** 1. EXTERNAL WEIGHT FACTORS:
** CALCULATION OF POPULATION WEIGHTS FOR LASA-WAVES.
** EXAMPLE: LASAH.

** PRELIMINARY: When do you use external weights?
** Answer: When you calculate simple frequencies or cross-tabs and you want to achieve a representation of the general population.
** You can interpret the values of the weighting variable to indicate the number of observations represented by single cases in your data file; for LASA they do mostly not exceed 2 and do not fall below 0; the mean weight should be equal to 1.
** Note, that weights are not needed (and, in SPSS, not correctly treated, see https://www.ibm.com/support/knowledgecenter/en/SSLVMB_24.0.0/spss/base/idh_weig.html) when you analyse associations; for association analyses, you need to control for age and sex in order to achieve generalizable results.

** Merge lasaz002, lasaz004 and lasaz008 into one file.

```
match files  /file = 'pad\LASAZ002.SAV'
                        /file = 'pad\LASAZ004.SAV'
                        /file = 'pad\LASAZ008.SAV'
                        /by respnr
                        /keep respnr  hresult sex hage.
```

** Select the cases for which data are from H wave (data collected in 2011-2012).
** N.B. It is very important that if your analytic sample includes fewer participants than the complete sample, you need to first select this smaller sample before calculating the weights for it.

```
FILTER OFF.
USE ALL.
SELECT IF (hresult = 5).
EXECUTE.
```

```
desc hage.
```

** Recode hage into 5-year categories, as population weights are used for 5-year age categories.
** Minimum age is 63.97; participants aged up to 65 are added to the category 65-69.99.
** Maximum age is 104.21; participants aged 95 and over are defined in a separate category 95+.
** It is important that you assign high numbers as codes to the new categories, because in the next step you multiply these by 1 and 2 (the codes for sex), and the results of the multiplication should not overlap; i.e., each combination of age-category and sex should have a unique value.

```
compute hage5= hage.
```

```
recode hage5 (LOWEST thru 69.99=65) (70 thru 74.99 = 70)
 (75 thru 79.99 = 75) (80 thru 84.99 = 80) (85 thru 89.99=85)
 (90 thru 94.99=90) (95 thru HIGHEST = 95).
value lab hage5 65 '64-70' 70 '70-75' 75 '75-80' 80 '80-85'
                   85 '85-90' 90 '90-95' 95 '95+'.
fre hage5.

** Preparation for non-parametric test to examine the deviation of the
sex-age distribution in LASA ("observed frequency) from the
distribution in the Dutch population ("expected frequency").
** The expected frequencies are derived from Statistics Netherlands
(population composition as of 01-01-2012 by sex and 5-year age group).
** Construct one variable hsexlft combining the 2x7 sex-age categories.

compute hsexlft = sex * hage5.

** This is the variable that you test against the population
distribution ('expected' numbers), using a non-parametric test for
goodness-of-fit (chi-square with one degree of freedom).
** Note, that the first row of expected numbers contains the population
numbers in successive age categories for men, and the second row
contains the respective numbers for women.
** The attached excel file (converted to PDF) contains the population
numbers for the successive LASA waves from 1996 onwards; the numbers
used for LASAH are marked in yellow in the column for 2012.

npar test chisquare = hsexlft
 /expected =  432989 310115 224507 144421 69526 20912 3376
              441067 339638 282037 224161 146361 61381 15877.

** When the npar test is significant, the deviation of the sample
distribution from the population distribution is significant.
** This test provides observed and expected numbers for each category.
** The quotients of these numbers (1 / observed / expected) constitute
the sample weights for each sex-age category.
** This is how you calculate the weight variable hwsexlft.

compute hewsexlft = 0.
if (hsexlft = 65) hewsexlft = (1/(207 /242.6)).
if (hsexlft = 70) hewsexlft = (1/(168/173.8)).
if (hsexlft = 75) hewsexlft = (1/(102/125.8)).
if (hsexlft = 80) hewsexlft = (1/(94 / 80.9)).
if (hsexlft = 85) hewsexlft = (1/( 53 / 39.0)).
if (hsexlft = 90) hewsexlft = (1/( 16/ 11.7)).
if (hsexlft = 95) hewsexlft = (1/( 11 / 1.9)).
if (hsexlft = 130) hewsexlft = (1/(242/247.1)).
if (hsexlft = 140) hewsexlft = (1/(189 /190.3)).
if (hsexlft = 150) hewsexlft = (1/(148 /158.0)).
if (hsexlft = 160) hewsexlft = (1/(134/125.6)).
if (hsexlft = 170) hewsexlft = (1/( 93/ 82.0)).
if (hsexlft = 180) hewsexlft = (1/( 45 / 34.4)).
```

```
if (hsexlft = 190) hewsexlft = (1/( 20/ 8.9)).

var label hewsexlft ‘population weights as of 01-01-2012’.

** Check the calculation by applying the weight variable hewsexlft.

weight by hewsexlft.
cro tab sex by hlft.

** The cross table should show the weighted numbers.

weight off.

desc hewsexlft.

** If your mean weight is NOT equal to 1, check your calculation!.

** Once you apply a weight variable, it remains in effect until you
select another weight variable or turn off weighting. If you save a
weighted data file, weighting information is saved with the data file.
** You can turn off weighting at any time, even after the file has been
saved in weighted form.
** Regardless, it is best to switch the weight variable off, as soon as
you do not need it anymore.
```