

```

name: <unnamed>
log: F:\workshop\long format\event_05_21_18.log
log type: text
opened on: 21 May 2018, 12:16:05

```

```
. use "F:\workshop\long format\command_long.dta", clear
```

```
. des
```

```

Contains data from F:\workshop\long format\command_long.dta
obs:          18
vars:          6          27 Feb 2017 09:13
size:         126

```

variable name	storage type	display format	value label	variable label
id	byte	%10.0g		id
year	int	%10.0g		year
month	byte	%10.0g		month
marst	byte	%10.0g		marst
sex	byte	%10.0g		sex
age	byte	%10.0g		age

```
Sorted by:
```

```
. list, sepby(id)
```

	id	year	month	marst	sex	age
1.	1	2010	1	6	1	35
2.	1	2010	2	6	1	35
3.	1	2010	3	6	1	35
4.	1	2010	4	1	1	36
5.	1	2010	5	.	1	36
6.	1	2010	6	1	1	36
7.	2	2010	1	6	2	30
8.	2	2010	2	6	2	30
9.	2	2010	3	1	2	30
10.	2	2010	4	1	2	30
11.	2	2010	5	4	2	30
12.	2	2010	6	9	2	30
13.	2	2010	7	9	2	31
14.	2	2010	8	9	2	31
15.	2	2010	9	1	2	31
16.	2	2010	10	1	2	31
17.	2	2010	11	1	2	31
18.	2	2010	12	1	2	31

```

. *****
. * Sort data
. *****

```

```
. sort id year month
```

```

. *****
. * Create an indicator variable for each timepoint within the person
. *****
. by id: gen time = _n

```

```

. label variable time "time points"

.
. *****
. *Calculate the total number of records per person
. *****
. by id: gen t_time = _N

. label variable t_time "total time points"

.
.
. list, sepby(id)

```

	id	year	month	marst	sex	age	time	t_time
1.	1	2010	1	6	1	35	1	6
2.	1	2010	2	6	1	35	2	6
3.	1	2010	3	6	1	35	3	6
4.	1	2010	4	1	1	36	4	6
5.	1	2010	5	.	1	36	5	6
6.	1	2010	6	1	1	36	6	6
7.	2	2010	1	6	2	30	1	12
8.	2	2010	2	6	2	30	2	12
9.	2	2010	3	1	2	30	3	12
10.	2	2010	4	1	2	30	4	12
11.	2	2010	5	4	2	30	5	12
12.	2	2010	6	9	2	30	6	12
13.	2	2010	7	9	2	31	7	12
14.	2	2010	8	9	2	31	8	12
15.	2	2010	9	1	2	31	9	12
16.	2	2010	10	1	2	31	10	12
17.	2	2010	11	1	2	31	11	12
18.	2	2010	12	1	2	31	12	12

```

. *****
. * Check whether each record is a unique one
. *****
.
. duplicates report id year month

```

Duplicates in terms of id year month

copies	observations	surplus
1	18	0

```

.
. *****
. * Reshape data: from long format to wide format
. *****
.
. reshape wide year month marst sex age t_time, i(id) j(time)
(note: j = 1 2 3 4 5 6 7 8 9 10 11 12)

```

Data	long	->	wide
Number of obs.	18	->	2
Number of variables	8	->	73
j variable (12 values)	time	->	(dropped)
xij variables:			
	year	->	year1 year2 ... year12
	month	->	month1 month2 ... month12


```
2. |-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   | 1         2         3         4         5         6         7         8         9         10
   |-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
.
```

```
. list marst* if id ==2
```

```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| marst1  marst2  marst3  marst4  marst5  marst6  marst7  marst8  marst9  marst10
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
2. | 6         6         1         1         4         9         9         9         1         1
   |-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
.
```

```
. list sex* if id ==2
```

```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| sex1    sex2    sex3    sex4    sex5    sex6    sex7    sex8    sex9    sex10   sex11   sex12
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
2. | 2         2         2         2         2         2         2         2         2         2         2         2
   |-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
.
```

```
. list age* if id ==2
```

```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| age1    age2    age3    age4    age5    age6    age7    age8    age9    age10   age11   age12
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
2. | 30        30        30        30        30        30        31        31        31        31        31        31
   |-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
.
```

```
. list t_time* if id ==2
```

```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| t_time1  t_time2  t_time3  t_time4  t_time5  t_time6  t_time7  t_time8  t_time9
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
2. | 12         12         12         12         12         12         12         12         12
   |-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
.
```

```

. *****
. * Reshape data: from long format to wide format
. *****
.
```

```
. reshape long year month marst sex age t_time, i(id) j(time)
(note: j = 1 2 3 4 5 6 7 8 9 10 11 12)
```

```

Data                wide  ->  long
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
Number of obs.      2      ->   24
Number of variables  73     ->    8
j variable (12 values) ->  time
xij variables:
    year1 year2 ... year12 ->  year
    month1 month2 ... month12 -> month
    marst1 marst2 ... marst12 -> marst
    sex1 sex2 ... sex12 ->  sex
    age1 age2 ... age12 ->  age
    t_time1 t_time2 ... t_time12 -> t_time
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
.
```

```
. list, sepby(id)
```

```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| id  time  year  month  marst  sex  age  t_time
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
1. | 1      1    2010     1      6     1    35     6
2. | 1      2    2010     2      6     1    35     6
3. | 1      3    2010     3      6     1    35     6
4. | 1      4    2010     4      1     1    36     6
5. | 1      5    2010     5      .     1    36     6
6. | 1      6    2010     6      1     1    36     6
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
.
```

7.	1	7
8.	1	8
9.	1	9
10.	1	10
11.	1	11
12.	1	12

13.	2	1	2010	1	6	2	30	12
14.	2	2	2010	2	6	2	30	12
15.	2	3	2010	3	1	2	30	12
16.	2	4	2010	4	1	2	30	12
17.	2	5	2010	5	4	2	30	12
18.	2	6	2010	6	9	2	30	12
19.	2	7	2010	7	9	2	31	12
20.	2	8	2010	8	9	2	31	12
21.	2	9	2010	9	1	2	31	12
22.	2	10	2010	10	1	2	31	12
23.	2	11	2010	11	1	2	31	12
24.	2	12	2010	12	1	2	31	12

```
. drop if t_time ==.
(6 observations deleted)
```

```
. list, sepby(id)
```

	id	time	year	month	marst	sex	age	t_time
1.	1	1	2010	1	6	1	35	6
2.	1	2	2010	2	6	1	35	6
3.	1	3	2010	3	6	1	35	6
4.	1	4	2010	4	1	1	36	6
5.	1	5	2010	5	.	1	36	6
6.	1	6	2010	6	1	1	36	6

7.	2	1	2010	1	6	2	30	12
8.	2	2	2010	2	6	2	30	12
9.	2	3	2010	3	1	2	30	12
10.	2	4	2010	4	1	2	30	12
11.	2	5	2010	5	4	2	30	12
12.	2	6	2010	6	9	2	30	12
13.	2	7	2010	7	9	2	31	12
14.	2	8	2010	8	9	2	31	12
15.	2	9	2010	9	1	2	31	12
16.	2	10	2010	10	1	2	31	12
17.	2	11	2010	11	1	2	31	12
18.	2	12	2010	12	1	2	31	12

```
.
.
.
. *****
. * sort: Arrange data in a special order
. *****
.
. sort id time
. list, sepby(id)
```

	id	time	year	month	marst	sex	age	t_time
1.	1	1	2010	1	6	1	35	6
2.	1	2	2010	2	6	1	35	6
3.	1	3	2010	3	6	1	35	6
4.	1	4	2010	4	1	1	36	6
5.	1	5	2010	5	.	1	36	6
6.	1	6	2010	6	1	1	36	6

```

7. | 2      1  2010      1      6      2  30      12 |
8. | 2      2  2010      2      6      2  30      12 |
9. | 2      3  2010      3      1      2  30      12 |
10. | 2      4  2010      4      1      2  30      12 |
11. | 2      5  2010      5      4      2  30      12 |
12. | 2      6  2010      6      9      2  30      12 |
13. | 2      7  2010      7      9      2  31      12 |
14. | 2      8  2010      8      9      2  31      12 |
15. | 2      9  2010      9      1      2  31      12 |
16. | 2     10  2010     10      1      2  31      12 |
17. | 2     11  2010     11      1      2  31      12 |
18. | 2     12  2010     12      1      2  31      12 |
+-----+

```

```

.
.
. *****
. * Handling the missing value
. *****
.
. list id time year month marst, sepby(id)

```

```

+-----+
| id  time  year  month  marst |
+-----+
1. | 1      1  2010      1      6 |
2. | 1      2  2010      2      6 |
3. | 1      3  2010      3      6 |
4. | 1      4  2010      4      1 |
5. | 1      5  2010      5      . |
6. | 1      6  2010      6      1 |
+-----+
7. | 2      1  2010      1      6 |
8. | 2      2  2010      2      6 |
9. | 2      3  2010      3      1 |
10. | 2      4  2010      4      1 |
11. | 2      5  2010      5      4 |
12. | 2      6  2010      6      9 |
13. | 2      7  2010      7      9 |
14. | 2      8  2010      8      9 |
15. | 2      9  2010      9      1 |
16. | 2     10  2010     10      1 |
17. | 2     11  2010     11      1 |
18. | 2     12  2010     12      1 |
+-----+

```

```

.
. gen marst_r = marst
(1 missing value generated)

. label variable marst_r "recoded marital status"

.
. replace marst_r = marst[_n-1] if time ~=1 & marst_r ==. & marst_r[_n-1] ~=.
(1 real change made)

.
. list id time year month marst marst_r, sepby(id)

```

```

+-----+
| id  time  year  month  marst  marst_r |
+-----+
1. | 1      1  2010      1      6      6 |
2. | 1      2  2010      2      6      6 |
3. | 1      3  2010      3      6      6 |
4. | 1      4  2010      4      1      1 |
5. | 1      5  2010      5      .      1 |
6. | 1      6  2010      6      1      1 |
+-----+
7. | 2      1  2010      1      6      6 |
8. | 2      2  2010      2      6      6 |

```

```

9. | 2      3  2010      3      1      1 |
10. | 2      4  2010      4      1      1 |
11. | 2      5  2010      5      4      4 |
12. | 2      6  2010      6      9      9 |
13. | 2      7  2010      7      9      9 |
14. | 2      8  2010      8      9      9 |
15. | 2      9  2010      9      1      1 |
16. | 2     10  2010     10      1      1 |
17. | 2     11  2010     11      1      1 |
18. | 2     12  2010     12      1      1 |
+-----+

```

```

.
.
. *****
. * Code the transition into marriage
. *****
.
. gen c_mar =0
.
. label variable c_mar "change into marriage"
.
. by id: replace c_mar = 1 if marst_r[_n-1] >=4 & (marst_r ==1 | marst_r ==2| marst_r ==3 )
(3 real changes made)
.
. list id time year month marst_r c_mar, sepby(id)

```

```

+-----+
| id   time   year   month   marst_r   c_mar |
+-----+
1. | 1     1     2010     1       6       0 |
2. | 1     2     2010     2       6       0 |
3. | 1     3     2010     3       6       0 |
4. | 1     4     2010     4       1       1 |
5. | 1     5     2010     5       1       0 |
6. | 1     6     2010     6       1       0 |
+-----+
7. | 2     1     2010     1       6       0 |
8. | 2     2     2010     2       6       0 |
9. | 2     3     2010     3       1       1 |
10. | 2     4     2010     4       1       0 |
11. | 2     5     2010     5       4       0 |
12. | 2     6     2010     6       9       0 |
13. | 2     7     2010     7       9       0 |
14. | 2     8     2010     8       9       0 |
15. | 2     9     2010     9       1       1 |
16. | 2    10     2010    10       1       0 |
17. | 2    11     2010    11       1       0 |
18. | 2    12     2010    12       1       0 |
+-----+

```

```

.
.
. *****
. *Create an indicator for the times that respondents entered marriage
. *****
.
. by id: gen i_c_mar = sum(c_mar)
.
. label variable i_c_mar "indicator for the times of entering marriage"
.
. *****
. * Create an indicator for the total number of times of entering marriage"
. *****
.
. by id: egen s_c_mar = sum(c_mar)

```

```
. label variable s_c_mar "indicator for total number of times of entering marriage"
```

```
.  
.  
list id time year month marst_r c_mar i_c_mar s_c_mar, sepby(id)
```

	id	time	year	month	marst_r	c_mar	i_c_mar	s_c_mar
1.	1	1	2010	1	6	0	0	1
2.	1	2	2010	2	6	0	0	1
3.	1	3	2010	3	6	0	0	1
4.	1	4	2010	4	1	1	1	1
5.	1	5	2010	5	1	0	1	1
6.	1	6	2010	6	1	0	1	1
7.	2	1	2010	1	6	0	0	2
8.	2	2	2010	2	6	0	0	2
9.	2	3	2010	3	1	1	1	2
10.	2	4	2010	4	1	0	1	2
11.	2	5	2010	5	4	0	1	2
12.	2	6	2010	6	9	0	1	2
13.	2	7	2010	7	9	0	1	2
14.	2	8	2010	8	9	0	1	2
15.	2	9	2010	9	1	1	2	2
16.	2	10	2010	10	1	0	2	2
17.	2	11	2010	11	1	0	2	2
18.	2	12	2010	12	1	0	2	2

```
.  
.  
.  
*****  
*Extract the time when the first marriage took place  
*****  
.  
by id: gen time_mar1 = time if c_mar ==1 & i_c_mar ==1  
(16 missing values generated)
```

```
. label variable time_mar1 "time for the first marriage"
```

```
. list id time year month marst_r c_mar i_c_mar s_c_mar time_mar1, sepby(id)
```

	id	time	year	month	marst_r	c_mar	i_c_mar	s_c_mar	time_m~1
1.	1	1	2010	1	6	0	0	1	.
2.	1	2	2010	2	6	0	0	1	.
3.	1	3	2010	3	6	0	0	1	.
4.	1	4	2010	4	1	1	1	1	4
5.	1	5	2010	5	1	0	1	1	.
6.	1	6	2010	6	1	0	1	1	.
7.	2	1	2010	1	6	0	0	2	.
8.	2	2	2010	2	6	0	0	2	.
9.	2	3	2010	3	1	1	1	2	3
10.	2	4	2010	4	1	0	1	2	.
11.	2	5	2010	5	4	0	1	2	.
12.	2	6	2010	6	9	0	1	2	.
13.	2	7	2010	7	9	0	1	2	.
14.	2	8	2010	8	9	0	1	2	.
15.	2	9	2010	9	1	1	2	2	.
16.	2	10	2010	10	1	0	2	2	.
17.	2	11	2010	11	1	0	2	2	.
18.	2	12	2010	12	1	0	2	2	.


```

.
.
. *****
. * Expand the time of the first transition into marriage for all records of the individual
. *****
.
. by id: egen m_time_mar1 = max(time_mar1)
.
. label variable m_time_mar1 "time for the first marriage"
.
. list id time year month marst_r c_mar i_c_mar s_c_mar time_mar1 m_time_mar1, sepby(id)

```

	id	time	year	month	marst_r	c_mar	i_c_mar	s_c_mar	time_m~1	m_time~1
1.	1	1	2010	1	6	0	0	1	.	4
2.	1	2	2010	2	6	0	0	1	.	4
3.	1	3	2010	3	6	0	0	1	.	4
4.	1	4	2010	4	1	1	1	1	4	4
5.	1	5	2010	5	1	0	1	1	.	4
6.	1	6	2010	6	1	0	1	1	.	4
7.	2	1	2010	1	6	0	0	2	.	3
8.	2	2	2010	2	6	0	0	2	.	3
9.	2	3	2010	3	1	1	1	2	3	3
10.	2	4	2010	4	1	0	1	2	.	3
11.	2	5	2010	5	4	0	1	2	.	3
12.	2	6	2010	6	9	0	1	2	.	3
13.	2	7	2010	7	9	0	1	2	.	3
14.	2	8	2010	8	9	0	1	2	.	3
15.	2	9	2010	9	1	1	2	2	.	3
16.	2	10	2010	10	1	0	2	2	.	3
17.	2	11	2010	11	1	0	2	2	.	3
18.	2	12	2010	12	1	0	2	2	.	3

```

.
.
. *****
. * Remove records that occurred after the first transition into marriage
. *****
.
. list id time year month marst_r c_mar i_c_mar s_c_mar time_mar1 m_time_mar1, sepby(id)

```

	id	time	year	month	marst_r	c_mar	i_c_mar	s_c_mar	time_m~1	m_time~1
1.	1	1	2010	1	6	0	0	1	.	4
2.	1	2	2010	2	6	0	0	1	.	4
3.	1	3	2010	3	6	0	0	1	.	4
4.	1	4	2010	4	1	1	1	1	4	4
5.	1	5	2010	5	1	0	1	1	.	4
6.	1	6	2010	6	1	0	1	1	.	4
7.	2	1	2010	1	6	0	0	2	.	3
8.	2	2	2010	2	6	0	0	2	.	3
9.	2	3	2010	3	1	1	1	2	3	3
10.	2	4	2010	4	1	0	1	2	.	3
11.	2	5	2010	5	4	0	1	2	.	3
12.	2	6	2010	6	9	0	1	2	.	3
13.	2	7	2010	7	9	0	1	2	.	3
14.	2	8	2010	8	9	0	1	2	.	3
15.	2	9	2010	9	1	1	2	2	.	3
16.	2	10	2010	10	1	0	2	2	.	3
17.	2	11	2010	11	1	0	2	2	.	3
18.	2	12	2010	12	1	0	2	2	.	3

```
. by id: drop if time > m_time_mar1
(11 observations deleted)
```

```
. list id time year month marst_r c_mar i_c_mar s_c_mar time_mar1 m_time_mar1, sepby(id)
```

	id	time	year	month	marst_r	c_mar	i_c_mar	s_c_mar	time_m~1	m_time~1
1.	1	1	2010	1	6	0	0	1	.	4
2.	1	2	2010	2	6	0	0	1	.	4
3.	1	3	2010	3	6	0	0	1	.	4
4.	1	4	2010	4	1	1	1	1	4	4
5.	2	1	2010	1	6	0	0	2	.	3
6.	2	2	2010	2	6	0	0	2	.	3
7.	2	3	2010	3	1	1	1	2	3	3

```
. *****
. * Save the event history data
. *****
```

```
. list id time year month marst_r sex age, sepby(id)
```

	id	time	year	month	marst_r	sex	age
1.	1	1	2010	1	6	1	35
2.	1	2	2010	2	6	1	35
3.	1	3	2010	3	6	1	35
4.	1	4	2010	4	1	1	36
5.	2	1	2010	1	6	2	30
6.	2	2	2010	2	6	2	30
7.	2	3	2010	3	1	2	30

```
. save "F:\workshop\long format\event.dta", replace
file F:\workshop\long format\event.dta saved
```

```
. log close
name: <unnamed>
log: F:\workshop\long format\event_05_21_18.log
log type: text
closed on: 21 May 2018, 12:16:05
```