```
------------------------------------------------------------------------------------------------------------------------
      name:  <unnamed>
       log:  k:\family workshop\stata.log
  log type:  text
 opened on:  24 Jun 2013, 15:53:25

.
. ****************************
. * The workshop data
. ****************************
.
. use C:\temp\family_workshop.dta, clear

.
. ************************
. * 1. The size of a family
. ************************
. use C:\temp\family_workshop.dta, clear

. sort family_id

. by family_id: gen family_N = _N  /* use the variable, family_N, to indicate the total number of people in family *
> /

. list family_id epppnum person_id family_N

     +-----------------------------------------+
     | family~d   epppnum   person~d   family_N |
     |-----------------------------------------|
  1. |        1       101       1001          2 |
  2. |        1       102       1002          2 |
  3. |        2       101       1003          3 |
  4. |        2       102       1004          3 |
  5. |        2       103       1005          3 |
     |-----------------------------------------|
  6. |        3       101       1006          4 |
  7. |        3       102       1007          4 |
  8. |        3       103       1008          4 |
  9. |        3       104       1009          4 |
     +-----------------------------------------+

.
. ************************
. * 2. The number of kids in the family
. ************************
. sort family_id

. gen kid = 1 if age < 18 /* Note: the variable, kid, indicates whether a family member is a kid or not */
(6 missing values generated)

.
. by family_id: egen t_kid =sum(kid) /* Note: the variable, t_kid, indicates the total number of kids within the fam
> ily */
```

```
. list family_id epppnum person_id age epnmom etypmom epndad etypdad kid t_kid

     +-------------------------------------------------------------------------------------------+
     | family~d   epppnum   person~d   age   epnmom   etypmom   epndad   etypdad   kid   t_kid |
     |-------------------------------------------------------------------------------------------|
  1. |        1       101       1001    23        .         .        .         .     .       0 |
  2. |        1       102       1002    23        .         .        .         .     .       0 |
  3. |        2       101       1003    29        .         .        .         .     .       1 |
  4. |        2       102       1004    32        .         .        .         .     .       1 |
  5. |        2       103       1005     6      102         1      101         1     1       1 |
     |-------------------------------------------------------------------------------------------|
  6. |        3       101       1006    51        .         .        .         .     .       2 |
  7. |        3       102       1007    53        .         .        .         .     .       2 |
  8. |        3       103       1008    15      101         1      102         2     1       2 |
  9. |        3       104       1009     3      101         1      102         1     1       2 |
     +-------------------------------------------------------------------------------------------+
.
.
. **************************
. * 3. The total family income
. **************************
. sort family_id

. by family_id:  egen t_income =sum(income) /* Note: the variable, t_income, indicates the total income of the famil
> y */

. list family_id epppnum person_id income t_income

     +-------------------------------------------------+
     | family~d   epppnum   person~d   income   t_income |
     |-------------------------------------------------|
  1. |        1       101       1001     3000       3000 |
  2. |        1       102       1002        .       3000 |
  3. |        2       101       1003        .       5000 |
  4. |        2       102       1004     5000       5000 |
  5. |        2       103       1005        .       5000 |
     |-------------------------------------------------|
  6. |        3       101       1006     2000       5000 |
  7. |        3       102       1007     3000       5000 |
  8. |        3       103       1008        .       5000 |
  9. |        3       104       1009        .       5000 |
     +-------------------------------------------------+


.
.
. ****************************************
. * 4. The ID of the respondent's spouse?
. ****************************************
.
.         ****************************************************************
.         * Generate the indicator for husbands and wives within the family
.         ****************************************************************
```

```
.
. gen epnwife = epnspous if sex ==1 & epnspous ~=.
(6 missing values generated)

. gen epnhus =  epnspous if sex ==2 & epnspous ~=.
(6 missing values generated)

.
.
.         ****************************************
.         * Expand the indicators to the whole family
.         ****************************************
.
. sort family_id

. by family_id: egen epnwife2 =max(epnwife)

. by family_id: egen epnhus2 = max(epnhus)

.
.
.         ****************************************************
.         * Extract the IDs for husband and wife in the family
.         ****************************************************
.
. gen id_wife = person_id if epppnum == epnwife2
(6 missing values generated)

. gen id_hus  = person_id if epppnum == epnhus2
(6 missing values generated)

.
.         ****************************************
.         * Expand the IDs for husband and wive to the whole family
.         ****************************************
.
. sort family_id

. by family_id: egen id_wife2 =max(id_wife)

. by family_id: egen id_hus2 = max(id_hus)

.
.
.         ************************************
.         * Remove unnecessary information
.         ************************************
.
. replace id_wife2 =. if epnspous ==.
(3 real changes made, 3 to missing)

. replace id_wife2 =. if epnspous ~=.   & sex==2
(3 real changes made, 3 to missing)
```

```
. replace id_hus2  =. if epnspous ==.
(3 real changes made, 3 to missing)

. replace id_hus2  =. if epnspous ~=.   & sex ==1
(3 real changes made, 3 to missing)

. 
. 
. list family_id epppnum person_id sex epnspous id_wife2 id_hus2

     +-------------------------------------------------------------------------+
     | family~d   epppnum   person~d   sex   epnspous   id_wife2   id_hus2 |
     |-------------------------------------------------------------------------|
  1. |        1       101       1001     1        102       1002         . |
  2. |        1       102       1002     2        101          .      1001 |
  3. |        2       101       1003     1        102       1004         . |
  4. |        2       102       1004     2        101          .      1003 |
  5. |        2       103       1005     1          .          .         . |
     |-------------------------------------------------------------------------|
  6. |        3       101       1006     2        102          .      1007 |
  7. |        3       102       1007     1        101       1006         . |
  8. |        3       103       1008     2          .          .         . |
  9. |        3       104       1009     1          .          .         . |
     +-------------------------------------------------------------------------+



. 
. ********************************************
. * 5. How many biological kids do respondents have?
. ********************************************
. 
.         ************************************************************************
.         * count the possible number of biological kids for a father in the family
.         ************************************************************************
. 
. gen bio_kid =1 if etypdad ==1
(7 missing values generated)

. sort family_id

. by family_id: egen t_bio_kid = sum(bio_kid)

. 
. 
. list family_id epppnum person_id sex epndad etypdad t_bio_kid

     +-----------------------------------------------------------------+
     | family~d   epppnum   person~d   sex   epndad   etypdad   t_bio_~d |
     |-----------------------------------------------------------------|
  1. |        1       101       1001     1        .         .          0 |
  2. |        1       102       1002     2        .         .          0 |
  3. |        2       101       1003     1        .         .          1 |
  4. |        2       102       1004     2        .         .          1 |
```

```
  5. |        2       103       1005     1      101         1          1 |
     |-----------------------------------------------------------------|
  6. |        3       101       1006     2        .         .          1 |
  7. |        3       102       1007     1        .         .          1 |
  8. |        3       103       1008     2      102         2          1 |
  9. |        3       104       1009     1      102         1          1 |
     +-----------------------------------------------------------------+


.
.          ********************************************************************
.          * Extract the ID of biological father's from the kid's report
.          ********************************************************************
.
.
. gen bio_dad_epn = epndad if bio_kid ==1
(7 missing values generated)

.
. list family_id epppnum person_id sex epndad etypdad bio_kid t_bio_kid bio_dad_epn

     +-------------------------------------------------------------------------------------------+
     | family~d   epppnum   person~d   sex   epndad   etypdad   bio_kid   t_bio_~d   bio_da~n |
     |-------------------------------------------------------------------------------------------|
  1. |        1       101       1001     1        .         .         .          0          . |
  2. |        1       102       1002     2        .         .         .          0          . |
  3. |        2       101       1003     1        .         .         .          1          . |
  4. |        2       102       1004     2        .         .         .          1          . |
  5. |        2       103       1005     1      101         1         1          1        101 |
     |-------------------------------------------------------------------------------------------|
  6. |        3       101       1006     2        .         .         .          1          . |
  7. |        3       102       1007     1        .         .         .          1          . |
  8. |        3       103       1008     2      102         2         .          1          . |
  9. |        3       104       1009     1      102         1         1          1        102 |
     +-------------------------------------------------------------------------------------------+


.
.
.          ************************************************************
.          * Expand the ID of biological father's to the whole family
.          ************************************************************
.
. sort family_id

. by family_id: egen m_bio_dad_epn = max(bio_dad_epn)
(2 missing values generated)

.
. list family_id epppnum person_id sex epndad etypdad bio_kid t_bio_kid bio_dad_epn m_bio_dad_epn

     +-------------------------------------------------------------------------------------------------+
     | family~d   epppnum   person~d   sex   epndad   etypdad   bio_kid   t_bio_~d   bio_da~n   m_bio_~n |
     |-------------------------------------------------------------------------------------------------|
  1. |        1       101       1001     1        .         .         .          0          .          . |
  2. |        1       102       1002     2        .         .         .          0          .          . |
  3. |        2       101       1003     1        .         .         .          1          .        101 |
  4. |        2       102       1004     2        .         .         .          1          .        101 |
```

```
  5. |        2       103       1005     1      101         1         1          1        101        101 |
     |-------------------------------------------------------------------------------------------------|
  6. |        3       101       1006     2        .         .         .          1          .        102 |
  7. |        3       102       1007     1        .         .         .          1          .        102 |
  8. |        3       103       1008     2      102         2         .          1          .        102 |
  9. |        3       104       1009     1      102         1         1          1        102        102 |
     +-------------------------------------------------------------------------------------------------+


.
.          ****************************************************
.          * Create an indicator of whether adult's epppnum correspond to the ID
.          * number that children reported for their biological father.
.          ****************************************************8***
.
. gen nbio_kid = 0

. replace nbio_kid = 1 if epppnum == m_bio_dad_epn & sex ==1
(2 real changes made)

. label variable nbio_kid "number of biological kids for the father"

.
. list family_id epppnum person_id sex epndad etypdad kid m_bio_dad_epn nbio_kid, sepby(family_id)

     +-------------------------------------------------------------------------------+
     | family~d   epppnum   person~d   sex   epndad   etypdad   kid   m_bio_~n   nbio_kid |
     |-------------------------------------------------------------------------------|
  1. |        1       101       1001     1        .         .     .          .          0 |
  2. |        1       102       1002     2        .         .     .          .          0 |
     |-------------------------------------------------------------------------------|
  3. |        2       101       1003     1        .         .     .        101          1 |
  4. |        2       102       1004     2        .         .     .        101          0 |
  5. |        2       103       1005     1      101         1     1        101          0 |
     |-------------------------------------------------------------------------------|
  6. |        3       101       1006     2        .         .     .        102          0 |
  7. |        3       102       1007     1        .         .     .        102          1 |
  8. |        3       103       1008     2      102         2     1        102          0 |
  9. |        3       104       1009     1      102         1     1        102          0 |
     +-------------------------------------------------------------------------------+


.
. log close
      name:  <unnamed>
       log:  k:\family workshop\stata.log
  log type:  text
 closed on:  24 Jun 2013, 15:53:26
-------------------------------------------------------------------------------------------------------------------
```