The 500 Family Study data

Zhengguo Gu

The original raw data – "500 Family Study" data – is publicly available at http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/4549?searchSource=revise&q=500+family+study. To obtain the data used in this article, follow the following steps:

Step 1. Download data

Download the DS3 and DS4 raw data (SPSS format) from the website. Use SPSS software to save the raw data as comma separated values file.

Step 2. Process data

A function for replacing missing values in the raw data with NA

We first write a simple function for dealing with the coding missing value in the raw data. All missing values are replaced with the mean

```
sim_impu <- function(DATA, M_values, repl){
  # M_values: missing values in the raw data. Some missing values are coded as 9, some are coded as 999
  # repl = T, missing values are replaced with mean.

  if(missing(repl)){
    repl = F
  }

  nrow <- dim(DATA)[1]
  ncol <- dim(DATA)[2]
  DATA <- data.matrix(DATA)

  for(i in 1:length(M_values)){
    DATA[which(DATA == M_values[i])] <- NA
  }

  if(repl == T){

    for(n in 1:nrow){
    for(m in 1:ncol){
      if(is.na(DATA[n, m])){
          DATA[n, m] <- mean(DATA[n, ], na.rm = T)
      }
    }
   }
  }


 return(DATA)
}
```

Load raw data

```
#DS3 = read.csv("D:\\Dropbox\\Dropbox\\tilburg office\\Research SCA\\Project 2 software Simultaneous\\n
#DS4 = read.csv("D:\\Dropbox\\Dropbox\\tilburg office\\Research SCA\\Project 2 software Simultaneous\\n

DS3 = read.csv("D:/Dropbox/Tilburg office/Research SCA/Project 2 software Simultaneous/newdata/04549-00
DS4 = read.csv("D:/Dropbox/Tilburg office/Research SCA/Project 2 software Simultaneous/newdata/04549-00
```

Process DS3 data

Questionnaire #74: Please tell us how much agree with each of the following statements about your relationship with your spouse/parter.

Example question: 74a. My partner and I understand each other perfectly. (strongly disagree to strongly agree).

The following questions in this questionnaire are to be reverse coded:

- b
- e
- h
- i
- l
- n

```
DS3[, c(352:367)] <- sim_impu(DS3[, c(352:367)], 9)
DS3[, c(353, 356, 359, 360, 363, 365)] <- 6 - DS3[, c(353, 356, 359, 360, 363, 365)] #reverse coding

DS3[, c(352:367)] <- sim_impu(DS3[, c(352:367)], 9, repl = T) #NA replaced with mean
summary(DS3[, c(352:367)])
```

```
##     UNDPAR_1        HABPAR_1        RESPAR_1       MUDPAR_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.00   Min.   :1.000
##  1st Qu.:3.000   1st Qu.:3.000   1st Qu.:4.00   1st Qu.:2.000
##  Median :4.000   Median :4.000   Median :4.00   Median :3.000
##  Mean   :3.606   Mean   :3.912   Mean   :3.85   Mean   :2.895
##  3rd Qu.:4.000   3rd Qu.:5.000   3rd Qu.:5.00   3rd Qu.:4.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.00   Max.   :5.000
##  NA's   :96      NA's   :96      NA's   :96     NA's   :96
##     CMTPAR_1        SUCREL_1        CONFLC_1        FINPOS_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:3.000   1st Qu.:2.000   1st Qu.:3.000   1st Qu.:2.000
##  Median :4.000   Median :4.000   Median :4.000   Median :4.000
##  Mean   :3.806   Mean   :3.275   Mean   :3.689   Mean   :3.592
##  3rd Qu.:5.000   3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:5.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
##  NA's   :96      NA's   :96      NA's   :96      NA's   :96
##     HNEEDS_1        LEISUR_1        HAPSEX_1        PARENT_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000   1st Qu.:3.000
##  Median :2.000   Median :4.000   Median :4.000   Median :4.000
##  Mean   :2.873   Mean   :3.374   Mean   :3.377   Mean   :3.976
##  3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:5.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
```

```
##  NA's   :96      NA's   :96      NA's   :96      NA's   :96
##     NORGRT_1        OTHRLS_1        RELPRC_1        OVERAL_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:2.000   1st Qu.:3.000   1st Qu.:3.000   1st Qu.:4.000
##  Median :4.000   Median :4.000   Median :4.000   Median :5.000
##  Mean   :3.309   Mean   :3.905   Mean   :3.769   Mean   :4.318
##  3rd Qu.:5.000   3rd Qu.:5.000   3rd Qu.:5.000   3rd Qu.:5.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
##  NA's   :96      NA's   :96      NA's   :96      NA's   :96
```

Note: 96 rows in the data matrix above contain NaN vectors, that is, the entire row is NA.

After reverse coding, **the higher a score, the happier/more satisfied a person in his/her relationship**.

Questionnaire #75: There are various ways that couples deal with serious disagreements. When you have a serious disagreement with your spouse/partner, how often do you. . .

Example question: 75a. Just keep your opinions to yourself (Never to very often).

The following question is to be reverse coded:

- c
- d

```
DS3[, c(368:371)] <- sim_impu(DS3[, c(368:371)], 9)
DS3[, c(370, 371)] <- 6 - DS3[, c(370, 371)] #so the higher the value the calmer.
DS3[, c(368:371)] <- sim_impu(DS3[, c(368:371)], 9, repl = T)
summary(DS3[, c(368:371)])
```

```
##     DONTLK_1        TALKNG_1       SCREAM_1        HITOTH_1
##  Min.   :1.000   Min.   :1.00   Min.   :1.000   Min.   :3.000
##  1st Qu.:2.000   1st Qu.:3.00   1st Qu.:3.000   1st Qu.:5.000
##  Median :3.000   Median :3.00   Median :4.000   Median :5.000
##  Mean   :2.507   Mean   :3.53   Mean   :3.641   Mean   :4.941
##  3rd Qu.:3.000   3rd Qu.:4.00   3rd Qu.:4.000   3rd Qu.:5.000
##  Max.   :5.000   Max.   :5.00   Max.   :5.000   Max.   :5.000
##  NA's   :96      NA's   :96     NA's   :96      NA's   :96
```

Note: 96 rows in the data matrix above contain NaN vectors, that is, the entire row is NA.

After reverse coding, **the higher a score, the less violent during a disagreement with the spouse/partner**.

Questionnaire #80: What do you think the chances are that the child participating in the study will..

Example question: a. Graduate from high school (very low to very high).

```
DS3[, c(379:386)] <- sim_impu(DS3[, c(379:386)], c(9, -8), repl = T)
summary(DS3[, c(379:386)])
```

```
##     CGRDHS_1        CGRDCO_1        CMARRY_1        CCHILD_1
##  Min.   :2.000   Min.   :1.000   Min.   :2.000   Min.   :2.000
##  1st Qu.:5.000   1st Qu.:5.000   1st Qu.:4.000   1st Qu.:4.000
##  Median :5.000   Median :5.000   Median :5.000   Median :5.000
##  Mean   :4.941   Mean   :4.711   Mean   :4.446   Mean   :4.374
```

```
##  3rd Qu.:5.000   3rd Qu.:5.000   3rd Qu.:5.000   3rd Qu.:5.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
##  NA's   :31      NA's   :31      NA's   :31      NA's   :31
##     COWNHM_1        CJOYWK_1        CJOYHM_1        CSTBMR_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:4.000   1st Qu.:4.000   1st Qu.:4.000   1st Qu.:4.000
##  Median :5.000   Median :4.000   Median :4.000   Median :4.000
##  Mean   :4.475   Mean   :4.326   Mean   :4.331   Mean   :4.107
##  3rd Qu.:5.000   3rd Qu.:5.000   3rd Qu.:5.000   3rd Qu.:5.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
##  NA's   :31      NA's   :31      NA's   :31      NA's   :31
```

Note: 31 rows in the data matrix above contain NaN vectors, that is, the entire row is NA.

No reverse coding needed. **The higher the score, the stronger the parent feels that the child has a bright future/happy life**.

Questionnaire #81: How often do you and the child in the study participate in the following activities together:

Example question: a. Talking about everyday events in your lives.

```
DS3[, c(387:404)] <- sim_impu(DS3[, c(387:404)], c(9, -8), repl = T)
summary(DS3[, c(387:404)])
```

```
##     TKEVEN_1        TKISSU_1        TKNEWS_1        COOKTG_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:4.000   1st Qu.:2.000   1st Qu.:2.000   1st Qu.:1.000
##  Median :4.000   Median :3.000   Median :3.000   Median :2.000
##  Mean   :3.786   Mean   :2.833   Mean   :2.629   Mean   :2.241
##  3rd Qu.:4.000   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000
##  Max.   :4.000   Max.   :4.000   Max.   :4.000   Max.   :4.000
##  NA's   :33      NA's   :33      NA's   :33      NA's   :33
##     HMWKTG_1        EATMTG_1        WATVTG_1        SHOPTG_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:2.000   1st Qu.:4.000   1st Qu.:2.000   1st Qu.:2.000
##  Median :2.000   Median :4.000   Median :3.000   Median :2.000
##  Mean   :2.287   Mean   :3.736   Mean   :2.844   Mean   :2.194
##  3rd Qu.:3.000   3rd Qu.:4.000   3rd Qu.:3.000   3rd Qu.:3.000
##  Max.   :4.000   Max.   :4.000   Max.   :4.000   Max.   :4.000
##  NA's   :33      NA's   :33      NA's   :33      NA's   :33
##     SPRTTG_1        MUSCTG_1        ARTCTG_1        VOLUTG_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.00
##  1st Qu.:1.000   1st Qu.:2.000   1st Qu.:1.000   1st Qu.:1.00
##  Median :2.000   Median :2.000   Median :1.000   Median :1.00
##  Mean   :2.067   Mean   :2.394   Mean   :1.699   Mean   :1.29
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:2.000   3rd Qu.:2.00
##  Max.   :4.000   Max.   :4.000   Max.   :4.000   Max.   :4.00
##  NA's   :33      NA's   :33      NA's   :33      NA's   :33
##     RELGTG_1       HOBYTG_1        READTG_1        PFRDTG_1
##  Min.   :1.00   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:1.00   1st Qu.:1.000   1st Qu.:2.143   1st Qu.:2.000
##  Median :2.00   Median :2.000   Median :2.571   Median :2.357
##  Mean   :1.89   Mean   :1.894   Mean   :2.777   Mean   :2.357
##  3rd Qu.:3.00   3rd Qu.:2.615   3rd Qu.:3.714   3rd Qu.:2.786
```

```
##  Max.   :4.00   Max.   :4.000   Max.   :4.000   Max.   :4.000
##  NA's   :33     NA's   :33      NA's   :33      NA's   :33
##     PLAYTG_1        PARKTG_1
##  Min.   :1.000   Min.   :1.000
##  1st Qu.:2.071   1st Qu.:2.000
##  Median :2.500   Median :2.357
##  Mean   :2.646   Mean   :2.427
##  3rd Qu.:3.000   3rd Qu.:2.857
##  Max.   :4.000   Max.   :4.000
##  NA's   :33      NA's   :33
```

Note: 33 rows in the data matrix above contain NaN vectors, that is, the entire row is NA.

No reverse coding needed. **The higher the score, the more often the parent does activities together with the child**.

Questionnaire #85: The statements below describe lots of events that routinely occur in families with young children. These events sometimes make life diffcult. Please read each item and circle how much of a 'hassle' you feel this is for you. If you have more than one child, these events can include any or all of your children.

Example question: a. Continually clearning up messess of toys or food.

```
DS3[, c(441:460)] <- sim_impu(DS3[, c(441:460)], c(9, -8), repl = T)
summary(DS3[, c(441:460)])
```

```
##     CMESSE_1        CNAGGD_1        CPICKY_1        CNOLST_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:3.000   1st Qu.:3.000   1st Qu.:2.000   1st Qu.:3.000
##  Median :3.000   Median :4.000   Median :3.000   Median :3.000
##  Mean   :3.424   Mean   :3.575   Mean   :2.914   Mean   :3.454
##  3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:4.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
##  NA's   :565     NA's   :565     NA's   :565     NA's   :565
##     XBABST_1        CSCHED_1        CARGUE_1        CPLAYW_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:1.000   1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000
##  Median :2.000   Median :2.000   Median :3.000   Median :2.000
##  Mean   :2.434   Mean   :2.272   Mean   :2.872   Mean   :2.462
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:4.000   3rd Qu.:3.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
##  NA's   :565     NA's   :565     NA's   :565     NA's   :565
##     CBEDTM_1        CEVRYW_1        CWATCH_1        CINTRU_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:2.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:2.000
##  Median :2.000   Median :2.000   Median :2.000   Median :3.000
##  Mean   :2.594   Mean   :1.999   Mean   :2.048   Mean   :2.694
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
##  NA's   :565     NA's   :565     NA's   :565     NA's   :565
##     CNEEDS_1        CDIRTY_1        CPRVCY_1        CPUBLC_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000
##  Median :2.000   Median :1.000   Median :2.000   Median :2.000
##  Mean   :1.949   Mean   :1.534   Mean   :2.046   Mean   :2.064
```

```
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:3.000   3rd Qu.:3.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
##  NA's   :565     NA's   :565     NA's   :565     NA's   :565
##     CREADY_1        CLEAVE_1        CPBFRN_1        CERRND_1
##  Min.   :1.00   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:2.00   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000
##  Median :3.00   Median :1.000   Median :1.000   Median :2.000
##  Mean   :2.84   Mean   :1.697   Mean   :1.716   Mean   :1.987
##  3rd Qu.:4.00   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:3.000
##  Max.   :5.00   Max.   :5.000   Max.   :5.000   Max.   :5.000
##  NA's   :565    NA's   :565     NA's   :565     NA's   :565
```

Note: 565 rows in the data matrix above contain NaN vectors, that is, the entire row is NA. We should consider remove this questionnaire.

No reverse coding needed.

Questionnaire #87: Now we're going to present a few more statements about parenting. How true do you feel each of the following statements is in your life?

Example question: a. Being a parent is harder than I thought it would be. (Never true to always true)

```
DS3[, c(540:543)] <- sim_impu(DS3[, c(540:543)], 9, repl = T)
DS3[, c(540:543)] <- 6 - DS3[, c(540:543)] #reverse coding
summary(DS3[, c(540:543)])
```

```
##     PARHRD_1        CHBOTH_1        GIVEUP_1        TRAPPD_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:2.000   1st Qu.:3.000   1st Qu.:2.000   1st Qu.:3.000
##  Median :3.000   Median :3.000   Median :3.000   Median :4.000
##  Mean   :2.715   Mean   :3.398   Mean   :3.209   Mean   :4.035
##  3rd Qu.:3.000   3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:5.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
##  NA's   :30      NA's   :30      NA's   :30      NA's   :30
```

Note: 30 rows in the data matrix above contain NaN vectors, that is, the entire row is NA.

The higher the score, the more positive feeling about parenting.

Questionnaire #88: In general, I am a parent who...

Example question: a. Makes my children feel better when they talk over their worries with me... (Never true to always true)

```
DS3[, c(544:549)] <- sim_impu(DS3[, c(544:549)], 9, repl = T)
summary(DS3[, c(544:549)])
```

```
##     CHBETT_1        TKWCHD_1        JOYTKC_1        JOYCHI_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:4.000   1st Qu.:4.000   1st Qu.:4.000   1st Qu.:4.000
##  Median :4.000   Median :4.000   Median :4.000   Median :4.000
##  Mean   :3.879   Mean   :4.152   Mean   :4.251   Mean   :4.315
##  3rd Qu.:4.000   3rd Qu.:5.000   3rd Qu.:5.000   3rd Qu.:5.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
##  NA's   :31      NA's   :31      NA's   :31      NA's   :31
##     CHEERS_1        GOODTM_1
```

```
##  Min.   :1.000   Min.   :1.000
##  1st Qu.:4.000   1st Qu.:4.000
##  Median :4.000   Median :4.000
##  Mean   :4.016   Mean   :4.087
##  3rd Qu.:5.000   3rd Qu.:5.000
##  Max.   :5.000   Max.   :5.000
##  NA's   :31      NA's   :31
```

Note: 30 rows in the data matrix above contain NaN vectors, that is, the entire row is NA.

No reverse coding needed. **The higher the score, the more the parent cheers the child up**.

Questionnaire #89: When you and one of your children have had a disagreement, how often have you. . .

Example question: a. Just kept your opinion to your self (Never to Very often)

```
DS3[, c(550:553)] <- sim_impu(DS3[, c(550:553)], 9)
DS3[, c(550, 551)] <- 5 - DS3[, c(550, 551)] #note, category 0 to 4
DS3[, c(550:553)] <- sim_impu(DS3[, c(550:553)], 9, repl = T)
DS3[, c(550:553)] <- 5 - DS3[, c(550:553)]  # reverse coding again
summary(DS3[, c(550:553)])
```

```
##     KPOPN__1        DISCUS_1        ARGUEC_1        HITCHL_1
##  Min.   :0.000   Min.   :0.000   Min.   :1.000   Min.   :1.667
##  1st Qu.:1.000   1st Qu.:2.000   1st Qu.:3.000   1st Qu.:4.000
##  Median :1.000   Median :3.000   Median :3.000   Median :5.000
##  Mean   :1.299   Mean   :2.859   Mean   :3.469   Mean   :4.596
##  3rd Qu.:2.000   3rd Qu.:3.000   3rd Qu.:4.000   3rd Qu.:5.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
##  NA's   :33      NA's   :33      NA's   :33      NA's   :33
```

Note: 33 rows in the data matrix above contain NaN vectors, that is, the entire row is NA.

After reverse coding, **the higher the score, the less often the parent (aggressively) discusses issues with the child**.

Questionnaire #91: Now we're going to list statements that deal with ways that you may perceive yourself. Indicate how often these statements apply to you.

Example question: a. I feel good about myself.

```
DS3[, c(555:568)] <- sim_impu(DS3[, c(555:568)], 9)
DS3[, c(556:562, 565, 566)] <- 4 -  DS3[, c(556:562, 565, 566)]
DS3[, c(555:568)] <- sim_impu(DS3[, c(555:568)], 9, repl = T)
summary(DS3[, c(555:568)])
```

```
##     FEELGD_1        NOPROU_1        EDGE___1       FORGET_1
##  Min.   :1.000   Min.   :0.000   Min.   :0.00   Min.   :0.000
##  1st Qu.:3.000   1st Qu.:3.000   1st Qu.:2.00   1st Qu.:2.000
##  Median :3.000   Median :3.000   Median :3.00   Median :3.000
##  Mean   :3.032   Mean   :3.047   Mean   :2.87   Mean   :2.456
##  3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:3.00   3rd Qu.:3.000
##  Max.   :4.000   Max.   :4.000   Max.   :4.00   Max.   :4.000
##  NA's   :29      NA's   :29      NA's   :29     NA's   :29
##     NERVOU_1        CONCEN_1        UNPRED_1        MOANGR_1
```

```
##  Min.   :0.000   Min.   :0.000   Min.   :0.000   Min.   :0.00
##  1st Qu.:2.000   1st Qu.:2.000   1st Qu.:3.000   1st Qu.:2.00
##  Median :3.000   Median :3.000   Median :3.000   Median :3.00
##  Mean   :2.962   Mean   :2.768   Mean   :3.139   Mean   :2.73
##  3rd Qu.:4.000   3rd Qu.:3.000   3rd Qu.:4.000   3rd Qu.:3.00
##  Max.   :4.000   Max.   :4.000   Max.   :4.000   Max.   :4.00
##  NA's   :29      NA's   :29      NA's   :29      NA's   :29
##     EXANGR_1        ONTOPO_1        STRESS_1        NOCOPE_1
##  Min.   :0.000   Min.   :0.000   Min.   :0.000   Min.   :0.000
##  1st Qu.:1.000   1st Qu.:2.000   1st Qu.:1.000   1st Qu.:2.000
##  Median :2.000   Median :3.000   Median :2.000   Median :2.000
##  Mean   :1.859   Mean   :2.666   Mean   :1.802   Mean   :2.301
##  3rd Qu.:2.000   3rd Qu.:3.000   3rd Qu.:2.000   3rd Qu.:3.000
##  Max.   :4.000   Max.   :4.000   Max.   :4.000   Max.   :4.000
##  NA's   :29      NA's   :29      NA's   :29      NA's   :29
##     CONFID_1        CONFWK_1
##  Min.   :0.000   Min.   :0.000
##  1st Qu.:3.000   1st Qu.:3.000
##  Median :3.000   Median :3.000
##  Mean   :3.063   Mean   :3.302
##  3rd Qu.:4.000   3rd Qu.:4.000
##  Max.   :4.000   Max.   :4.000
##  NA's   :29      NA's   :29
```

Note: 29 rows in the data matrix above contain NaN vectors, that is, the entire row is NA.

After the reverse coding, **the higher the score, the more confident (and the less anger) the parent feels about him/herself**.

Combining all the questionnaires above

```
DS3_NEW <- rowMeans(DS3[, c(352:367)])
DS3_NEW <- cbind(DS3_NEW, rowMeans(DS3[, c(368:371)]))
DS3_NEW <- cbind(DS3_NEW, rowMeans(DS3[, c(379:386)]))
DS3_NEW <- cbind(DS3_NEW, rowMeans(DS3[, c(387:404)]))
DS3_NEW <- cbind(DS3_NEW, rowMeans(DS3[, c(441:460)]))
DS3_NEW <- cbind(DS3_NEW, rowMeans(DS3[, c(540:543)]))
DS3_NEW <- cbind(DS3_NEW, rowMeans(DS3[, c(544:549)]))
DS3_NEW <- cbind(DS3_NEW, rowMeans(DS3[, c(550:553)]))
DS3_NEW <- cbind(DS3_NEW, rowMeans(DS3[, c(555:568)]))
DS3_NEW <- cbind(DS3[, 1:3],DS3_NEW)

colnames(DS3_NEW)[4:12] <- c("Relationship with partners", "Argue with partners", "Childs bright future
                             "Activities with children", "Family hassles", "Feeling about parenting",
                             "Communation with children", "Argue with children", "Confidence about on

summary(DS3_NEW)
```

```
##     ï..ID_1          PERID__1        FAMID__1
##  Min.   :100101   Min.   :1.000   Min.   :1001
##  1st Qu.:113627   1st Qu.:1.000   1st Qu.:1136
##  Median :128202   Median :2.000   Median :1282
##  Mean   :156564   Mean   :1.572   Mean   :1565
##  3rd Qu.:206576   3rd Qu.:2.000   3rd Qu.:2065
```

```
##  Max.   :303102   Max.   :2.000   Max.   :3031
##                                   NA's   :1
##  Relationship with partners Argue with partners Childs bright future
##  Min.   :1.000              Min.   :2.000       Min.   :2.250
##  1st Qu.:3.125              1st Qu.:3.500       1st Qu.:4.125
##  Median :3.688              Median :3.750       Median :4.625
##  Mean   :3.595              Mean   :3.655       Mean   :4.464
##  3rd Qu.:4.188              3rd Qu.:4.000       3rd Qu.:5.000
##  Max.   :5.000              Max.   :5.000       Max.   :5.000
##  NA's   :96                 NA's   :96          NA's   :31
##  Activities with children Family hassles  Feeling about parenting
##  Min.   :1.143            Min.   :1.000   Min.   :1.000
##  1st Qu.:2.143            1st Qu.:2.050   1st Qu.:3.000
##  Median :2.471            Median :2.450   Median :3.250
##  Mean   :2.444            Mean   :2.429   Mean   :3.339
##  3rd Qu.:2.722            3rd Qu.:2.750   3rd Qu.:3.750
##  Max.   :4.000            Max.   :4.300   Max.   :5.000
##  NA's   :33               NA's   :565     NA's   :30
##  Communation with children Argue with children Confidence about oneself
##  Min.   :1.000             Min.   :1.500       Min.   :1.143
##  1st Qu.:3.833             1st Qu.:2.750       1st Qu.:2.429
##  Median :4.000             Median :3.000       Median :2.714
##  Mean   :4.117             Mean   :3.056       Mean   :2.714
##  3rd Qu.:4.500             3rd Qu.:3.250       3rd Qu.:3.071
##  Max.   :5.000             Max.   :5.000       Max.   :3.929
##  NA's   :31                NA's   :33          NA's   :29
```

Note that item "family hassles" contains 565 missing values (i.e., 565 persons did not fill in the famlily hassles questionnaire at all), and thus this item will be removed.

Process DS4 data

!! Note: For this dataset, the corresponding questionnaires are not included in the datafile folders, nor are they available online. However, based on the labels in the SPSS file and also in the Code book, we are able to tell which items belong together and form a questionnaire.

Questionnaire #29 (title unknown; tentative title: Self confidence/esteem)

example question: a. it is difficult to be myself.

```
DS4[, c(176:183)] <- sim_impu(DS4[, c(176:183)], 9)
DS4[, c(176, 181, 182)] <- 3 - DS4[, c(176, 181, 182)]  #recoding, answer category 0~3
DS4[, c(176:183)] <- sim_impu(DS4[, c(176:183)], 9, repl = T)
summary(DS4[, c(176:183)])
```

```
##     NORELX_1        FAMCRE_1        FLGOOD_1        SPECHP_1
##  Min.   :0.000   Min.   :0.000   Min.   :0.000   Min.   :0.000
##  1st Qu.:1.000   1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000
##  Median :2.000   Median :2.000   Median :3.000   Median :2.000
##  Mean   :1.927   Mean   :2.176   Mean   :2.437   Mean   :2.031
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000
##  Max.   :3.000   Max.   :3.000   Max.   :3.000   Max.   :3.000
##  NA's   :38      NA's   :38      NA's   :38      NA's   :38
##     NOEMBR_1        PROBAT_1        HAPBIR_1         ALWLUV_1
```

```
##  Min.   :0.000   Min.   :0.000   Min.   :0.0000   Min.   :0.000
##  1st Qu.:2.000   1st Qu.:2.000   1st Qu.:0.0000   1st Qu.:3.000
##  Median :3.000   Median :2.000   Median :0.0000   Median :3.000
##  Mean   :2.486   Mean   :2.182   Mean   :0.3789   Mean   :2.764
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:1.0000   3rd Qu.:3.000
##  Max.   :3.000   Max.   :3.000   Max.   :3.0000   Max.   :3.000
##  NA's   :38      NA's   :38      NA's   :38       NA's   :38
```

Note: 38 rows in the data matrix above contain NaN vectors, that is, the entire row is NA.

After reverse coding, **the higher the score, the higher level of self confidence/esteem**.

Question #44: Grades on most recent report card

```
DS4[which(DS4[, 304] == 99), 304]<- NA #99:missing value
DS4[which(DS4[, 304] == 9), 304] <- NA #9: no grade
DS4[, 304] <- 9 - DS4[, 304]
summary(DS4[, 304])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.000   6.000   7.000   6.771   8.000   8.000      50
```

Note: 50 persons don't have a score for this question.

After reverse coding, **the higher the score, the better the academic performance**.

Questionnaire #51: (title unknown; tentative title: social life and extracurricular activities)

```
DS4[, c(326:342)] <- sim_impu(DS4[, c(326:342)], 9, repl = T)
summary(DS4[, c(326:342)])
```

```
##     FRIEND_1       TV_____1        GAMES__1        READIN_1
##  Min.   :1.00   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:2.00   1st Qu.:3.000   1st Qu.:1.000   1st Qu.:1.000
##  Median :3.00   Median :3.000   Median :2.000   Median :2.000
##  Mean   :2.67   Mean   :3.242   Mean   :2.348   Mean   :2.386
##  3rd Qu.:3.00   3rd Qu.:4.000   3rd Qu.:3.000   3rd Qu.:3.000
##  Max.   :4.00   Max.   :4.000   Max.   :4.000   Max.   :4.000
##  NA's   :46     NA's   :46      NA's   :46      NA's   :46
##     HOB____1        YUTGRP_1        SPORTF_1        POOL___1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000
##  Median :2.000   Median :1.000   Median :2.000   Median :2.000
##  Mean   :2.056   Mean   :1.812   Mean   :2.376   Mean   :2.153
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000
##  Max.   :4.000   Max.   :4.000   Max.   :4.000   Max.   :4.000
##  NA's   :46      NA's   :46      NA's   :46      NA's   :46
##     VOLWRK_1        DRIVIN_1        TELEPH_1        TLKWAD_1
##  Min.   :1.000   Min.   :1.00   Min.   :1.000   Min.   :1.000
##  1st Qu.:1.000   1st Qu.:1.00   1st Qu.:3.000   1st Qu.:2.000
##  Median :1.000   Median :2.00   Median :3.000   Median :2.000
##  Mean   :1.592   Mean   :2.05   Mean   :3.184   Mean   :2.471
##  3rd Qu.:2.000   3rd Qu.:3.00   3rd Qu.:4.000   3rd Qu.:3.000
##  Max.   :4.000   Max.   :4.00   Max.   :4.000   Max.   :4.000
```

```
##  NA's   :46      NA's   :46     NA's   :46      NA's   :46
##     ARTNCL_1        ARTOCL_1        RELCLS_1        LESSON_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000
##  Median :1.000   Median :1.000   Median :1.000   Median :1.000
##  Mean   :1.866   Mean   :1.828   Mean   :1.643   Mean   :1.636
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:2.822   3rd Qu.:2.000
##  Max.   :4.000   Max.   :4.000   Max.   :4.000   Max.   :4.000
##  NA's   :46      NA's   :46      NA's   :46      NA's   :46
##     ATHLET_1
##  Min.   :1.000
##  1st Qu.:1.000
##  Median :3.000
##  Mean   :2.431
##  3rd Qu.:4.000
##  Max.   :4.000
##  NA's   :46
```

Note: 46 rows in the data matrix above contain NaN vectors, that is, the entire row is NA.

No reverse coding needed, **the higher the score, the more time for social life and extracurricular activities**.

Questionnaire #61: (title unknown; tentative title: (importance of) friendship)

Example question: a. I have lots of friends.

```
DS4[, c(359:367)] <- sim_impu(DS4[, c(359:367)], 9)
DS4[, c(361, 362)] <- 6 - DS4[, c(361, 362)]
DS4[, c(359:367)] <- sim_impu(DS4[, c(359:367)], 9, repl = T)
summary(DS4[, c(359:367)])
```

```
##     MANY___1        POPULR_1       NEW____1        ESTRAN_1
##  Min.   :1.000   Min.   :1.00   Min.   :1.000   Min.   :1.000
##  1st Qu.:4.000   1st Qu.:2.00   1st Qu.:3.000   1st Qu.:3.000
##  Median :4.000   Median :3.00   Median :4.000   Median :4.000
##  Mean   :4.186   Mean   :3.11   Mean   :3.893   Mean   :3.703
##  3rd Qu.:5.000   3rd Qu.:4.00   3rd Qu.:5.000   3rd Qu.:4.000
##  Max.   :5.000   Max.   :5.00   Max.   :5.000   Max.   :5.000
##  NA's   :57      NA's   :57     NA's   :57      NA's   :57
##     DOCARE_1        DOTRST_1        DOTALK_1        DOCOUN_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:4.000   1st Qu.:4.000   1st Qu.:3.000   1st Qu.:4.000
##  Median :4.000   Median :4.000   Median :4.000   Median :5.000
##  Mean   :4.064   Mean   :4.074   Mean   :3.987   Mean   :4.283
##  3rd Qu.:5.000   3rd Qu.:5.000   3rd Qu.:5.000   3rd Qu.:5.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
##  NA's   :57      NA's   :57      NA's   :57      NA's   :57
##     DOLIKE_1
##  Min.   :1.00
##  1st Qu.:4.00
##  Median :4.50
##  Mean   :4.34
##  3rd Qu.:5.00
##  Max.   :5.00
```

```
##  NA's   :57
```

Note: no missing values

After reverse coding, **the higher the score, the more important/the more friends/the more a person trusts his/her friend**.

Questionnaire #64: (title unknown; tentative title: self image)

Example question: a. I feel good about myself.

```
DS4[, c(381:397)] <- sim_impu(DS4[, c(381:397)], 9)
DS4[, c(383,384, 386:388, 390:394, 396,397)] <- 4 - DS4[, c(383,384, 386:388, 390:394, 396,397)] #rever
DS4[, c(381:397)] <- sim_impu(DS4[, c(381:397)], 9, repl = T)
summary(DS4[, c(381:397)])
```

```
##     FEELGD_1        IGOODU_1        NOPROU_1        LOSER__1
##  Min.   :0.000   Min.   :0.000   Min.   :0.000   Min.   :0.000
##  1st Qu.:2.000   1st Qu.:3.000   1st Qu.:2.000   1st Qu.:2.000
##  Median :3.000   Median :3.000   Median :3.000   Median :3.000
##  Mean   :3.032   Mean   :3.137   Mean   :2.752   Mean   :2.701
##  3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:3.000   3rd Qu.:3.000
##  Max.   :4.000   Max.   :4.000   Max.   :4.000   Max.   :4.000
##  NA's   :54      NA's   :54      NA's   :54      NA's   :54
##     CALM___1        EDGE___1        WORRY__1        FORGET_1
##  Min.   :0.000   Min.   :0.000   Min.   :0.000   Min.   :0.0
##  1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.0
##  Median :3.000   Median :3.000   Median :2.000   Median :3.0
##  Mean   :2.567   Mean   :2.772   Mean   :2.255   Mean   :2.4
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.0
##  Max.   :4.000   Max.   :4.000   Max.   :4.000   Max.   :4.0
##  NA's   :54      NA's   :54      NA's   :54      NA's   :54
##     CONTRO_1        NERVOU_1        CONCEN_1        UNPRED_1
##  Min.   :0.000   Min.   :0.000   Min.   :0.000   Min.   :0.000
##  1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000
##  Median :3.000   Median :3.000   Median :2.000   Median :3.000
##  Mean   :2.841   Mean   :2.632   Mean   :2.356   Mean   :2.806
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:4.000
##  Max.   :4.000   Max.   :4.000   Max.   :4.000   Max.   :4.000
##  NA's   :54      NA's   :54      NA's   :54      NA's   :54
##     MOANGR_1        EXANGR_1        ONTOPO_1        STRESS_1
##  Min.   :0.000   Min.   :0.000   Min.   :0.000   Min.   :0.000
##  1st Qu.:2.000   1st Qu.:1.000   1st Qu.:2.000   1st Qu.:1.000
##  Median :3.000   Median :2.000   Median :3.000   Median :2.000
##  Mean   :2.607   Mean   :2.022   Mean   :2.583   Mean   :1.792
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:2.909
##  Max.   :4.000   Max.   :4.000   Max.   :4.000   Max.   :4.000
##  NA's   :54      NA's   :54      NA's   :54      NA's   :54
##     NOCOPE_1
##  Min.   :0.000
##  1st Qu.:2.000
##  Median :2.000
##  Mean   :2.214
##  3rd Qu.:3.000
##  Max.   :4.000
```

```
##  NA's   :54
```

Note: 54 rows in the data matrix above contain NaN vectors, that is, the entire row is NA.

After reverse coding, **the higher the score, the more positive self image one has (the more calm, the less aggressive etc.)**.

Questionnaire #66 (title unknown; tentative title: Happiness)

Example question: a. I was bothered by things that are usually ok.

```
DS4[, c(399:418)] <- sim_impu(DS4[, c(399:418)], 9)
DS4[, c(399, 400, 402, 403, 405:411, 413, 414, 416:418)] <- 3 - DS4[, c(399, 400, 402, 403, 405:411, 413
DS4[, c(399:418)] <- sim_impu(DS4[, c(399:418)], 9, repl = T)
summary(DS4[, c(399:418)])
```

```
##     BOTHRD_1       APPETI_1        HOPEFU_1        BLUESH_1
##  Min.   :0.00   Min.   :0.000   Min.   :0.000   Min.   :0.000
##  1st Qu.:2.00   1st Qu.:2.000   1st Qu.:1.000   1st Qu.:2.000
##  Median :3.00   Median :3.000   Median :2.000   Median :3.000
##  Mean   :2.37   Mean   :2.483   Mean   :1.809   Mean   :2.365
##  3rd Qu.:3.00   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000
##  Max.   :3.00   Max.   :3.000   Max.   :3.000   Max.   :3.000
##  NA's   :57     NA's   :57      NA's   :57      NA's   :57
##     KEEPMN_1        HAPPYA_1        DEPRES_1       SPELLS_1
##  Min.   :0.000   Min.   :0.000   Min.   :0.00   Min.   :0.000
##  1st Qu.:1.000   1st Qu.:2.000   1st Qu.:2.00   1st Qu.:3.000
##  Median :2.000   Median :2.000   Median :2.00   Median :3.000
##  Mean   :1.896   Mean   :2.229   Mean   :2.25   Mean   :2.686
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.00   3rd Qu.:3.000
##  Max.   :3.000   Max.   :3.000   Max.   :3.00   Max.   :3.000
##  NA's   :57      NA's   :57      NA's   :57     NA's   :57
##     FEARFL_1        DISLIK_1        RESTLS_1        TKLESS_1
##  Min.   :0.000   Min.   :0.000   Min.   :0.000   Min.   :0.000
##  1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000
##  Median :3.000   Median :2.000   Median :3.000   Median :3.000
##  Mean   :2.581   Mean   :2.259   Mean   :2.259   Mean   :2.333
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000
##  Max.   :3.000   Max.   :3.000   Max.   :3.000   Max.   :3.000
##  NA's   :57      NA's   :57      NA's   :57      NA's   :57
##     LONELY_1        JOYLIF_1        FELTSD_1        NOTGOO_1
##  Min.   :0.000   Min.   :0.000   Min.   :0.000   Min.   :0.000
##  1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000
##  Median :3.000   Median :2.000   Median :2.000   Median :2.000
##  Mean   :2.318   Mean   :2.192   Mean   :2.186   Mean   :2.236
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000
##  Max.   :3.000   Max.   :3.000   Max.   :3.000   Max.   :3.000
##  NA's   :57      NA's   :57      NA's   :57      NA's   :57
##     GOOD___1        FAILUR_1        UNFRDL_1        EFFORT_1
##  Min.   :0.000   Min.   :0.000   Min.   :0.000   Min.   :0.000
##  1st Qu.:2.000   1st Qu.:3.000   1st Qu.:2.000   1st Qu.:1.000
##  Median :2.000   Median :3.000   Median :2.000   Median :2.000
##  Mean   :2.148   Mean   :2.712   Mean   :2.269   Mean   :1.708
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000
##  Max.   :3.000   Max.   :3.000   Max.   :3.000   Max.   :3.000
```

```
##  NA's   :57      NA's   :57      NA's   :57      NA's   :57
```

Note: 57 rows in the data matrix above contain NaN vectors, that is, the entire row is NA.

After reverse coding, **the higher the score, the more happier the person was**.

Questionnaire #77: (title unknown; tentative title: Confidence about the future)

Example question: a. you will have a job that pays well.

```
DS4[, c(468:480)] <- sim_impu(DS4[, c(468:480)], 9, repl = T)
summary(DS4[, c(468:480)])
```

```
##     GOODJB_1        TNOWN__1       FUNJOB_1       LVWFAM_1
##  Min.   :1.000   Min.   :1.00   Min.   :1.00   Min.   :1.000
##  1st Qu.:4.000   1st Qu.:4.00   1st Qu.:4.00   1st Qu.:3.000
##  Median :4.000   Median :4.00   Median :4.00   Median :3.000
##  Mean   :4.046   Mean   :4.21   Mean   :4.18   Mean   :3.139
##  3rd Qu.:5.000   3rd Qu.:5.00   3rd Qu.:5.00   3rd Qu.:4.000
##  Max.   :5.000   Max.   :5.00   Max.   :5.00   Max.   :5.000
##  NA's   :50      NA's   :50     NA's   :50     NA's   :50
##     MARRIE_1        CHILDR_1        HAPFAM_1        HEALTH_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:4.000   1st Qu.:4.000   1st Qu.:4.000   1st Qu.:4.000
##  Median :4.000   Median :4.000   Median :4.000   Median :4.000
##  Mean   :4.208   Mean   :4.059   Mean   :4.191   Mean   :4.124
##  3rd Qu.:5.000   3rd Qu.:5.000   3rd Qu.:5.000   3rd Qu.:5.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
##  NA's   :50      NA's   :50      NA's   :50      NA's   :50
##     FREEDO_1        RESPEC_1        FRIEGO_1        BETTER_1
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:3.000   1st Qu.:4.000   1st Qu.:4.000   1st Qu.:3.000
##  Median :4.000   Median :4.000   Median :4.000   Median :3.000
##  Mean   :3.939   Mean   :4.037   Mean   :4.276   Mean   :3.526
##  3rd Qu.:5.000   3rd Qu.:5.000   3rd Qu.:5.000   3rd Qu.:4.000
##  Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
##  NA's   :50      NA's   :50      NA's   :50      NA's   :50
##     BEST___1
##  Min.   :1.000
##  1st Qu.:3.000
##  Median :3.000
##  Mean   :3.446
##  3rd Qu.:4.000
##  Max.   :5.000
##  NA's   :50
```

Note: 50 rows in the data matrix above contain NaN vectors, that is, the entire row is NA.

No reverse coding needed; **the higher the score, the more confident a person is about the future**.

Combine all the data above

```
DS4_NEW <- rowMeans(DS4[, 176:183]) #self confidence/esteem
DS4_NEW <- cbind(DS4_NEW, DS4[, 304]) #academic performance
DS4_NEW <- cbind(DS4_NEW, rowMeans(DS4[, 326:342])) # social/extracurricular
```

```
DS4_NEW <- cbind(DS4_NEW, rowMeans(DS4[, 359:367])) # importance of friendship
DS4_NEW <- cbind(DS4_NEW, rowMeans(DS4[, 381:397])) # self image
DS4_NEW <- cbind(DS4_NEW, rowMeans(DS4[, 399:418])) # happiness
DS4_NEW <- cbind(DS4_NEW, rowMeans(DS4[, 468:480])) # confidence about future
DS4_NEW <- cbind(DS4[, 1:3], DS4_NEW)

colnames(DS4_NEW)[4:10] <- c("Self confidence/esteem", "Academic performance", "Social life and extracu
                             "Self Image", "Happiness", "Confidence about the future")
summary(DS4_NEW)
```

```
##     ï..ID_1          PERID__1        FAMID__1    Self confidence/esteem
##  Min.   :100103   Min.   :3.000   Min.   :1001   Min.   :0.000
##  1st Qu.:110804   1st Qu.:3.000   1st Qu.:1108   1st Qu.:1.750
##  Median :119903   Median :3.000   Median :1199   Median :2.125
##  Mean   :131786   Mean   :3.359   Mean   :1318   Mean   :2.048
##  3rd Qu.:128904   3rd Qu.:4.000   3rd Qu.:1289   3rd Qu.:2.375
##  Max.   :303104   Max.   :6.000   Max.   :3031   Max.   :3.000
##                                                  NA's   :38
##  Academic performance Social life and extracurricular
##  Min.   :1.000        Min.   :1.294
##  1st Qu.:6.000        1st Qu.:1.941
##  Median :7.000        Median :2.235
##  Mean   :6.771        Mean   :2.220
##  3rd Qu.:8.000        3rd Qu.:2.471
##  Max.   :8.000        Max.   :4.000
##  NA's   :50           NA's   :46
##  Importance of friendship   Self Image       Happiness
##  Min.   :1.000            Min.   :0.5882   Min.   :0.250
##  1st Qu.:3.667            1st Qu.:2.2353   1st Qu.:2.000
##  Median :4.000            Median :2.5882   Median :2.400
##  Mean   :3.960            Mean   :2.5570   Mean   :2.265
##  3rd Qu.:4.361            3rd Qu.:2.9412   3rd Qu.:2.600
##  Max.   :5.000            Max.   :3.8235   Max.   :3.000
##  NA's   :57               NA's   :54       NA's   :57
##  Confidence about the future
##  Min.   :1.462
##  1st Qu.:3.615
##  Median :3.923
##  Mean   :3.952
##  3rd Qu.:4.308
##  Max.   :5.000
##  NA's   :50
```

Step 3. Combine DS3 and DS4 data

Because "family hassels" in DS3_NEW contains too much NA values, better remove this column.

```
DS3_NEW <- DS3_NEW[, -8]
```

Remove NA's.

```
DS3_NEW_Final <- DS3_NEW[-sort(unique(which(is.na(DS3_NEW), arr.ind = T)[, 1])), ]
DS4_NEW_Final <- DS4_NEW[-sort(unique(which(is.na(DS4_NEW), arr.ind = T)[, 1])), ]
```

```
summary(DS3_NEW_Final)
```

```
##     ï..ID_1          PERID__1        FAMID__1
##  Min.   :100101   Min.   :1.000   Min.   :1001
##  1st Qu.:113052   1st Qu.:1.000   1st Qu.:1130
##  Median :128752   Median :2.000   Median :1288
##  Mean   :157267   Mean   :1.538   Mean   :1573
##  3rd Qu.:206676   3rd Qu.:2.000   3rd Qu.:2067
##  Max.   :303102   Max.   :2.000   Max.   :3031
##  Relationship with partners Argue with partners Childs bright future
##  Min.   :1.000              Min.   :2.000       Min.   :2.250
##  1st Qu.:3.125              1st Qu.:3.500       1st Qu.:4.125
##  Median :3.688              Median :3.750       Median :4.625
##  Mean   :3.594              Mean   :3.662       Mean   :4.472
##  3rd Qu.:4.188              3rd Qu.:4.000       3rd Qu.:5.000
##  Max.   :5.000              Max.   :5.000       Max.   :5.000
##  Activities with children Feeling about parenting
##  Min.   :1.143            Min.   :1.000
##  1st Qu.:2.146            1st Qu.:3.000
##  Median :2.485            Median :3.250
##  Mean   :2.449            Mean   :3.347
##  3rd Qu.:2.722            3rd Qu.:3.750
##  Max.   :4.000            Max.   :5.000
##  Communation with children Argue with children Confidence about oneself
##  Min.   :1.000             Min.   :1.50        Min.   :1.143
##  1st Qu.:3.833             1st Qu.:2.75        1st Qu.:2.429
##  Median :4.000             Median :3.00        Median :2.760
##  Mean   :4.111             Mean   :3.07        Mean   :2.722
##  3rd Qu.:4.500             3rd Qu.:3.50        3rd Qu.:3.071
##  Max.   :5.000             Max.   :5.00        Max.   :3.929
```

```
summary(DS4_NEW_Final)
```

```
##     ï..ID_1          PERID__1        FAMID__1    Self confidence/esteem
##  Min.   :100103   Min.   :3.000   Min.   :1001   Min.   :0.000
##  1st Qu.:110103   1st Qu.:3.000   1st Qu.:1101   1st Qu.:1.875
##  Median :119603   Median :3.000   Median :1196   Median :2.125
##  Mean   :130468   Mean   :3.344   Mean   :1305   Mean   :2.069
##  3rd Qu.:129253   3rd Qu.:4.000   3rd Qu.:1292   3rd Qu.:2.375
##  Max.   :303104   Max.   :6.000   Max.   :3031   Max.   :3.000
##  Academic performance Social life and extracurricular
##  Min.   :1.000        Min.   :1.294
##  1st Qu.:6.000        1st Qu.:1.941
##  Median :7.000        Median :2.235
##  Mean   :6.806        Mean   :2.219
##  3rd Qu.:8.000        3rd Qu.:2.471
##  Max.   :8.000        Max.   :3.235
##  Importance of friendship   Self Image       Happiness
##  Min.   :1.556            Min.   :0.5882   Min.   :0.250
##  1st Qu.:3.667            1st Qu.:2.2353   1st Qu.:2.025
##  Median :4.000            Median :2.5882   Median :2.400
##  Mean   :3.958            Mean   :2.5652   Mean   :2.270
##  3rd Qu.:4.333            3rd Qu.:2.9412   3rd Qu.:2.600
##  Max.   :5.000            Max.   :3.8235   Max.   :3.000
```

```
##  Confidence about the future
##  Min.   :1.615
##  1st Qu.:3.615
##  Median :3.923
##  Mean   :3.953
##  3rd Qu.:4.308
##  Max.   :5.000
```

Devide DS3 dataset into to datasets – one for the mother and one for the father, so that in the end, we have three datasets (i.e., mother, father, child).

```
DS3_Mom <- DS3_NEW_Final[DS3_NEW_Final[, 2]==2, ]
DS3_Dad <- DS3_NEW_Final[DS3_NEW_Final[, 2]==1, ]
DS4_Kid <- DS4_NEW_Final

family_index <- intersect(intersect(DS3_Mom[, 3], DS3_Dad[, 3]), DS4_Kid[, 3])
DS3_Mom_Final <- DS3_Mom[DS3_Mom[, 3]==family_index[1], ]
for(i in 2:length(family_index)){
  DS3_Mom_Final <- rbind(DS3_Mom_Final, DS3_Mom[DS3_Mom[, 3]==family_index[i], ])
}
DS3_Dad_Final <- DS3_Dad[DS3_Dad[, 3]==family_index[1], ]
for(i in 2:length(family_index)){
  DS3_Dad_Final <- rbind(DS3_Dad_Final, DS3_Dad[DS3_Dad[, 3]==family_index[i], ])
}
DS4_Kid_Final <- DS4_Kid[DS4_Kid[, 3]==family_index[1], ]
for(i in 2:length(family_index)){
  DS4_Kid_Final <- rbind(DS4_Kid_Final, DS4_Kid[DS4_Kid[, 3]==family_index[i], ])
}

DS4_Kid_Final <- DS4_Kid_Final[!duplicated(DS4_Kid_Final[, 3]), ]  # a family can have more than 1 chil
```

We now have the final datasets for mother, father, and child. Each row representing the mother/father/child from the SAME family.

```
family_data <- list("mom" = DS3_Mom_Final[, -c(1:3)], "dad" = DS3_Dad_Final[, -c(1:3)], "child" = DS4_K
colnames(family_data[[1]]) <- c("M: Relationship with partners", "M: Argue with partners", "M: Childs b
                     "M: Activities with children", "M: Feeling about parenting",
                     "M: Communation with children", "M: Argue with children",
                     "M: Confidence about oneself")
colnames(family_data[[2]]) <- c("D: Relationship with partners", "D: Argue with partners", "D: Childs b
                    "D: Activities with children", "D: Feeling about parenting",
                    "D: Communation with children", "D: Argue with children",
                    "D: Confidence about oneself")

#save(family_data, file = "D:\\Dropbox\\Dropbox\\tilburg office\\Research SCA\\Project 2 software Simul
save(family_data, file = "D:\\Dropbox\\Tilburg office\\Research SCA\\Project 2 software Simultaneous\\n
```