(pernum)
(j = 1 2 3 4 5 6 7 8)
Data                Long   ->   Wide
-----
Number of observations      10   ->   2
Number of variables        3     ->   9
j variable (8 values)      pernum ->   (dropped)
xij variables:
                        sex   ->   sex1 sex2 ... sex8
-----

. list, sepby(family_id)
+-----+
| family~d  sex1  sex2  sex3  sex4  sex5  sex6  sex7  sex8 |
|-----|
1. |          9    2    2    .    .    .    .    . |
|-----|
2. |       8212    1    2    1    1    2    2    1    2 |
+-----+

. sort family_id

. save "D:\Jason\workshop\long format data\wide format.dta", replace
file D:\Jason\workshop\long format data\wide format.dta saved
.
.
. *****
. * Possible Error #1: mixed types of variables to be combined into one variable
. *****
.
. clear

. use "D:\Jason\workshop\long format data\reshapel.dta", clear

. des

Contains data from D:\Jason\workshop\long format data\reshapel.dta
Observations:      2
Variables:         9          6 Nov 2023 10:51
-----
Variable   Storage  Display  Value  Variable label
  name      type    format  label
-----
family_id  int      %10.0g
sex1       byte    %10.0g
sex2       byte    %10.0g
sex3       byte    %10.0g
sex4       byte    %10.0g
sex5       byte    %10.0g
sex6       byte    %10.0g
sex7       byte    %10.0g
sex8       str1    %9s
-----

Sorted by:

.
. reshape long sex, i(family_id) j(pernum)
(j = 1 2 3 4 5 6 7 8)
variable sex8 type mismatch with other sex variables
r(198);

end of do-file

r(198);

. destring sex8, replace
sex8: all characters numeric; replaced as byte

```

(1 missing value generated)

.
. des

Contains data from D:\Jason\workshop\long format data\reshape1.dta

Observations: 2
Variables: 9 6 Nov 2023 10:51

Variable name	Storage type	Display format	Value label	Variable label
family_id	int	%10.0g		family_id
sex1	byte	%10.0g		sex1
sex2	byte	%10.0g		sex2
sex3	byte	%10.0g		sex3
sex4	byte	%10.0g		sex4
sex5	byte	%10.0g		sex5
sex6	byte	%10.0g		sex6
sex7	byte	%10.0g		sex7
sex8	byte	%10.0g		sex8

Sorted by:
Note: Dataset has changed since last saved.

.
. reshape long sex, i(family_id) j(pernum)
(j = 1 2 3 4 5 6 7 8)

Data	Wide	->	Long
Number of observations	2	->	16
Number of variables	9	->	3
j variable (8 values)		->	pernum
xij variables:	sex1 sex2 ... sex8	->	sex

. list, sepby(family_id)

	family~d	pernum	sex
1.	9	1	2
2.	9	2	2
3.	9	3	.
4.	9	4	.
5.	9	5	.
6.	9	6	.
7.	9	7	.
8.	9	8	.
9.	8212	1	1
10.	8212	2	2
11.	8212	3	1
12.	8212	4	1
13.	8212	5	2
14.	8212	6	2
15.	8212	7	1
16.	8212	8	2

. *****
. * Possible Error #2: the name of new variable has already been used
. *****

. use "D:\Jason\workshop\long format data\reshape2.dta", clear

. des

Contains data from D:\Jason\workshop\long format data\reshape2.dta

Observations: 2
 Variables: 9 6 Nov 2023 11:00

Variable name	Storage type	Display format	Value label	Variable label
family_id	int	%10.0g		family_id
sex1	byte	%10.0g		sex1
sex2	byte	%10.0g		sex2
sex3	byte	%10.0g		sex3
sex4	byte	%10.0g		sex4
sex5	byte	%10.0g		sex5
sex6	byte	%10.0g		sex6
sex7	byte	%10.0g		sex7
sex	byte	%10.0g		sex

Sorted by:

```
.
. reshape long sex, i(family_id) j(pernum)
(j = 1 2 3 4 5 6 7)
variable sex already defined
r(110);
```

```
. rename sex sex8
```

```
. label variable sex8 "sex8"
```

```
. reshape long sex, i(family_id) j(pernum)
(j = 1 2 3 4 5 6 7 8)
```

Data	Wide	->	Long
Number of observations	2	->	16
Number of variables	9	->	3
j variable (8 values)		->	pernum
xij variables:	sex1 sex2 ... sex8	->	sex

```
.
. *****
. * Possible Error #3: Variable names do not contain numerical values
. *****
.
.
.
. use "D:\Jason\workshop\long format data\reshape3.dta", clear
.
. des
```

Contains data from D:\Jason\workshop\long format data\reshape3.dta

Observations: 2
 Variables: 9 6 Nov 2023 11:00

Variable name	Storage type	Display format	Value label	Variable label
family_id	int	%10.0g		family_id
sex_a	byte	%10.0g		sex_a
sex_b	byte	%10.0g		sex_b
sex_c	byte	%10.0g		sex_c
sex_d	byte	%10.0g		sex_d
sex_e	byte	%10.0g		sex_e
sex_f	byte	%10.0g		sex_f
sex_g	byte	%10.0g		sex_g
sex_h	byte	%10.0g		sex_h

Sorted by:

```
. reshape long sex_, i(family_id) j(pernum)
variable pernum contains all missing values
r(498);
```

```
. rename sex_a sex_1
```

```
. rename sex_b sex_2
```

```
. rename sex_c sex_3
```

```
. rename sex_d sex_4
```

```
. rename sex_e sex_5
```

```
. rename sex_f sex_6
```

```
. rename sex_g sex_7
```

```
. rename sex_h sex_8
```

```
. label variable sex_1 "sex_1"
```

```
. label variable sex_2 "sex_2"
```

```
. label variable sex_3 "sex_3"
```

```
. label variable sex_4 "sex_4"
```

```
. label variable sex_5 "sex_5"
```

```
. label variable sex_6 "sex_6"
```

```
. label variable sex_7 "sex_7"
```

```
. label variable sex_8 "sex_8"
```

```
. reshape long sex_, i(family_id) j(pernum)
(j = 1 2 3 4 5 6 7 8)
```

```
Data                                Wide  ->  Long
-----
Number of observations                2  ->  16
Number of variables                   9  ->   3
j variable (8 values)                 -> pernum
xij variables:
      sex_1 sex_2 ... sex_8  ->  sex_
```

```
. rename sex_ sex
```

```
. *****
```

```
. * Merge data in long format with data in wide format
```

```
. *****  
.   
.   
. *****
```

```
. * look at the data file
```

```
. *****
```

```
. use "D:\Jason\workshop\long format data\household_income.dta", clear
```

```
. list
```

```
+-----+  
| family~d  hoseho~e |  
+-----+  
1. |          9      4000 |  
2. |         8212      6000 |  
3. |         9999      8000 |  
+-----+
```

```
. use "D:\Jason\workshop\long format data\long format.dta", clear
```

```
. list
```

```
+-----+  
| family~d  pernum  sex |  
+-----+  
1. |          9        1   2 |  
2. |          9        2   2 |  
3. |         8212       1   1 |  
4. |         8212       2   2 |  
5. |         8212       3   1 |  
+-----+  
6. |         8212       4   1 |  
7. |         8212       5   2 |  
8. |         8212       6   2 |  
9. |         8212       7   1 |  
10. |         8212       8   2 |  
+-----+
```

```
. use "D:\Jason\workshop\long format data\wide format.dta", clear
```

```
. list
```

```
+-----+  
| family~d  sex1  sex2  sex3  sex4  sex5  sex6  sex7  sex8 |  
+-----+  
1. |          9    2    2    .    .    .    .    . |  
2. |         8212    1    2    1    1    2    2    1 |  
+-----+
```

```
. *****
```

```
. * Merging data: data in wide format to data in wide format
```

```
. *****
.
.
. use "D:\Jason\workshop\long format data\household_income.dta", clear
.
. merge 1:1 family_id using "D:\Jason\workshop\long format data\wide format.dta"
```

Result	Number of obs	
Not matched	1	
from master	1	(_merge==1)
from using	0	(_merge==2)
Matched	2	(_merge==3)

```
. sort family_id
.
. list, sepby(family_id)
```

	family~d	hoseho~e	sex1	sex2	sex3	sex4	sex5	sex6	sex7	sex8	_merge
1.	9	4000	2	2	Matched (3)
2.	8212	6000	1	2	1	1	2	2	1	2	Matched (3)
3.	9999	8000	Master only (1)

```
. save "D:\Jason\workshop\long format data\wide_wide.dta", replace
file D:\Jason\workshop\long format data\wide_wide.dta saved
```

```
. *****
.
. * Merging data: data in wide format to data in long format
```

```
. *****
.
. use "D:\Jason\workshop\long format data\household_income.dta", clear
.
. merge 1:m family_id using "D:\Jason\workshop\long format data\long format.dta"
```

Result	Number of obs	
Not matched	1	
from master	1	(_merge==1)
from using	0	(_merge==2)
Matched	10	(_merge==3)

```
. sort family_id pernum
.
. list, sepby(family_id)
```

	family~d	hoseho~e	pernum	sex	_merge
1.	9	4000	1	2	Matched (3)
2.	9	4000	2	2	Matched (3)
3.	8212	6000	1	1	Matched (3)
4.	8212	6000	2	2	Matched (3)
5.	8212	6000	3	1	Matched (3)

6.		8212	6000	4	1	Matched (3)	
7.		8212	6000	5	2	Matched (3)	
8.		8212	6000	6	2	Matched (3)	
9.		8212	6000	7	1	Matched (3)	
10.		8212	6000	8	2	Matched (3)	

11.		9999	8000	.	.	Master only (1)	

```
. save "D:\Jason\workshop\long format data\wide_long.dta", replace
file D:\Jason\workshop\long format data\wide_long.dta saved
```

```
.
.
.
.
.
.
.
. *****
```

```
. * Merging data: data in long format to data in wide format
```

```
. *****
```

```
. use "D:\Jason\workshop\long format data\long format.dta", clear
```

```
. merge m:1 family_id using "D:\Jason\workshop\long format data\household_income.dta"
```

Result	Number of obs	

Not matched	1	
from master	0	(_merge==1)
from using	1	(_merge==2)
Matched	10	(_merge==3)

```
. sort family_id pernum
```

```
. list, sepby(family_id)
```

		family~d	pernum	sex	hoseho~e	_merge	

1.		9	1	2	4000	Matched (3)	
2.		9	2	2	4000	Matched (3)	

3.		8212	1	1	6000	Matched (3)	
4.		8212	2	2	6000	Matched (3)	
5.		8212	3	1	6000	Matched (3)	
6.		8212	4	1	6000	Matched (3)	
7.		8212	5	2	6000	Matched (3)	
8.		8212	6	2	6000	Matched (3)	
9.		8212	7	1	6000	Matched (3)	
10.		8212	8	2	6000	Matched (3)	

11.		9999	.	.	8000	Using only (2)	

```
. save "D:\Jason\workshop\long format data\long_wide.dta", replace
file D:\Jason\workshop\long format data\long_wide.dta saved
```

```
.
.
.
.
.
.
.
.
.
.
.
. log close
name: <unnamed>
log: D:\Jason\workshop\long format data\reshape and merge2.log
log type: text
```


closed on: 6 Nov 2023, 11:08:34
