Adjusting or Normalizing Weights "On the Fly" in SPSS[1]

Non-random samples usually require weighting. The weight a case has is usually a function of the likelihood of inclusion in the sample. Many data sets data include a variable that should be used as a weight.

In analysis, the weighted number of cases should equal the unweighted number of cases, and the mean of the weights should be 1. If analysis is done separately by subgroup, the mean of weights for each subgroup should be 1. If the weighted number of cases is very different from the unweighted number of cases, tests of significance will not be valid. SPSS does not adjust weights automatically.

Weights can be adjusted by dividing the weight by the mean of weights. The relative values of the weights are not changed, but they are adjusted so that the mean is 1, and the sum of weights equals the N of cases. The mean of the weights by running descriptives on the weight variable.

Following is a method of adjusting weights "on the fly" in SPSS so that the weights will have a mean of 1.

```
** ADJUST SPSSX.  ADJUST WEIGHTS "ON THE FLY" IN SPSSX .
* KEEP ONLY THE VARIABLES YOU WILL USE TO SAVE TIME AND SPACE IN THE SORT.
GET FILE= FARM40 / KEEP WEIGHT RACE SEX MARSTAT2.
* TURN WEIGHTS OFF.
WEIGHT OFF.

** IMPORTANT - BEFORE ADJUSTING WEIGHTS -- SELECT CASES YOU WANT TO USE IN THE ANALYSIS.
SELECT IF NOT(MISSING(RACE)) AND NOT(MISSING(SEX)) AND NOT(MISSING(MARSTAT2)) AND AGECAT=1 .

* IF YOU ARE NOT DOING SEPARATE ANALYSIS FOR SUBGROUPS, .
*  CREATE A CONSTANT TO USE IN "AGGREGATE" AND "MATCH FILES".
COMPUTE CONSTANT=1.

* NO NEED TO SORT CASES IF NOT DOING SEPARATE ANALYSIS FOR SUBGROUPS.
SORT CASES BY RACE SEX .
SAVE OUT=  'd:\temp\tempall.sav'  .

* CREATE A NEW DATASET "tempagg.sav" WHICH WILL HAVE .
* CASES FOR EACH SUBGROUP.  THE VARIABLES WILL BE
* race sex and meanwgt, which will be the mean of weights for the group.
* SUBSTITUTE "CONSTANT" FOR "RACE SEX" IF NOT DOING SEPARATE ANALYSIS BY SUBGROUP.

aggregate outfile = 'd:\temp\tempagg.sav'
  / break = race sex
  / meanwgt = mean( weight ) .

* MERGE NEW DATASET WITH ORIGINAL, BY SUBGROUPS
* SUBSTITUTE "CONSTANT" FOR "RACE SEX" IF NOT DOING SEPARATE ANALYSIS BY SUBGROUP.
match files
    table = 'd:\temp\tempagg.sav'
  / file =  'd:\temp\tempall.sav'
  / by race sex  .

* CREATE ADJUSTED WEIGHT.
compute adjwt = weight / meanwgt .
VARIABLE LABEL ADJWT 'WEIGHT ADJ FOR RACE SEX'.

* IF DOING SEPARATE ANALYSIS BY SUBGROUP - .
SPLIT FILE BY RACE SEX.

* UNWEIGHTED MEANS WILL BE 1.
DESC VAR = ADJWT.

* TURN WEIGHTS ON .
WEIGHT BY ADJWT .
FREQ VAR=MARSTAT2.
```

---

[1]Prepared by Patty Glynn, University of Washington. March 12, 2001, updated 6/16/02, updated 7/8/04.