



17

SAS[®] Macro

Introduction	248
Structure of the SAS [®] Macro	248



INTRODUCTION

This chapter presents the 17 SAS® macros used in the previous chapters. These are also available from www.pisa.oecd.org. Table 17.1 presents a summary of the 17 SAS® macros. The file names are in blue and the macro names as well as their arguments are in black.

STRUCTURE OF THE SAS® MACROS

All SAS® macros have the same structure.

- The first step consists of:
 - Saving the INFILE data file into a temporary data file denoted BRRDATA and dropping all variables that are not necessary for the analysis; and
 - Sorting that file by all variables specified in the BYVAR argument.
- The second step is the iterative part of the macro:
 - The requested statistics is computed 81 times for the computation without plausible values or 405 times for the computation with plausible values;
 - At each run, the results are stored in a temporary file, with the number of the replicate and the number of the plausible values if needed then added to that file; and
 - That file is merged with another temporary file that keeps the results of all runs.
- The third step consists of data file transformations to allow the computation of the final estimate and its respective standard error. This step is quite specific to each SAS® macro.
- The fourth step is devoted to the computation of the final statistic and its respective standard error:
 - The final estimates and 80 replicates are separated, *i.e.* BRR_TEMP1 is divided into BRR_TEMP2 (final estimates) and BRR_TEMP3 (replicate estimates). In the case of the plausible values, the five estimates are averaged and the measurement error is computed;
 - The two files are then merged to create BRR_TEMP4 and the squared differences between the final estimate and the 80 replicates are computed;
 - Through a PROC UNIVARIATE procedure, the sum of the squared differences is computed and divided by 20. These results are saved in BRR_TEMP5; and
 - BRR_TEMP3 (finale estimates and, in the case of the plausible values, the measurement variance) and BRR_TEMP5 (the sampling variance estimates) are combined. The final estimates and their respective standard errors are saved in the OUTFILE datafile.
- The final step flags the statistics for not reaching minimum sample size requirements that researchers set. This step is not included in some macros.
 - The unweighted numbers of students, the unweighted number of schools and the weighted percentage of the population are firstly computed per population estimates returned by the macro;
 - The unweighted number of schools and the weighted percentage of the population are then compared with the benchmarks provided by the researchers; and
 - If these numbers are below the benchmarks, the results are flagged.

The SAS® syntax is presented hereafter.



Table 17.1
Synthesis of the 17 SAS® macros

Requested statistics	Without plausible values	With plausible values
Mean, STD, VAR, Quartiles, Median, Percentiles	PROC_MEANS_NO_PV.sas %BRR_PROCMEAN (INFILE= , REPLI_ROOT= , BYVAR= , VAR= , STAT= , LIMIT= , LIMIT_CRITERIA= , ID_SCHOOL= , OUTFILE=);	PROC_MEANS_PV.sas %BRR_PROCMEAN_PV (INFILE= , REPLI_ROOT= , BYVAR= , PV_ROOT= , STAT= , LIMIT= , LIMIT_CRITERIA= , ID_SCHOOL= , OUTFILE=);
Frequencies	PROC_FREQ_NO_PV.sas %BRR_FREQ (INFILE= , REPLI_ROOT= , BYVAR= , VAR= , LIMIT= , LIMIT_CRITERIA= , ID_SCHOOL= , OUTFILE=);	PROC_FREQ_PV.sas %BRR_FREQ_PV (INFILE= , REPLI_ROOT= , BYVAR= , PV_ROOT= , LIMIT= , LIMIT_CRITERIA= , ID_SCHOOL= , OUTFILE=);
Regression	PROC_REG_NO_PV.sas %BRR_REG (INFILE= , REPLI_ROOT= , VARDEP= , EXPLICA= , BYVAR= , LIMIT= , LIMIT_CRITERIA= , ID_SCHOOL= , OUTFILE=);	PROC_REG_PV.sas %BRR_REG_PV (INFILE= , REPLI_ROOT= , EXPLICA= , BYVAR= , PV_ROOT= , LIMIT= , LIMIT_CRITERIA= , ID_SCHOOL= , OUTFILE=);
Correlation coefficients	PROC_CORR_NO_PV.sas %BRR_CORR (INFILE= , REPLI_ROOT= , BYVAR= , VAR1= , VAR2= , LIMIT= , LIMIT_CRITERIA= , ID_SCHOOL= , OUTFILE=);	PROC_CORR_PV.sas %BRR_CORR_PV (INFILE= , REPLI_ROOT= , BYVAR= , EXPLICA= , PV_ROOT= , LIMIT= , LIMIT_CRITERIA= , ID_SCHOOL= , OUTFILE=);
Differences on mean, STD, VAR, Quartiles, Median, Percentiles	PROC_DIF_NO_PV.sas %BRR_PROCMEAN_DIF (INFILE= , REPLI_ROOT= , BYVAR= , VAR= , COMPARE= , CATEGORY= , STAT= , OUTFILE=);	PROC_DIF_PV.sas %BRR_PROCMEAN_DIF_PV (INFILE= , REPLI_ROOT= , BYVAR= , PV_ROOT= , COMPARE= , CATEGORY= , STAT= , OUTFILE=);
Quartiles		QUARTILE_PV.sas %QUARTILE_PV (INFILE= , REPLI_ROOT= , BYVAR= , PV_ROOT= , INDEX= , LIMIT= , LIMIT_CRITERIA= , ID_SCHOOL= , OUTFILE=);
Relative risk	RELATIVE_RISK_NO_PV.sas %BRR_RR (INFILE= , REPLI_ROOT= , BYVAR= , ANTECEDENT= , OUTCOME= , LIMIT= , LIMIT_CRITERIA= , ID_SCHOOL= , OUTFILE=);	RELATIVE_RISK_PV.sas %BRR_RR_PV (INFILE= , REPLI_ROOT= , BYVAR= , ANTECEDENT_ROOT= , OUTCOME_ROOT= , LIMIT= , LIMIT_CRITERIA= , ID_SCHOOL= , OUTFILE=);
Effect size	EFFECT_SIZE_NO_PV.sas %BRR_EFFECT (INFILE= , REPLI_ROOT= , BYVAR= , VAR= , EFFECT= , OUTFILE=);	EFFECT_SIZE_PV.sas %BRR_EFFECT_PV (INFILE= , REPLI_ROOT= , BYVAR= , PV_ROOT= , EFFECT= , OUTFILE=);
Multilevel regression	PROC_MIXED_NO_PV.sas %BRR_MIXED (INFILE= , REPLI_ROOT= , VARDEP= , FIXEFP= , RANEF= , BYVAR= , LEVEL2= , OUTSCREEN= , OUTFILE=);	PROC_MIXED_PV.sas %BRR_MIXED_PV (INFILE= , REPLI_ROOT= , PV_ROOT= , FIXEFP= , RANEF= , BYVAR= , LEVEL2= , OUTSCREEN= , OUTFILE=);



Box 17.1 [1/3] SAS® macro of PROC_MEANS_NO_PV.sas

```

%MACRO BRR_PROCMEAN(INFILE = ,
                    REPLI_ROOT = ,
                    BYVAR = ,
                    VAR = ,
                    STAT = ,
                    LIMIT=,
                    LIMIT_CRITERIA=,
                    ID_SCHOOL=,
                    OUTFILE = );

/*
MEANING OF THE MACRO ARGUMENTS

INFILE = INPUT DATA FILE.
REPLI_ROOT = ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL WEIGHT
VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
BYVAR = BREAKDOWN VARIABLES.
VAR = VARIABLE ON WHICH THE REQUESTED STATISTIC WILL BE COMPUTED.
STAT = REQUESTED STATISTIQUE.
SUMWGT = SUM OF THE WEIGHT
MEAN = MEAN
VAR = VARIANCE
STD = STANDARD DEVIATION
CV = COEFFICIENT OF VARIATION
MEDIAN = MEDIAN
Q1 = FIRST QUARTILE
Q3 = THIRD QUARTILE
QRANGE = RANGE BETWEEN Q1 AND Q3
PX = PERCENTILE, WITH P1, P5, P10, P25, P50, P75, P90, P95 and P99
LIMIT = FLAGGING YES OR NO.
LIMIT_CRITERIA = 1) NUMBER OF STUDENTS 2) NUMBER OF SCHOOLS 3) PERCENTAGE OF STUDENTS
AND 4) NUMBER OF VARIABLES FROM THE BYVAR ARGUMENT FOR DEFINING THE POPULATION OF
REFERENCE.
ID_SCHOOL = VARIABLE NAME FOR THE SCHOOL IDENTIFICATION.
OUTFILE = FILE WITH THE STATISTIC ESTIMATES AND THEIR STANDARD ERROR ESTIMATES.

*/

OPTIONS NONOTES;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP1;
RUN;

PROC SORT DATA=&INFILE OUT=BRRDATA(KEEP=&REPLI_ROOT.0-&REPLI_ROOT.80 &BYVAR &VAR
&ID_SCHOOL);
    BY &BYVAR;
RUN;

%DO I = 0 %TO 80;
    PROC MEANS DATA=BRRDATA VARDEF=WGT NOPRINT;
        VAR &VAR ;
        BY &BYVAR;
        WEIGHT &REPLI_ROOT&I;
        OUTPUT OUT=MEAN_TEMP &STAT=PV;
    RUN;

    DATA MEAN_TEMP;
        SET MEAN_TEMP;
        L=&I;
    RUN;

    PROC APPEND BASE = BRR_TEMP1 DATA=MEAN_TEMP;
    RUN;

%END;

```



Box 17.1 [2/3] SAS® macro of PROC_MEANS_NO_PV.sas

```

DATA BRR_TEMP4;
    MERGE BRR_TEMP2 BRR_TEMP3;
    BY &BYVAR;
    VARI=( (PV-STAT)**2)*(1/20);
RUN;

PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
    VAR VARI;
    BY &BYVAR;
    OUTPUT OUT=BRR_TEMP5 SUM=SS;
RUN;

DATA BRR_TEMP6;
    MERGE BRR_TEMP3 BRR_TEMP5;
    BY &BYVAR;
    SESTAT=(SS)**0.5;
    FORMAT STAT F10.2;
    FORMAT SESTAT F10.2;
    KEEP &BYVAR STAT SESTAT;
RUN;

%IF (%UPCASE(&LIMIT)=NO) %THEN %DO;

    DATA &OUTFILE;
        SET BRR_TEMP6;
    RUN;

%END;
%ELSE %DO;
    DATA BRR_TEMP7;
        SET BRRDATA;
        NB_MISS=0;
        ARRAY LIST_VAR (1) &VAR;
        DO K=1 TO 1;
            IF (LIST_VAR(K) IN (.,.I,.M,.N)) THEN NB_MISS=NB_MISS+1;
        END;
        IF (NB_MISS>1) THEN NB_MISS=1;
    RUN;
    PROC FREQ DATA=BRR_TEMP7 NOPRINT;
        TABLE NB_MISS /OUT=BRR_TEMP8;
        BY &BYVAR;
        WHERE (NB_MISS=0);
    RUN;
    %LET FLAG_STUD=%SCAN(&LIMIT_CRITERIA,1);
    DATA BRR_TEMP8;
        SET BRR_TEMP8;
        IF (COUNT < &FLAG_STUD) THEN FLAG_STUD=1;
        ELSE FLAG_STUD=0;
        KEEP &BYVAR FLAG_STUD;
    RUN;
    PROC SORT DATA=BRR_TEMP7;
        BY &BYVAR &ID_SCHOOL;
    RUN;
    PROC FREQ DATA=BRR_TEMP7 NOPRINT;
        TABLE NB_MISS /OUT=BRR_TEMP9;
        BY &BYVAR &ID_SCHOOL;
        WHERE (NB_MISS=0);
    RUN;
    PROC FREQ DATA=BRR_TEMP9 NOPRINT;
        TABLE NB_MISS /OUT=BRR_TEMP10;
        BY &BYVAR;
    RUN;

```



Box 17.1 [3/3] SAS® macro of PROC_MEANS_NO_PV.sas

```

%LET FLAG_SCH=%SCAN(&LIMIT_CRITERIA,2);
DATA BRR_TEMP10;
  SET BRR_TEMP10;
  IF (COUNT < &FLAG_SCH) THEN FLAG_SCH=1;
  ELSE FLAG_SCH=0;
  KEEP &BYVAR FLAG_SCH;
RUN;
PROC SORT DATA=BRR_TEMP7;
  BY &BYVAR NB_MISS;
RUN;
PROC FREQ DATA=BRR_TEMP7 NOPRINT;
  TABLE NB_MISS/OUT=BRR_TEMP11;
  BY &BYVAR;
  WEIGHT &REPLI_ROOT.0;
RUN;
%LET K=1;
%LET POPREF=;
%LET NBVAR=%SCAN(&LIMIT_CRITERIA,4);
%DO %WHILE(&K <= &NBVAR);
  %LET POPREF_ADD=%SCAN(&BYVAR,&K);
  %LET POPREF=&POPREF &POPREF_ADD;
  %LET K=%EVAL(&K+1);
%END;

PROC MEANS DATA=BRR_TEMP11 NOPRINT;
  VAR COUNT;
  BY &POPREF;
  OUTPUT OUT=BRR_TEMP12 SUM=SOMWGT;
RUN;
%LET FLAG_PCT=%SCAN(&LIMIT_CRITERIA,3);
DATA BRR_TEMP13;
  MERGE BRR_TEMP11 BRR_TEMP12;
  BY &POPREF;
  PCT=(COUNT/SOMWGT)*100;
  IF (PCT < &FLAG_PCT) THEN FLAG_PCT=1;
  ELSE FLAG_PCT=0;
  IF (NB_MISS=0);
  KEEP &BYVAR FLAG_PCT;
RUN;
DATA &OUTFILE;
  MERGE BRR_TEMP6 BRR_TEMP8 BRR_TEMP10 BRR_TEMP13;
  BY &BYVAR;
RUN;
PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP7 BRR_TEMP8 BRR_TEMP9 BRR_TEMP10 BRR_TEMP11 BRR_TEMP12
BRR_TEMP13;
RUN;

%END;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5 BRR_TEMP6 MEAN_
TEMP BRRDATA;
RUN;

OPTIONS NOTES;

%MEND BRR_PROCMEAN;

```



Box 17.2 [1/3] SAS® macro of PROC_MEANS_PV.sas

```

%MACRO BRR_PROCMEAN_PV(INFILE =,
                        REPLI_ROOT =,
                        BYVAR =,
                        PV_ROOT =,
                        STAT =,
                        LIMIT=,
                        LIMIT_CRITERIA=,
                        ID_SCHOOL=,
                        OUTFILE =);

/*

MEANING OF THE MACRO ARGUMENTS

INFILE = INPUT DATA FILE.
REPLI_ROOT = ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL
WEIGHT VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
BYVAR = BREAKDOWN VARIABLES
PV_ROOT = ROOT OF THE 5 PLAUSIBLE VALUES VARIABLES NAMES
STAT = REQUESTED STATISTIQUE.
SUMWGT = SUM OF THE WEIGHT
MEAN = MEAN
VAR = VARIANCE
STD = STANDARD DEVIATION
CV = COEFFICIENT OF VARIATION
MEDIAN = MEDIAN
Q1 = FIRST QUANTILE
Q3 = THIRD QUANTILE
QRANGE = RANGE BETWEEN Q1 AND Q3
PX = PERCENTILE, WITH P1, P5, P10, P25, P50, P75, P90, P95 and P99
LIMIT = FLAGGING YES OR NO.
LIMIT_CRITERIA = 1) NUMBER OF STUDENTS 2) NUMBER OF SCHOOLS 3) PERCENTAGE OF
STUDENTS AND 4) NUMBER OF VARIABLES FROM THE BYVAR ARGUMENT FOR DEFINING THE
POPULATION OF REFERENCE.
ID_SCHOOL = VARIABLE NAME FOR THE SCHOOL IDENTIFICATION
OUTFILE = FILE WITH THE STATISTIC ESTIMATES AND THEIR STANDARD ERROR ESTIMATES.

*/

OPTIONS NONOTES;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP1;
RUN;

PROC SORT DATA=&INFILE
    OUT=BRRDATA(KEEP=&REPLI_ROOT.0-&REPLI_ROOT.80 &BYVAR &PV_ROOT.1-&PV_ROOT.5
&ID_SCHOOL);
    BY &BYVAR;
RUN;

%DO I = 0 %TO 80;
    PROC MEANS DATA=BRRDATA VARDEF=WGT NOPRINT;
        VAR &PV_ROOT.1 - &PV_ROOT.5 ;
        BY &BYVAR;
        WEIGHT &REPLI_ROOT&I;
        OUTPUT OUT=MEAN_TEMP &STAT=PV1 - PV5;
    RUN;

    DATA MEAN_TEMP;
        SET MEAN_TEMP;
        L=&I;
    RUN;

    PROC APPEND BASE = BRR_TEMP1 DATA=MEAN_TEMP;
    RUN;

%END;

```



Box 17.2 [2/3] SAS® macro of PROC_MEANS_PV.sas

```

DATA BRR_TEMP2 (DROP=STAT FIN1-FIN5 MESVAR) BRR_TEMP3 (KEEP=&BYVAR STAT FIN1-FIN5
MESVAR) ;
  SET BRR_TEMP1;
  IF L > 0 THEN OUTPUT BRR_TEMP2;
  ELSE DO;
    STAT = (PV1+PV2+PV3+PV4+PV5) /5;
    FIN1=PV1;
    FIN2=PV2;
    FIN3=PV3;
    FIN4=PV4;
    FIN5=PV5;
    MESVAR= (( (STAT-FIN1)**2) + ((STAT-FIN2)**2) + ((STAT-FIN3)**2) + ((STAT-
FIN4)**2) + ((STAT-FIN5)**2)) /4;
    OUTPUT BRR_TEMP3;
  END;
RUN;
PROC SORT DATA=BRR_TEMP2;
  BY &BYVAR;
RUN;
PROC SORT DATA=BRR_TEMP3;
  BY &BYVAR;
RUN;
DATA BRR_TEMP4;
  MERGE BRR_TEMP2 BRR_TEMP3;
  BY &BYVAR;
  ARRAY A (5)
    PV1-PV5;
  ARRAY B (5)
    FIN1-FIN5;
  ARRAY C (5)
    VAR1-VAR5;
  DO I=1 TO 5;
    C(I) = (1/20) * ((A(I) - B(I)) **2) ;
  END;
RUN;
PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
  VAR VAR1 VAR2 VAR3 VAR4 VAR5;
  BY &BYVAR;
  OUTPUT OUT=BRR_TEMP5 SUM=SS1 SS2 SS3 SS4 SS5;
RUN;
DATA BRR_TEMP6;
  MERGE BRR_TEMP3 BRR_TEMP5;
  BY &BYVAR;
  SAMP= (SS1+SS2+SS3+SS4+SS5) /5;
  FINVAR= (SAMP+ (1.2*MESVAR)) ;
  SESTAT= (FINVAR)**0.5;
  FORMAT STAT F10.1;
  FORMAT SESTAT F10.2;
  sestat2=(ss1+(1.2*MESVAR))**0.5;
  KEEP &BYVAR STAT SESTAT sestat2 ;
RUN;
%IF (%UPCASE(&LIMIT)=NO) %THEN %DO;

  DATA &OUTFILE;
    SET BRR_TEMP6;
  RUN;
%END;
%ELSE %DO;
  DATA BRR_TEMP7;
    SET BRCDATA;
    NB_MISS=0;
  RUN;
  PROC FREQ DATA=BRR_TEMP7 NOPRINT;
    TABLE NB MISS /OUT=BRR_TEMP8;
    BY &BYVAR;
  RUN;

```




Box 17.2 [3/3] SAS® macro of PROC_MEANS_PV.sas

```

%LET FLAG_STUD=%SCAN(&LIMIT_CRITERIA,1);
DATA BRR_TEMP8;
    SET BRR_TEMP8;
    IF (COUNT < &FLAG_STUD) THEN FLAG_STUD=1;
    ELSE FLAG_STUD=0;
    KEEP &BYVAR FLAG_STUD;
RUN;
PROC SORT DATA=BRR_TEMP7;
    BY &BYVAR &ID_SCHOOL;
RUN;
PROC FREQ DATA=BRR_TEMP7 NOPRINT;
    TABLE NB_MISS /OUT=BRR_TEMP9;
    BY &BYVAR &ID_SCHOOL;
RUN;
PROC FREQ DATA=BRR_TEMP9 NOPRINT;
    TABLE NB_MISS /OUT=BRR_TEMP10;
    BY &BYVAR;
RUN;
%LET FLAG_SCH=%SCAN(&LIMIT_CRITERIA,2);
DATA BRR_TEMP10;
    SET BRR_TEMP10;
    IF (COUNT < &FLAG_SCH) THEN FLAG_SCH=1;
    ELSE FLAG_SCH=0;
    KEEP &BYVAR FLAG_SCH;
RUN;
PROC SORT DATA=BRR_TEMP7;
    BY &BYVAR NB_MISS;
RUN;
PROC FREQ DATA=BRR_TEMP7 NOPRINT;
    TABLE NB_MISS/OUT=BRR_TEMP11;
    BY &BYVAR;
    WEIGHT &REPLI_ROOT.0;
RUN;
%LET K=1;
%LET POPREF=;
%LET NBVAR=%SCAN(&LIMIT_CRITERIA,4);
%DO %WHILE(&K <= &NBVAR);
    %LET POPREF_ADD=%SCAN(&BYVAR,&K);
    %LET POPREF=&POPREF &POPREF_ADD;
    %LET K=%EVAL(&K+1);
%END;
PROC MEANS DATA=BRR_TEMP11 NOPRINT;
    VAR COUNT;
    BY &POPREF;
    OUTPUT OUT=BRR_TEMP12 SUM=SOMWGT;
RUN;
%LET FLAG_PCT=%SCAN(&LIMIT_CRITERIA,3);
DATA BRR_TEMP13;
    MERGE BRR_TEMP11 BRR_TEMP12;
    BY &POPREF;
    PCT=(COUNT/SOMWGT)*100;
    IF (PCT < &FLAG_PCT) THEN FLAG_PCT=1;
    ELSE FLAG_PCT=0;
    KEEP &BYVAR FLAG_PCT;
RUN;
DATA &OUTFILE;
    MERGE BRR_TEMP6 BRR_TEMP8 BRR_TEMP10 BRR_TEMP13;
    BY &BYVAR;
RUN;
PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP7 BRR_TEMP8 BRR_TEMP9 BRR_TEMP10 BRR_TEMP11 BRR_
TEMP12 BRR_TEMP13;
RUN;
%END;
PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5 BRR_TEMP6 MEAN_
TEMP BRRDATA;
RUN;
OPTIONS NOTES;
%MEND BRR_PROCMEAN_PV;

```



Box 17.3 [1/3] SAS® macro of PROC_FREQ_NO_PV.sas

```

%macro BRR_FREQ(INFILE=,
                REPLI_ROOT=,
                BYVAR=,
                VAR=,
                LIMIT=,
                LIMIT_CRITERIA=,
                ID_SCHOOL=,
                OUTFILE=);

/*
MEANING OF THE MACRO ARGUMENTS

INFILE = INPUT DATA FILE.
REPLI_ROOT = ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL
WEIGHT VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
BYVAR = BREAKDOWN VARIABLES.
VAR = VARIABLES ON WHICH PERCENTAGES WILL BE COMPUTED.
LIMIT = FLAGGING YES OR NO.
LIMIT_CRITERIA = 1) NUMBER OF STUDENTS 2) NUMBER OF SCHOOLS 3) PERCENTAGE OF
STUDENTS AND 4) NUMBER OF VARIABLES FROM THE BYVAR ARGUMENT FOR DEFINING THE
POPULATION OF REFERENCE.
ID_SCHOOL = VARIABLE NAME FOR THE SCHOOL IDENTIFICATION.
OUTFILE = FILE WITH THE STATISTIC ESTIMATES AND THEIR STANDARD ERROR ESTIMATES.
*/
OPTIONS NONOTES;
PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP0;
RUN;
PROC SORT DATA=&INFILE
    OUT=BRRDATA(KEEP=&REPLI_ROOT.0-&REPLI_ROOT.80 &BYVAR &VAR &ID_SCHOOL);
    BY &BYVAR &VAR;
RUN;
%DO I = 0 %TO 80;
    PROC FREQ DATA=BRRDATA NOPRINT;
        TABLE &VAR /OUT=FREQ_TEMP ;
        BY &BYVAR;
        WEIGHT &REPLI_ROOT&I;
    RUN;
    DATA FREQ_TEMP;
        SET FREQ_TEMP;
        L=&I;
        KEEP &BYVAR L &VAR PERCENT;
    RUN;
    PROC APPEND BASE = BRR_TEMP0 DATA=FREQ_TEMP;
    RUN;
%END;
PROC SORT DATA=BRR_TEMP0 OUT=BRR_TEMP1 (RENAME=(PERCENT=PV));
    BY &BYVAR L &VAR;
RUN;
DATA BRR_TEMP2(KEEP=&BYVAR &VAR PV)BRR_TEMP3(KEEP=&BYVAR &VAR STAT);
    SET BRR_TEMP1;
    IF L > 0 THEN OUTPUT BRR_TEMP2;
    ELSE DO;
        STAT =PV;
        OUTPUT BRR_TEMP3;
    END;
RUN;
PROC SORT DATA=BRR_TEMP2;
    BY &BYVAR &VAR;
RUN;
PROC SORT DATA=BRR_TEMP3;
    BY &BYVAR &VAR;
RUN;

```



Box 17.3 [2/3] SAS® macro of PROC_FREQ_NO_PV.sas

```

DATA BRR_TEMP4;
    MERGE BRR_TEMP2 BRR_TEMP3;
    BY &BYVAR &VAR;
    VARI=((PV-STAT)**2)*(1/20);
RUN;

PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
    VAR VARI;
    BY &BYVAR &VAR;
    OUTPUT OUT=BRR_TEMP5 SUM=SS;
RUN;

DATA BRR_TEMP6;
    MERGE BRR_TEMP3 BRR_TEMP5;
    BY &BYVAR &VAR;
    SESTAT=(SS)**0.5;
    FORMAT STAT F10.2;
    FORMAT SESTAT F10.2;
    KEEP &BYVAR &VAR STAT SESTAT;
RUN;

%IF (%UPCASE(&LIMIT)=NO) %THEN %DO;

    DATA &OUTFILE;
        SET BRR_TEMP6;
    RUN;

%END;
%ELSE %DO;

    PROC FREQ DATA=BRRDATA NOPRINT;
        TABLE &VAR /OUT=BRR_TEMP7;
        BY &BYVAR;
    RUN;
    %LET FLAG_STUD=%SCAN(&LIMIT_CRITERIA,1);
    DATA BRR_TEMP7;
        SET BRR_TEMP7;
        IF (COUNT < &FLAG_STUD) THEN FLAG_STUD=1;
        ELSE FLAG_STUD=0;
        KEEP &BYVAR &VAR FLAG_STUD;
    RUN;
    PROC SORT DATA=BRRDATA;
        BY &BYVAR &ID_SCHOOL;
    RUN;
    PROC FREQ DATA=BRRDATA NOPRINT;
        TABLE &VAR /OUT=BRR_TEMP8;
        BY &BYVAR &ID_SCHOOL;
    RUN;
    PROC FREQ DATA=BRR_TEMP8 NOPRINT;
        TABLE &VAR /OUT=BRR_TEMP9;
        BY &BYVAR;
    RUN;
    %LET FLAG_SCH=%SCAN(&LIMIT_CRITERIA,2);
    DATA BRR_TEMP9;
        SET BRR_TEMP9;
        IF (COUNT < &FLAG_SCH) THEN FLAG_SCH=1;
        ELSE FLAG_SCH=0;
        KEEP &BYVAR &VAR FLAG_SCH;
    RUN;

    PROC FREQ DATA=BRRDATA NOPRINT;
        TABLE &VAR/OUT=BRR_TEMP10;
        BY &BYVAR;
        WEIGHT &REPLI_ROOT.0;
    RUN;

```



Box 17.3 [3/3] SAS® macro of PROC_FREQ_NO_PV.sas

```

%LET K=1;
  %LET POPREF=;
  %LET NBVAR=%SCAN(&LIMIT_CRITERIA,4);
  %DO %WHILE(&K <= &NBVAR);
    %LET POPREF_ADD=%SCAN(&BYVAR,&K);
    %LET POPREF=&POPREF &POPREF_ADD;
    %LET K=%EVAL(&K+1);
  %END;

PROC MEANS DATA=BRR_TEMP10 NOPRINT;
  VAR COUNT;
  BY &POPREF;
  OUTPUT OUT=BRR_TEMP11 SUM=SOMWGT;
RUN;
%LET FLAG_PCT=%SCAN(&LIMIT_CRITERIA,3);
DATA BRR_TEMP12;
  MERGE BRR_TEMP10 BRR_TEMP11;
  BY &POPREF;
  PCT=(COUNT/SOMWGT)*100;
  IF (PCT < &FLAG_PCT) THEN FLAG_PCT=1;
  ELSE FLAG_PCT=0;
  KEEP &BYVAR &VAR FLAG_PCT;
RUN;
DATA &OUTFILE;
  MERGE BRR_TEMP6 BRR_TEMP7 BRR_TEMP9 BRR_TEMP12;
  BY &BYVAR &VAR;
  IF (STAT NE .);
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP7 BRR_TEMP8 BRR_TEMP9 BRR_TEMP10 BRR_TEMP11 BRR_TEMP12;
RUN;
%END;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP0 BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5 BRR_TEMP6
  FREQ_TEMP BRRDATA;
RUN;

OPTIONS NOTES;

%MEND BRR_FREQ;

```



Box 17.4 [1/4] SAS® macro of PROC_FREQ_PV.sas

```

%macro BRR_FREQ_PV(INFILE=,
                   REPLI_ROOT=,
                   BYVAR=,
                   PV_ROOT=,
                   LIMIT=,
                   LIMIT_CRITERIA=,
                   ID_SCHOOL=,
                   OUTFILE=);

/*

MEANING OF THE MACRO ARGUMENTS
INFILE = INPUT DATA FILE.
REPLI_ROOT = ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL
WEIGHT VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
BYVAR = BREAKDOWN VARIABLES.
PV_ROOT = ROOT OF THE 5 PROFICIENCY LEVEL VARIABLES NAMES.
LIMIT = FLAGGING YES OR NO.
LIMIT_CRITERIA = 1) NUMBER OF STUDENTS 2) NUMBER OF SCHOOLS 3) PERCENTAGE OF
STUDENTS AND 4) NUMBER OF VARIABLES FROM THE BYVAR ARGUMENT FOR DEFINING THE
POPULATION OF REFERENCE.
ID_SCHOOL = VARIABLE NAME FOR THE SCHOOL IDENTIFICATION.
OUTFILE = FILE WITH THE STATISTIC ESTIMATES AND THEIR STANDARD ERROR
ESTIMATES.

*/

OPTIONS NONOTES;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP0;
RUN;
%DO I = 1 %TO 5;
  PROC SORT DATA=&INFILE
            OUT=BRRDATA(KEEP=&REPLI_ROOT.0-&REPLI_ROOT.80 &BYVAR &PV_ROOT.1-&PV_
ROOT.5 &ID_SCHOOL);
    BY &BYVAR &PV_ROOT&I;
  RUN;
  %DO J = 0 %TO 80;
    PROC FREQ DATA=BRRDATA NOPRINT;
      TABLE &PV_ROOT&I /OUT=FREQ_TEMP ;
      BY &BYVAR;
      WEIGHT &REPLI_ROOT&J;
    RUN;
    DATA FREQ_TEMP;
      SET FREQ_TEMP;
      K=&I;
      L=&J;
      &PV_ROOT=&PV_ROOT&I;
      KEEP &BYVAR K L PERCENT &PV_ROOT;
    RUN;
    PROC APPEND BASE = BRR_TEMP0 DATA=FREQ_TEMP;
  RUN;
  %END;
%END;
PROC SORT DATA=BRR_TEMP0;
  BY &BYVAR L &PV_ROOT;
RUN;
PROC TRANSPOSE DATA=BRR_TEMP0 OUT=BRR_TEMP1 PREFIX=PV;
  BY &BYVAR L &PV_ROOT;
  VAR PERCENT;
RUN;
DATA BRR_TEMP1;
  SET BRR_TEMP1;
  DROP _NAME_ _LABEL_;
RUN;

```



Box 17.4 [2/4] SAS® macro of PROC_FREQ_PV.sas

```

DATA BRR_TEMP2 (DROP=STAT FIN1-FIN5 MESVAR) BRR_TEMP3 (KEEP=&BYVAR &PV_ROOT STAT FIN1-
FIN5 MESVAR);
  SET BRR_TEMP1;
  IF L > 0 THEN OUTPUT BRR_TEMP2;
  ELSE DO;
    STAT = (PV1+PV2+PV3+PV4+PV5) / 5;
    FIN1=PV1;
    FIN2=PV2;
    FIN3=PV3;
    FIN4=PV4;
    FIN5=PV5;
    MESVAR=(((STAT-FIN1)**2) + ((STAT-FIN2)**2) + ((STAT-FIN3)**2) + ((STAT-
FIN4)**2) + ((STAT-FIN5)**2)) / 4;
    OUTPUT BRR_TEMP3;
  END;
RUN;
PROC SORT DATA=BRR_TEMP2;
  BY &BYVAR &PV_ROOT;
RUN;
PROC SORT DATA=BRR_TEMP3;
  BY &BYVAR &PV_ROOT;
RUN;
DATA BRR_TEMP4;
  MERGE BRR_TEMP2 BRR_TEMP3;
  BY &BYVAR &PV_ROOT;
  ARRAY A (5)
    PV1-PV5;
  ARRAY B (5)
    FIN1-FIN5;
  ARRAY C (5)
    VAR1-VAR5;
  DO I=1 TO 5;
    C(I) = (1/20) * ((A(I) - B(I)) ** 2);
  END;
RUN;
PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
  VAR VAR1 VAR2 VAR3 VAR4 VAR5;
  BY &BYVAR &PV_ROOT;
  OUTPUT OUT=BRR_TEMP5 SUM=SS1 SS2 SS3 SS4 SS5;
RUN;
DATA BRR_TEMP6;
  MERGE BRR_TEMP3 BRR_TEMP5;
  BY &BYVAR &PV_ROOT;
  SAMP = (SS1+SS2+SS3+SS4+SS5) / 5;
  FINVAR = (SAMP + (1.2 * MESVAR));
  SESTAT = (FINVAR) ** 0.5;
  KEEP &BYVAR &PV_ROOT STAT SESTAT;
RUN;
%IF (%UPCASE(&LIMIT)=NO) %THEN %DO;

  DATA &OUTFILE;
    SET BRR_TEMP6;
  RUN;

%END;
%ELSE %DO;
  %DO M=1 %TO 5;
  PROC FREQ DATA=BRRDATA NOPRINT;
    TABLE &PV_ROOT&M / OUT=BRR_TEMP7;
    BY &BYVAR;
  RUN;
  DATA BRR_TEMP8;
    SET BRR_TEMP7;
    COUNT&M=COUNT;
    &PV_ROOT=&PV_ROOT&M;
    KEEP &BYVAR &PV_ROOT COUNT&M;
  RUN;

```



Box 17.4 [3/4] SAS® macro of PROC_FREQ_PV.sas

```

%IF (&M=1) %THEN %DO;
    DATA BRR_TEMP9;
        SET BRR_TEMP8;
    RUN;
%END;
%ELSE %DO;
    DATA BRR_TEMP9;
        MERGE BRR_TEMP9 BRR_TEMP8;
        BY &BYVAR &PV_ROOT;
    RUN;
%END;
%END;
%LET FLAG_STUD=%SCAN(&LIMIT_CRITERIA,1);
DATA BRR_TEMP9;
    SET BRR_TEMP9;
    COUNT=(COUNT1+COUNT2+COUNT3+COUNT4+COUNT5)/5;
    IF (COUNT < &FLAG_STUD) THEN FLAG_STUD=1;
    ELSE FLAG_STUD=0;
    KEEP &BYVAR &PV_ROOT FLAG_STUD;
RUN;
PROC SORT DATA=BRRDATA;
    BY &BYVAR &ID_SCHOOL;
RUN;
%DO M=1 %TO 5;
PROC FREQ DATA=BRRDATA NOPRINT;
    TABLE &PV_ROOT&M /OUT=BRR_TEMP10;
    BY &BYVAR &ID_SCHOOL;
RUN;
PROC FREQ DATA=BRR_TEMP10 NOPRINT;
    TABLE &PV_ROOT&M /OUT=BRR_TEMP11;
    BY &BYVAR;
RUN;
DATA BRR_TEMP12;
    SET BRR_TEMP11;
    COUNT&M=COUNT;
    &PV_ROOT=&PV_ROOT&M;
    KEEP &BYVAR &PV_ROOT COUNT&M;
RUN;
%IF (&M=1) %THEN %DO;
    DATA BRR_TEMP13;
        SET BRR_TEMP12;
    RUN;
%END;
%ELSE %DO;
    DATA BRR_TEMP13;
        MERGE BRR_TEMP13 BRR_TEMP12;
        BY &BYVAR &PV_ROOT;
    RUN;
%END;
%END;
%LET FLAG_SCH=%SCAN(&LIMIT_CRITERIA,2);
DATA BRR_TEMP13;
    SET BRR_TEMP13;
    COUNT=(COUNT1+COUNT2+COUNT3+COUNT4+COUNT5)/5;
    IF (COUNT < &FLAG_SCH) THEN FLAG_SCH=1;
    ELSE FLAG_SCH=0;
    KEEP &BYVAR &PV_ROOT FLAG_SCH;
RUN;
%DO M=1 %TO 5;
PROC FREQ DATA=BRRDATA NOPRINT;
    TABLE &PV_ROOT&M/OUT=BRR_TEMP14;
    BY &BYVAR;
    WEIGHT &REPLI_ROOT.0;
RUN;
DATA BRR_TEMP15;
    SET BRR_TEMP14;
    COUNT&M=COUNT;
    &PV_ROOT=&PV_ROOT&M;
    KEEP &BYVAR &PV_ROOT COUNT&M;
RUN;

```



Box 17.4 [4/4] SAS® macro of PROC_FREQ_PV.sas

```

%IF (&M=1) %THEN %DO;
    DATA BRR_TEMP16;
        SET BRR_TEMP15;
    RUN;
%END;
%ELSE %DO;
DATA BRR_TEMP16;
    MERGE BRR_TEMP16 BRR_TEMP15;
        BY &BYVAR &PV_ROOT;
    RUN;
%END;
%END;

%LET K=1;
%LET POPREF=;
%LET NBVAR=%SCAN(&LIMIT_CRITERIA,4);
%DO %WHILE(&K <= &NBVAR);
    %LET POPREF_ADD=%SCAN(&BYVAR,&K);
    %LET POPREF=&POPREF &POPREF_ADD;
    %LET K=%EVAL(&K+1);
%END;

PROC MEANS DATA=BRR_TEMP16 NOPRINT;
    VAR COUNT1-COUNT5;
    BY &POPREF;
    OUTPUT OUT=BRR_TEMP17 SUM=SOMWGT1-SOMWGT5;
RUN;
%LET FLAG_PCT=%SCAN(&LIMIT_CRITERIA,3);
DATA BRR_TEMP18;
    MERGE BRR_TEMP16 BRR_TEMP17;
    BY &POPREF;
    ARRAY B1 (5) PCT1-PCT5;
    ARRAY B2 (5) COUNT1-COUNT5;
    ARRAY B3 (5) SOMWGT1-SOMWGT5;
    DO A=1 TO 5;
        B1(A) = (B2(A)/B3(A))*100;
    END;
    PCT=(PCT1+PCT2+PCT3+PCT4+PCT5)/5;
    IF (PCT < &FLAG_PCT) THEN FLAG_PCT=1;
    ELSE FLAG_PCT=0;
    KEEP &BYVAR &PV_ROOT FLAG_PCT;
RUN;

DATA &OUTFILE;
    MERGE BRR_TEMP6 BRR_TEMP9 BRR_TEMP13 BRR_TEMP18;
    BY &BYVAR &PV_ROOT;
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
DELETE BRR_TEMP7 BRR_TEMP8 BRR_TEMP9 BRR_TEMP10 BRR_TEMP11 BRR_TEMP12 BRR_TEMP13
        BRR_TEMP14 BRR_TEMP15 BRR_TEMP16 BRR_TEMP17 BRR_TEMP18;
RUN;

%END;

PROC DATASETS LIBRARY=WORK NOLIST;
DELETE BRR_TEMP0 BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5 FREQ_TEMP
        BRR_TEMP6 MEAN_TEMP BRRDATA;
RUN;

OPTIONS NOTES;

%MEND BRR_FREQ_PV;

```




Box 17.5 [1/3] SAS® macro of PROC_REG_NO_PV.sas

```

%macro BRR_REG(INFILE=,
               REPLI_ROOT=,
               VARDEP=,
               EXPLICA=,
               BYVAR=,
               LIMIT=,
               LIMIT_CRITERIA=,
               ID_SCHOOL=,
               OUTFILE=);

/*
MEANING OF THE MACRO ARGUMENTS

INFILE = INPUT DATA FILE.
REPLI_ROOT = ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL
WEIGHT VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
VARDEP = DEPENDENT VARIABLE
EXPLICA = LIST OF INDEPENDENT VARIABLES
BYVAR = BREAKDOWN VARIABLES
LIMIT = FLAGGING YES OR NO
LIMIT_CRITERIA = 1) NUMBER OF STUDENTS 2) NUMBER OF SCHOOLS 3) PERCENTAGE OF
STUDENTS AND 4) NUMBER OF VARIABLES FROM THE BYVAR ARGUMENT FOR DEFINING THE
POPULATION OF REFERENCE.
ID_SCHOOL = VARIABLE NAME FOR THE SCHOOL IDENTIFICATION
OUTFILE = FILE WITH THE STATISTIC ESTIMATES AND THEIR STANDARD ERROR ESTIMATES.
*/
OPTIONS NONOTES;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP1;
RUN;
PROC SORT DATA=&INFILE OUT=BRRDATA (KEEP=&REPLI_ROOT.0-&REPLI_ROOT.80 &BYVAR
&VARDEP &EXPLICA &ID_SCHOOL);
    BY &BYVAR;
RUN;
%DO I = 0 %TO 80;

    PROC REG DATA=BRRDATA OUTEST=COEF_TEMP NOPRINT EDF;
        MODEL &VARDEP=&EXPLICA;
        WEIGHT &REPLI_ROOT&I;
        BY &BYVAR;
    RUN;
    DATA COEF_TEMP;
        SET COEF_TEMP;
        L=&I;
    RUN;
    PROC APPEND BASE = BRR_TEMP1 DATA=COEF_TEMP;
    RUN;
%END;
PROC SORT DATA=BRR_TEMP1;
    BY &BYVAR L;
RUN;
PROC TRANSPOSE DATA=BRR_TEMP1 OUT=BRR_TEMP;
    BY &BYVAR L;
    VAR INTERCEPT &EXPLICA _RSQ_;
RUN;
DATA BRR_TEMP1 ;
    SET BRR_TEMP (RENAME=( _NAME_ =CLASS COL1=COEF));
RUN;
%LET I=1;
%DO %WHILE(%LENGTH(%SCAN(&EXPLICA, &I)));
    %LET I=%EVAL(&I+1);
%END;

```



Box 17.5 [2/3] SAS® macro of PROC_REG_NO_PV.sas

```

%LET NB=%EVAL(&I-1);
%DO J=0 %to &NB;
  %IF &J=0 %THEN %DO;
    %LET INDEP=Intercept;
  %END;
  %IF &J>0 %THEN %DO;
    %LET INDEP=%SCAN(&EXPLICA,&J);
  %END;
  DATA BRR_TEMP1;
    SET BRR_TEMP1;
    IF (UPCASE(CLASS)=UPCASE("&INDEP")) THEN ORDRE=&J;
    IF (SUBSTR(CLASS,1,5)="_RSQ_") THEN ORDRE=&I;
  RUN;
%END;
DATA BRR_TEMP2 (KEEP=&BYVAR CLASS COEF) BRR_TEMP3 (KEEP=&BYVAR CLASS ORDRE STAT);
  SET BRR_TEMP1;
  IF L > 0 THEN OUTPUT BRR_TEMP2;
  ELSE DO;
    STAT =COEF;
    OUTPUT BRR_TEMP3;
  END;
RUN;
PROC SORT DATA=BRR_TEMP2;
  BY &BYVAR CLASS;
RUN;
PROC SORT DATA=BRR_TEMP3;
  BY &BYVAR CLASS;
RUN;
DATA BRR_TEMP4;
  MERGE BRR_TEMP2 BRR_TEMP3;
  BY &BYVAR CLASS;
  VARI=((COEF-STAT)**2)*(1/20);
RUN;
PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
  VAR VARI;
  BY &BYVAR CLASS;
  OUTPUT OUT=BRR_TEMP5 SUM=SS;
RUN;
DATA BRR_TEMP6;
  MERGE BRR_TEMP3 BRR_TEMP5;
  BY &BYVAR CLASS;
  SESTAT=(SS)**0.5;
  FORMAT STAT F10.2;
  FORMAT SESTAT F10.2;
  KEEP &BYVAR CLASS ORDRE STAT SESTAT;
RUN;
PROC SORT DATA=BRR_TEMP6 OUT=&OUTFILE (DROP=ORDRE);
  BY &BYVAR ORDRE;
RUN;
%IF (%UPCASE(&LIMIT)=YES) %THEN %DO;

  DATA BRR_TEMP7;
    SET BRRDATA;
    NB_MISS=0;
    ARRAY LIST_VAR (&I) &EXPLICA &VARDEP;
    DO K=1 TO &I;
      IF (LIST_VAR(K) IN (.,.I,.M,.N)) THEN NB_MISS=NB_MISS+1;
    END;
    IF (NB_MISS>1) THEN NB_MISS=1;
  RUN;
  PROC FREQ DATA=BRR_TEMP7 NOPRINT;
    TABLE NB_MISS /OUT=BRR_TEMP8;
    BY &BYVAR;
    WHERE (NB_MISS=0);
  RUN;

```



Box 17.5 [3/3] SAS® macro of PROC_REG_NO_PV.sas

```

%LET FLAG_STUD=%SCAN(&LIMIT_CRITERIA,1);
DATA BRR_TEMP8;
    SET BRR_TEMP8;
    IF (COUNT < &FLAG_STUD) THEN FLAG_STUD=1;
    ELSE FLAG_STUD=0;
    KEEP &BYVAR FLAG_STUD;
RUN;
PROC SORT DATA=BRR_TEMP7;
    BY &BYVAR &ID_SCHOOL;
RUN;
PROC FREQ DATA=BRR_TEMP7 NOPRINT;
    TABLE NB_MISS /OUT=BRR_TEMP9;
    BY &BYVAR &ID_SCHOOL;
    WHERE (NB_MISS=0);
RUN;
PROC FREQ DATA=BRR_TEMP9 NOPRINT;
    TABLE NB_MISS /OUT=BRR_TEMP10;
    BY &BYVAR;
RUN;
%LET FLAG_SCH=%SCAN(&LIMIT_CRITERIA,2);
DATA BRR_TEMP10;
    SET BRR_TEMP10;
    IF (COUNT < &FLAG_SCH) THEN FLAG_SCH=1;
    ELSE FLAG_SCH=0;
    KEEP &BYVAR FLAG_SCH;
RUN;
PROC SORT DATA=BRR_TEMP7;
    BY &BYVAR NB_MISS;
RUN;
PROC FREQ DATA=BRR_TEMP7 NOPRINT;
    TABLE NB_MISS/OUT=BRR_TEMP11;
    BY &BYVAR;
    WEIGHT &REPLI_ROOT.0;
RUN;
%LET K=1;
%LET POPREF=;
%LET NBVAR=%SCAN(&LIMIT_CRITERIA,4);
%DO %WHILE(&K <= &NBVAR);
    %LET POPREF_ADD=%SCAN(&BYVAR,&K);
    %LET POPREF=&POPREF &POPREF_ADD;
    %LET K=%EVAL(&K+1);
%END;
PROC MEANS DATA=BRR_TEMP11 NOPRINT;
    VAR COUNT;
    BY &POPREF;
    OUTPUT OUT=BRR_TEMP12 SUM=SOMWGT;
RUN;
%LET FLAG_PCT=%SCAN(&LIMIT_CRITERIA,3);
DATA BRR_TEMP13;
    MERGE BRR_TEMP11 BRR_TEMP12;
    BY &POPREF;
    PCT=(COUNT/SOMWGT)*100;
    IF (PCT < &FLAG_PCT) THEN FLAG_PCT=1;
    ELSE FLAG_PCT=0;
    IF (NB_MISS=0);
    KEEP &BYVAR FLAG_PCT;
RUN;
DATA &OUTFILE._CRITERIA;
    MERGE BRR_TEMP8 BRR_TEMP10 BRR_TEMP13;
    BY &BYVAR;
RUN;
PROC DATASETS LIBRARY=WORK NOLIST;
DELETE BRR_TEMP7 BRR_TEMP8 BRR_TEMP9 BRR_TEMP10 BRR_TEMP11 BRR_TEMP12 BRR_TEMP13;
RUN;
%END;
PROC DATASETS LIBRARY=WORK NOLIST;
DELETE BRR_TEMP BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5 BRR_TEMP6
COEF_TEMP BRRDATA;
RUN;
OPTIONS NOTES;
%MEND BRR_REG;

```



Box 17.6 [1/4] SAS® macro of PROC_REG_PV.sas

```

%macro BRR_REG_PV(INFILE=,
                 REPLI_ROOT=,
                 EXPLICA=,
                 BYVAR=,
                 PV_ROOT=,
                 LIMIT=,
                 LIMIT_CRITERIA=,
                 ID_SCHOOL=,
                 OUTFILE=);

/*
MEANING OF THE MACRO ARGUMENTS
INFILE = INPUT DATA FILE.
REPLI_ROOT = ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL
WEIGHT VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
EXPLICA = LIST OF INDEPENDENT VARIABLES
BYVAR = BREAKDOWN VARIABLES
PV_ROOT = ROOT OF THE 5 PLAUSIBLE VALUES VARIABLES NAMES
LIMIT = FLAGGING YES OR NO
LIMIT_CRITERIA = 1) NUMBER OF STUDENTS 2) NUMBER OF SCHOOLS 3) PERCENTAGE OF
STUDENTS AND 4) NUMBER OF VARIABLES FROM THE BYVAR ARGUMENT FOR DEFINING THE
POPULATION OF REFERENCE
ID_SCHOOL = VARIABLE NAME FOR THE SCHOOL IDENTIFICATION
OUTFILE = FILE WITH THE STATISTIC ESTIMATES AND THEIR STANDARD ERROR ESTIMATES.
*/
OPTIONS NONOTES;
PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP1;
RUN;
PROC SORT DATA=&INFILE OUT=BRRDATA (KEEP=&REPLI_ROOT.0-&REPLI_ROOT.80 &BYVAR &PV_
ROOT.1-&PV_ROOT.5 &EXPLICA &ID_SCHOOL);
    BY &BYVAR;
RUN;
%DO I = 0 %TO 80;
    PROC REG DATA=BRRDATA OUTEST=COEF_TEMP NOPRINT EDF;
        MODEL &PV_ROOT.1=&EXPLICA;
        MODEL &PV_ROOT.2=&EXPLICA;
        MODEL &PV_ROOT.3=&EXPLICA;
        MODEL &PV_ROOT.4=&EXPLICA;
        MODEL &PV_ROOT.5=&EXPLICA;
        WEIGHT &REPLI_ROOT&I;
        BY &BYVAR;

    RUN;
    DATA COEF_TEMP;
        SET COEF_TEMP;
        L=&I;

    RUN;
    PROC APPEND BASE = BRR_TEMP1 DATA=COEF_TEMP;
    RUN;

%END;
PROC SORT DATA=BRR_TEMP1;
    BY &BYVAR L;
RUN;
PROC TRANSPOSE DATA=BRR_TEMP1 OUT=BRR_TEMP PREFIX=PV;
    BY &BYVAR L;
    VAR INTERCEPT &EXPLICA _RSQ_;
RUN;
DATA BRR_TEMP1 ;
    SET BRR_TEMP (RENAME=( _NAME_ =CLASS));
    DROP _LABEL_;
RUN;

```



Box 17.6 [2/4] SAS® macro of PROC_REG_PV.sas

```

%LET I=1;
%DO %WHILE(%LENGTH(%SCAN(&EXPLICA,&I)));
    %LET I=%EVAL(&I+1);
%END;
%let NB=%EVAL(&I-1);

%DO J=0 %to &NB;
    %IF &J=0 %THEN %DO;
        %LET INDEP=Intercept;
    %END;
    %IF &J>0 %THEN %DO;
        %LET INDEP=%SCAN(&EXPLICA,&J);
    %END;

    DATA BRR_TEMP1;
        SET BRR_TEMP1;
        IF (UPCASE(CLASS)=UPCASE("&INDEP")) THEN ORDRE=&J;
        IF (SUBSTR(CLASS,1,5)="_RSQ_") THEN ORDRE=&I;
    RUN;

%END;
DATA BRR_TEMP2 (KEEP=&BYVAR CLASS ORDRE PV1-PV5) BRR_TEMP3 (KEEP=&BYVAR CLASS ORDRE
FIN1-FIN5 STAT MESVAR);
    SET BRR_TEMP1;
    IF L > 0 THEN OUTPUT BRR_TEMP2;
    ELSE DO;
        STAT = (PV1+PV2+PV3+PV4+PV5) / 5;
        FIN1=PV1;
        FIN2=PV2;
        FIN3=PV3;
        FIN4=PV4;
        FIN5=PV5;
        MESVAR=(((STAT-FIN1)**2)+((STAT-FIN2)**2)+((STAT-FIN3)**2)+((STAT-
FIN4)**2)+((STAT-FIN5)**2))/4;
        OUTPUT BRR_TEMP3;
    END;
RUN;
PROC SORT DATA=BRR_TEMP2;
    BY &BYVAR CLASS;
RUN;
PROC SORT DATA=BRR_TEMP3;
    BY &BYVAR CLASS;
RUN;
DATA BRR_TEMP4;
    MERGE BRR_TEMP2 BRR_TEMP3;
    BY &BYVAR CLASS;
    ARRAY A (5)
        PV1-PV5;
    ARRAY B (5)
        FIN1-FIN5;
    ARRAY C (5)
        VAR1-VAR5;
    DO I=1 TO 5;
        C(I)=(1/20)*(A(I)-B(I)**2);
    END;
RUN;
PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
    VAR VAR1 VAR2 VAR3 VAR4 VAR5;
    BY &BYVAR CLASS;
    OUTPUT OUT=BRR_TEMP5 SUM=SS1 SS2 SS3 SS4 SS5;
RUN;
DATA BRR_TEMP6;
    MERGE BRR_TEMP3 BRR_TEMP5;
    BY &BYVAR CLASS;
    SAMP=(SS1+SS2+SS3+SS4+SS5)/5;
    FINVAR=(SAMP+(1.2*MESVAR));
    SESTAT=(FINVAR)**0.5;
    FORMAT STAT F10.2;
    FORMAT SESTAT F10.2;
    KEEP &BYVAR CLASS STAT SESTAT ORDRE;
RUN;

```

Box 17.6 [3/4] SAS® macro of PROC_REG_PV.sas

```

PROC SORT DATA=BRR_TEMP6 OUT=&OUTFILE (DROP=ORDRE);
  BY &BYVAR ORDRE;
RUN;

%IF (%UPCASE(&LIMIT)=YES) %THEN %DO;

  %LET NB_M=%EVAL(&NB+5);

  DATA BRR_TEMP7;
    SET BRRDATA;
    NB_MISS=0;
    ARRAY LIST_VAR (&NB_M) &EXPLICA &PV_ROOT.1-&PV_ROOT.5;
    DO K=1 TO &NB;
      IF (LIST_VAR(K) IN (.,.I,.M,.N)) THEN NB_MISS=NB_MISS+1;
    END;
    IF (NB_MISS>1) THEN NB_MISS=1;
  RUN;
  PROC FREQ DATA=BRR_TEMP7 NOPRINT;
    TABLE NB_MISS /OUT=BRR_TEMP8;
    BY &BYVAR;
    WHERE (NB_MISS=0);
  RUN;
  %LET FLAG_STUD=%SCAN(&LIMIT_CRITERIA,1);
  DATA BRR_TEMP8;
    SET BRR_TEMP8;
    IF (COUNT < &FLAG_STUD) THEN FLAG_STUD=1;
    ELSE FLAG_STUD=0;
    KEEP &BYVAR FLAG_STUD;
  RUN;
  PROC SORT DATA=BRR_TEMP7;
    BY &BYVAR &ID_SCHOOL;
  RUN;
  PROC FREQ DATA=BRR_TEMP7 NOPRINT;
    TABLE NB_MISS /OUT=BRR_TEMP9;
    BY &BYVAR &ID_SCHOOL;
    WHERE (NB_MISS=0);
  RUN;
  PROC FREQ DATA=BRR_TEMP9 NOPRINT;
    TABLE NB_MISS /OUT=BRR_TEMP10;
    BY &BYVAR;
  RUN;
  %LET FLAG_SCH=%SCAN(&LIMIT_CRITERIA,2);
  DATA BRR_TEMP10;
    SET BRR_TEMP10;
    IF (COUNT < &FLAG_SCH) THEN FLAG_SCH=1;
    ELSE FLAG_SCH=0;
    KEEP &BYVAR FLAG_SCH;
  RUN;
  PROC SORT DATA=BRR_TEMP7;
    BY &BYVAR NB_MISS;
  RUN;
  PROC FREQ DATA=BRR_TEMP7 NOPRINT;
    TABLE NB_MISS/OUT=BRR_TEMP11;
    BY &BYVAR;
    WEIGHT &REPLI_ROOT.0;
  RUN;
  %LET K=1;
  %LET POPREF=;
  %LET NBVAR=%SCAN(&LIMIT_CRITERIA,4);
  %DO %WHILE(&K <= &NBVAR);
    %LET POPREF_ADD=%SCAN(&BYVAR,&K);
    %LET POPREF=&POPREF &POPREF_ADD;
    %LET K=%EVAL(&K+1);
  %END;

  PROC MEANS DATA=BRR_TEMP11 NOPRINT;
    VAR COUNT;
    BY &POPREF;
    OUTPUT OUT=BRR_TEMP12 SUM=SOMWGT;
  RUN;

```



Box 17.6 [4/4] SAS® macro of PROC_REG_PV.sas

```

%LET FLAG_PCT=%SCAN(&LIMIT_CRITERIA,3);
DATA BRR_TEMP13;
    MERGE BRR_TEMP11 BRR_TEMP12;
    BY &POPREF;
    PCT=(COUNT/SOMWGT)*100;
    IF (PCT < &FLAG_PCT) THEN FLAG_PCT=1;
    ELSE FLAG_PCT=0;
    IF (NB_MISS=0);
    KEEP &BYVAR FLAG_PCT;
RUN;
DATA &OUTFILE._CRITERIA;
    MERGE BRR_TEMP8 BRR_TEMP10 BRR_TEMP13;
    BY &BYVAR;
RUN;
PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP7 BRR_TEMP8 BRR_TEMP9 BRR_TEMP10 BRR_TEMP11 BRR_
TEMP12 BRR_TEMP13;
RUN;
%END;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5 BRR_TEMP6
COEF_TEMP BRRDATA;
RUN;

OPTIONS NOTES;

%MEND BRR_REG_PV;

```



Box 17.7 [1/3] SAS® macro of PROC_CORR_NO_PV.sas

```

%macro BRR_CORR(INFILE=,
                REPLI_ROOT=,
                BYVAR=,
                VAR1=,
                VAR2=,
                LIMIT=,
                LIMIT_CRITERIA=,
                ID_SCHOOL=,
                OUTFILE=);

/*
MEANING OF THE MACRO ARGUMENTS

INFILE = INPUT DATA FILE.
REPLI_ROOT = ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL
WEIGHT VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
BYVAR = BREAKDOWN VARIABLES.
VAR1 = FIRST NUMERIC VARIABLE.
VAR2 = SECOND NUMERIC VARIABLE.
LIMIT = FLAGGING YES OR NO.
LIMIT_CRITERIA = 1) NUMBER OF STUDENTS 2) NUMBER OF SCHOOLS 3) PERCENTAGE OF
STUDENTS AND 4) NUMBER OF VARIABLES FROM THE BYVAR ARGUMENT FOR DEFINING THE
POPULATION OF REFERENCE.
ID_SCHOOL = VARIABLE NAME FOR THE SCHOOL IDENTIFICATION.
OUTFILE = FILE WITH THE STATISTIC ESTIMATES AND THEIR RESPECTIVE STANDARD ERROR ESTIMATE.

*/

OPTIONS NONOTES;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP1;
RUN;

PROC SORT DATA=&INFILE OUT=BRRDATA (KEEP=&REPLI_ROOT.0-&REPLI_ROOT.80 &BYVAR &VAR1
&VAR2 &ID_SCHOOL);
    BY &BYVAR;
RUN;

%DO I = 0 %TO 80;

    PROC CORR DATA=BRRDATA VARDEF=WGT NOPRINT OUTP=CORR_TEMP;
        VAR &VAR1 ;
        WITH &VAR2;
        BY &BYVAR;
        WEIGHT &REPLI_ROOT&I;

    RUN;

    DATA CORR_TEMP;
        SET CORR_TEMP;
        L=&I;

    RUN;

    PROC APPEND BASE = BRR_TEMP1 DATA=CORR_TEMP;
    RUN;

%END;

DATA BRR_TEMP1;
    SET BRR_TEMP1;
    IF ( _TYPE_ NE "CORR") THEN DELETE;
    PV=&VAR1;
    KEEP &BYVAR L PV;
RUN;

DATA BRR_TEMP2 (KEEP=&BYVAR PV) BRR_TEMP3 (KEEP=&BYVAR STAT);
    SET BRR_TEMP1;
    IF L > 0 THEN OUTPUT BRR_TEMP2;
    ELSE DO;
        STAT =PV;
        OUTPUT BRR_TEMP3;
    END;
RUN;

```




Box 17.7 [2/3] SAS® macro of PROC CORR_NO_PV.sas

```

PROC SORT DATA=BRR_TEMP2;
  BY &BYVAR;
RUN;

PROC SORT DATA=BRR_TEMP3;
  BY &BYVAR;
RUN;

DATA BRR_TEMP4;
  MERGE BRR_TEMP2 BRR_TEMP3;
  BY &BYVAR;
  VARI=((PV-STAT)**2)*(1/20);
RUN;

PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
  VAR VARI;
  BY &BYVAR;
  OUTPUT OUT=BRR_TEMP5 SUM=SS;
RUN;

DATA BRR_TEMP6;
  MERGE BRR_TEMP3 BRR_TEMP5;
  BY &BYVAR;
  SESTAT=(SS)**0.5;
  FORMAT STAT F5.2;
  FORMAT SESTAT F5.2;
  KEEP &BYVAR STAT SESTAT;
RUN;

%IF (%UPCASE(&LIMIT)=NO) %THEN %DO;

  DATA &OUTFILE;
    SET BRR_TEMP6;
  RUN;

%END;
%ELSE %DO;
  DATA BRR_TEMP7;
    SET BRRDATA;
    NB_MISS=0;
    ARRAY LIST_VAR (2) &VAR1 &VAR2;
    DO K=1 TO 2;
      IF (LIST_VAR(K) IN (.,.I, .M, .N)) THEN NB_MISS=NB_MISS+1;
    END;
    IF (NB_MISS>1) THEN NB_MISS=1;
  RUN;
  PROC FREQ DATA=BRR_TEMP7 NOPRINT;
    TABLE NB_MISS /OUT=BRR_TEMP8;
    BY &BYVAR;
    WHERE (NB_MISS=0);
  RUN;
  %LET FLAG_STUD=%SCAN(&LIMIT_CRITERIA,1);
  DATA BRR_TEMP8;
    SET BRR_TEMP8;
    IF (COUNT < &FLAG_STUD) THEN FLAG_STUD=1;
    ELSE FLAG_STUD=0;
    KEEP &BYVAR FLAG_STUD;
  RUN;
  PROC SORT DATA=BRR_TEMP7;
    BY &BYVAR &ID_SCHOOL;
  RUN;
  PROC FREQ DATA=BRR_TEMP7 NOPRINT;
    TABLE NB_MISS /OUT=BRR_TEMP9;
    BY &BYVAR &ID_SCHOOL;
    WHERE (NB_MISS=0);
  RUN;
  PROC FREQ DATA=BRR_TEMP9 NOPRINT;
    TABLE NB_MISS /OUT=BRR_TEMP10;
    BY &BYVAR;
  RUN;

```



Box 17.7 [3/3] SAS® macro of PROC_CORR_NO_PV.sas

```

%LET FLAG_SCH=%SCAN(&LIMIT_CRITERIA,2);
DATA BRR_TEMP10;
    SET BRR_TEMP10;
    IF (COUNT < &FLAG_SCH) THEN FLAG_SCH=1;
    ELSE FLAG_SCH=0;
    KEEP &BYVAR FLAG_SCH;
RUN;
PROC SORT DATA=BRR_TEMP7;
    BY &BYVAR NB_MISS;
RUN;
PROC FREQ DATA=BRR_TEMP7 NOPRINT;
    TABLE NB_MISS/OUT=BRR_TEMP11;
    BY &BYVAR;
    WEIGHT &REPLI_ROOT.0;
RUN;
%LET K=1;
%LET POPREF=;
%LET NBVAR=%SCAN(&LIMIT_CRITERIA,4);
%DO %WHILE(&K <= &NBVAR);
    %LET POPREF_ADD=%SCAN(&BYVAR,&K);
    %LET POPREF=&POPREF &POPREF_ADD;
    %LET K=%EVAL(&K+1);
%END;

PROC MEANS DATA=BRR_TEMP11 NOPRINT;
    VAR COUNT;
    BY &POPREF;
    OUTPUT OUT=BRR_TEMP12 SUM=SOMWGT;
RUN;
%LET FLAG_PCT=%SCAN(&LIMIT_CRITERIA,3);
DATA BRR_TEMP13;
    MERGE BRR_TEMP11 BRR_TEMP12;
    BY &POPREF;
    PCT=(COUNT/SOMWGT)*100;
    IF (PCT < &FLAG_PCT) THEN FLAG_PCT=1;
    ELSE FLAG_PCT=0;
    IF (NB_MISS=0);
    KEEP &BYVAR FLAG_PCT;
RUN;
DATA &OUTFILE;
    MERGE BRR_TEMP6 BRR_TEMP8 BRR_TEMP10 BRR_TEMP13;
    BY &BYVAR;
RUN;
PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP7 BRR_TEMP8 BRR_TEMP9 BRR_TEMP10 BRR_TEMP11 BRR_
TEMP12 BRR_TEMP13;
RUN;

%END;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5 BRR_TEMP6 CORR_
TEMP BRRDATA;
RUN;

OPTIONS NOTES;

%MEND BRR_CORR;

```



Box 17.8 [1/3] SAS® macro of PROC_CORR_PV.sas

```

%macro BRR_CORR_PV(INFILE=,
                   REPLI_ROOT=,
                   BYVAR=,
                   EXPLICA=,
                   PV_ROOT=,
                   LIMIT=,
                   LIMIT_CRITERIA=,
                   ID_SCHOOL=,
                   OUTFILE=);

/*

MEANING OF THE MACRO ARGUMENTS

INFILE = INPUT DATA FILE.
REPLI_ROOT = ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL
WEIGHT VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
BYVAR = BREAKDOWN VARIABLES.
EXPLICA = NUMERIC VARIABLE.
PV_ROOT = ROOT OF THE 5 PLAUSIBLE VALUES VARIABLES NAMES
LIMIT = FLAGGING YES OR NO.
LIMIT_CRITERIA = 1) NUMBER OF STUDENTS 2) NUMBER OF SCHOOLS 3) PERCENTAGE OF
STUDENTS AND 4) NUMBER OF VARIABLES FROM THE BYVAR ARGUMENT FOR DEFINING THE
POPULATION OF REFERENCE.
ID_SCHOOL = VARIABLE NAME FOR THE SCHOOL IDENTIFICATION.
OUTFILE = FILE WITH THE STATISTIC ESTIMATES AND THEIR RESPECTIVE STANDARD ERROR ESTIMATE.

*/

OPTIONS NONOTES;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP1;
RUN;

PROC SORT DATA=&INFILE OUT=BRRDATA(KEEP=&REPLI_ROOT.0-&REPLI_ROOT.80 &BYVAR &PV_
ROOT.1-&PV_ROOT.5 &EXPLICA &ID_SCHOOL);
    BY &BYVAR;
RUN;

%DO I = 0 %TO 80;
    PROC CORR DATA=BRRDATA VARDEF=WGT NOPRINT OUTP=CORR_TEMP;
        VAR &PV_ROOT.1 - &PV_ROOT.5 ;
        WITH &EXPLICA;
    BY &BYVAR;
    WEIGHT &REPLI_ROOT&I;
RUN;

    DATA CORR_TEMP;
        SET CORR_TEMP;
        L=&I;
    RUN;

    PROC APPEND BASE = BRR_TEMP1 DATA=CORR_TEMP;
    RUN;

%END;

DATA BRR_TEMP1;
    SET BRR_TEMP1;
    IF (_TYPE_ NE "CORR") THEN DELETE;
    PV1=&PV_ROOT.1;
    PV2=&PV_ROOT.2;
    PV3=&PV_ROOT.3;
    PV4=&PV_ROOT.4;
    PV5=&PV_ROOT.5;
    KEEP &BYVAR L PV1-PV5;
RUN;

```

Box 17.8 [2/3] SAS® macro of PROC_CORR_PV.sas

```

DATA BRR_TEMP2 (DROP=STAT FIN1-FIN5 MESVAR) BRR_TEMP3 (KEEP=&BYVAR STAT FIN1-FIN5
MESVAR);
  SET BRR_TEMP1;
  IF L > 0 THEN OUTPUT BRR_TEMP2;
  ELSE DO;
    STAT = (PV1+PV2+PV3+PV4+PV5) / 5;
    FIN1=PV1;
    FIN2=PV2;
    FIN3=PV3;
    FIN4=PV4;
    FIN5=PV5;
    MESVAR=(((STAT-FIN1)**2)+((STAT-FIN2)**2)+((STAT-FIN3)**2)+((STAT-
FIN4)**2)+((STAT-FIN5)**2))/4;
    OUTPUT BRR_TEMP3;
  END;
RUN;
PROC SORT DATA=BRR_TEMP2;
  BY &BYVAR;
RUN;
PROC SORT DATA=BRR_TEMP3;
  BY &BYVAR;
RUN;
DATA BRR_TEMP4;
  MERGE BRR_TEMP2 BRR_TEMP3;
  BY &BYVAR;
  ARRAY A (5)
    PV1-PV5;
  ARRAY B (5)
    FIN1-FIN5;
  ARRAY C (5)
    VAR1-VAR5;
  DO I=1 TO 5;
    C(I) = (1/20) * ((A(I) - B(I)) ** 2);
  END;
RUN;
PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
  VAR VAR1 VAR2 VAR3 VAR4 VAR5;
  BY &BYVAR;
  OUTPUT OUT=BRR_TEMP5 SUM=SS1 SS2 SS3 SS4 SS5;
RUN;
DATA BRR_TEMP6;
  MERGE BRR_TEMP3 BRR_TEMP5;
  BY &BYVAR;
  SAMP=(SS1+SS2+SS3+SS4+SS5) / 5;
  FINVAR=(SAMP+(1.2*MESVAR));
  SESTAT=(FINVAR)**0.5;
  FORMAT STAT F5.2;
  FORMAT SESTAT F5.2;
  KEEP &BYVAR STAT SESTAT ;
RUN;
%IF (%UPCASE(&LIMIT)=NO) %THEN %DO;

  DATA &OUTFILE;
    SET BRR_TEMP6;
  RUN;
%END;
%ELSE %DO;
  DATA BRR_TEMP7;
    SET BRRDATA;
    NB_MISS=0;
    ARRAY LIST VAR (6) &EXPLICA &PV_ROOT.1-&PV_ROOT.5;
    DO K=1 TO 6;
      IF (LIST_VAR(K) IN (.,.I,.M,.N)) THEN NB_MISS=NB_MISS+1;
    END;
    IF (NB_MISS>1) THEN NB_MISS=1;
  RUN;
  PROC FREQ DATA=BRR_TEMP7 NOPRINT;
    TABLE NB_MISS /OUT=BRR_TEMP8;
    BY &BYVAR;
    WHERE (NB_MISS=0);
  RUN;

```



Box 17.8 [3/3] SAS® macro of PROC_CORR_PV.sas

```

%LET FLAG_STUD=%SCAN(&LIMIT_CRITERIA,1);
DATA BRR_TEMP8;
    SET BRR_TEMP8;
    IF (COUNT < &FLAG_STUD) THEN FLAG_STUD=1;
    ELSE FLAG_STUD=0;
    KEEP &BYVAR FLAG_STUD;
RUN;
PROC SORT DATA=BRR_TEMP7;
    BY &BYVAR &ID_SCHOOL;
RUN;
PROC FREQ DATA=BRR_TEMP7 NOPRINT;
    TABLE NB_MISS /OUT=BRR_TEMP9;
    BY &BYVAR &ID_SCHOOL;
    WHERE (NB_MISS=0);
RUN;
PROC FREQ DATA=BRR_TEMP9 NOPRINT;
    TABLE NB_MISS /OUT=BRR_TEMP10;
    BY &BYVAR;
RUN;
%LET FLAG_SCH=%SCAN(&LIMIT_CRITERIA,2);
DATA BRR_TEMP10;
    SET BRR_TEMP10;
    IF (COUNT < &FLAG_SCH) THEN FLAG_SCH=1;
    ELSE FLAG_SCH=0;
    KEEP &BYVAR FLAG_SCH;
RUN;
PROC SORT DATA=BRR_TEMP7;
    BY &BYVAR NB_MISS;
RUN;
PROC FREQ DATA=BRR_TEMP7 NOPRINT;
    TABLE NB_MISS/OUT=BRR_TEMP11;
    BY &BYVAR;
    WEIGHT &REPLI_ROOT.0;
RUN;
%LET K=1;
%LET POPREF=;
%LET NBVAR=%SCAN(&LIMIT_CRITERIA,4);
%DO %WHILE(&K <= &NBVAR);
    %LET POPREF_ADD=%SCAN(&BYVAR,&K);
    %LET POPREF=&POPREF &POPREF_ADD;
    %LET K=%EVAL(&K+1);
%END;
PROC MEANS DATA=BRR_TEMP11 NOPRINT;
    VAR COUNT;
    BY &POPREF;
    OUTPUT OUT=BRR_TEMP12 SUM=SOMWGT;
RUN;
%LET FLAG_PCT=%SCAN(&LIMIT_CRITERIA,3);
DATA BRR_TEMP13;
    MERGE BRR_TEMP11 BRR_TEMP12;
    BY &POPREF;
    PCT=(COUNT/SOMWGT)*100;
    IF (PCT < &FLAG_PCT) THEN FLAG_PCT=1;
    ELSE FLAG_PCT=0;
    IF (NB_MISS=0);
    KEEP &BYVAR FLAG_PCT;
RUN;
DATA &OUTFILE;
    MERGE BRR_TEMP6 BRR_TEMP8 BRR_TEMP10 BRR_TEMP13;
    BY &BYVAR;
RUN;
PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP7 BRR_TEMP8 BRR_TEMP9 BRR_TEMP10 BRR_TEMP11 BRR_
TEMP12 BRR_TEMP13;
RUN;
%END;
PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5 BRR_TEMP6 CORR_
TEMP BRRDATA;
RUN;
OPTIONS NOTES;
%MEND BRR_CORR_PV;

```



Box 17.9 [1/3] SAS® macro of PROC_DIF_NO_PV.sas

```

%MACRO BRR_PROCMEAN_DIF(INFILE =,
                        REPLI_ROOT =,
                        BYVAR =,
                        VAR =,
                        COMPARE =,
                        CATEGORY =,
                        STAT =,
                        OUTFILE =);

/*
MEANING OF THE MACRO ARGUMENTS

INFILE = INPUT DATA FILE.
REPLI_ROOT = ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL
WEIGHT VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
BYVAR = BREAKDOWN VARIABLES.
VAR = VARIABLES FOR WHICH THE STATISTIC IS REQUESTED.
COMPARE = BREAKDOWN VARIABLE NAME FOR WHICH CATEGORY CONTRASTS ARE REQUESTED.
CATEGORY = LIST OF THE "COMPARE" VARIABLE CATEGORIES FOR WHICH A CONTRAST IS
REQUESTED.
STAT = REQUESTED STATISTIC.
SUMWGT = SUM OF THE WEIGHT
MEAN = MEAN
VAR = VARIANCE
STD = STANDARD DEVIATION
CV = COEFFICIENT OF VARIATION
SKEWNESS = SKEWNESS
KURTOSIS = KURTOSIS
MEDIAN = MEDIAN
Q1 = FIRST QUARTILE
Q3 = THIRD QUARTILE
QRANGE = RANGE BETWEEN Q1 AND Q3
PX = PERCENTILE, WITH X BETWEEN 1 AND 99
OUTFILE = FILE WITH THE STATISTIC ESTIMATES AND THEIR STANDARD ERROR ESTIMATES.

*/
OPTIONS NONOTES;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP0 BRR_TEMP1 ;
RUN;

PROC SORT DATA=&INFILE OUT=BRRDATA(KEEP=&REPLI_ROOT.0-&REPLI_ROOT.80 &BYVAR
&COMPARE &VAR);
  BY &BYVAR &COMPARE;
RUN;

%DO I = 0 %TO 80;

  PROC MEANS DATA=BRRDATA VARDEF=WGT NOPRINT;
    VAR &VAR ;
    BY &BYVAR &COMPARE;
    WEIGHT &REPLI_ROOT&I;
    OUTPUT OUT=MEAN_TEMP &STAT=PV;

  RUN;

  DATA MEAN_TEMP;
    SET MEAN_TEMP;
    L=&I;

  RUN;

  PROC APPEND BASE = BRR_TEMP0 DATA=MEAN_TEMP;
  RUN;

%END;

```



Box 17.9 [2/3] SAS® macro of PROC_DIF_NO_PV.sas

```

%LET DEBUT=1;
%LET SUIVANT=2;
%LET COMPTE=1;

%LET I=1;

%DO %WHILE(%LENGTH(%SCAN(&CATEGORY, &I)));
    %LET I=%EVAL(&I+1);
%END;

%LET NBCAT=%EVAL(&I-1);
%LET NBDIF=%EVAL((&NBCAT*(&NBCAT-1))/2);
%LET ORDRE=1;

%DO J=1 %TO &NBDIF;

    %DO K=&DEBUT %TO &NBCAT-1;

        %LET CAT1=%SCAN(&CATEGORY, &DEBUT);
        %LET CAT2=%SCAN(&CATEGORY, &SUIVANT);

        DATA BRR_DIF1;
            SET BRR_TEMP0 (RENAME=(PV=M1PV));
            LENGTH CONTRAST $5;
            CONTRAST="&CAT1.-&CAT2";
            IF (&COMPARE=&CAT1);
            KEEP &BYVAR L CONTRAST M1PV &compare;

        RUN;

        DATA BRR_DIF2 ;
            SET BRR_TEMP0 (RENAME=(PV=M2PV));
            IF (&COMPARE=&CAT2);
            KEEP &BYVAR L M2PV &compare;

        RUN;

        PROC SORT DATA=BRR_DIF1;
            BY &BYVAR L;

        RUN;

        PROC SORT DATA=BRR_DIF2;
            BY &BYVAR L;

        RUN;

        DATA BRR_TEMP_DIF;
            MERGE BRR_DIF1 BRR_DIF2;
            BY &BYVAR L;
            IF (M1PV EQ . OR M2PV EQ .) THEN DELETE;
            PV=M1PV-M2PV;
            ORDRE=&ORDRE;
            KEEP &BYVAR CONTRAST L PV ORDRE;

        RUN;

        PROC APPEND BASE = BRR_TEMP1 DATA=BRR_TEMP_DIF;
        RUN;

        %LET SUIVANT=%EVAL(&SUIVANT+1);
        %LET COMPTE=%EVAL(&COMPTE+1);
        %LET ORDRE=%EVAL(&ORDRE+1);

    %END;

    %LET DEBUT=%EVAL(&DEBUT+1);
    %LET SUIVANT=%EVAL(&DEBUT+1);

%END;

```



Box 17.9 [3/3] SAS® macro of PROC_DIF_NO_PV.sas

```

DATA BRR_TEMP2 (KEEP=&BYVAR CONTRAST PV) BRR_TEMP3 (KEEP=&BYVAR CONTRAST ORDRE STAT);
  SET BRR_TEMP1;
  IF L > 0 THEN OUTPUT BRR_TEMP2;
  ELSE DO;
    STAT =PV;
    OUTPUT BRR_TEMP3;
  END;
RUN;

PROC SORT DATA=BRR_TEMP2;
  BY &BYVAR CONTRAST;
RUN;

PROC SORT DATA=BRR_TEMP3;
  BY &BYVAR CONTRAST;
RUN;

DATA BRR_TEMP4;
  MERGE BRR_TEMP2 BRR_TEMP3;
  BY &BYVAR CONTRAST;
  VARI=((PV-STAT)**2)*(1/20);
RUN;

PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
  VAR VARI;
  BY &BYVAR CONTRAST;
  OUTPUT OUT=BRR_TEMP5 SUM=SS;
RUN;

DATA BRR_TEMP6;
  MERGE BRR_TEMP3 BRR_TEMP5;
  BY &BYVAR CONTRAST;
  SESTAT=(SS)**0.5;
  FORMAT STAT F10.2;
  FORMAT SESTAT F10.2;
  KEEP &BYVAR CONTRAST ORDRE STAT SESTAT;
RUN;

PROC SORT DATA=BRR_TEMP6 OUT=&OUTFILE (DROP=ORDRE);
  BY &BYVAR ORDRE;
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP0 BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5 BRR_TEMP6
  MEAN_TEMP BRRDATA BRR_DIF1 BRR_DIF2 BRR_TEMP_DIF;
RUN;

OPTIONS NOTES;

%MEND BRR_PROCMEAN_DIF;

```




Box 17.10 [1/3] SAS® macro of PROC_DIF_PV.sas

```

%MACRO BRR_PROCMEAN_DIF_PV(INFILE = ,
                           REPLI_ROOT = ,
                           BYVAR = ,
                           PV_ROOT = ,
                           COMPARE = ,
                           CATEGORY = ,
                           STAT = ,
                           OUTFILE = );

/*

MEANING OF THE MACRO ARGUMENTS

INFILE =          INPUT DATA FILE.
REPLI_ROOT =     ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL WEI-
GHT VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
BYVAR =          BREAKDOWN VARIABLES.
PV_ROOT =        ROOT OF THE 5 PLAUSIBLE VALUES VARIABLES NAMES.
COMPARE =        BREAKDOWN VARIABLE NAME FOR WICH CATEGORY CONTRASTS ARE REQUESTED.
CATEGORY =       LIST OF THE "COMPARE" VARIABLE CATEGORIES FOR WHICH A
CONTRAT IS REQUESTED.
STAT =           REQUESTED STATISTIQUE.
SUMWGT =         SUM OF THE WEIGHT
MEAN =           MEAN
VAR =            VARIANCE
STD =            STANDARD DEVIATION
CV =             COEFFICIENT OF VARIATION
SKEWNESS =       SKEWNESS
KURTOSIS =       KURTOSIS
MEDIAN =         MEDIAN
Q1 =             FIRST QUARTILE
Q3 =             THIRD QUARTILE
QRANGE =         RANGE BETWEEN Q1 AND Q3
PX =             PERCENTILE, WITH X BETWEEN 1 AND 99
OUTFILE =        FILE WITH THE STATISTIC ESTIMATES AND THEIR STANDARD ERROR ESTIMATE

*/

OPTIONS NONOTES;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP0 BRR_TEMP1 ;
RUN;

PROC SORT DATA=&INFILE OUT=BRRDATA(KEEP=&REPLI_ROOT.0-&REPLI_ROOT.80 &BYVAR &COM-
PARE &PV_ROOT.1-&PV_ROOT.5);
    BY &BYVAR &COMPARE;
RUN;

%DO I = 0 %TO 80;

    PROC MEANS DATA=BRRDATA VARDEF=WGT NOPRINT;
        VAR &PV_ROOT.1 - &PV_ROOT.5 ;
        BY &BYVAR &COMPARE;
        WEIGHT &REPLI_ROOT&I;
        OUTPUT OUT=MEAN_TEMP &STAT=PV1 - PV5;
    RUN;

    DATA MEAN_TEMP;
        SET MEAN_TEMP;
        L=&I;
    RUN;

    PROC APPEND BASE = BRR_TEMP0 DATA=MEAN_TEMP;
    RUN;

%END;

```

Box 17.10 [2/3] SAS® macro of PROC_DIF_PV.sas

```

%LET DEBUT=1;
%LET SUIVANT=2;
%LET COMPTE=1;
%LET I=1;

%DO %WHILE(%LENGTH(%SCAN(&CATEGORY,&I)));
    %LET I=%EVAL(&I+1);
%END;

%LET NBCAT=%EVAL(&I-1);
%LET NBDIF=%EVAL((&NBCAT*(&NBCAT-1))/2);
%LET ORDRE=1;

%DO J=1 %TO &NBDIF;
    %DO K=&DEBUT %TO &NBCAT-1;
        %LET CAT1=%SCAN(&CATEGORY,&DEBUT);
        %LET CAT2=%SCAN(&CATEGORY,&SUIVANT);

        DATA BRR_DIF1;
            SET BRR_TEMP0 (RENAME=(PV1=M1PV1 PV2=M1PV2 PV3=M1PV3
PV4=M1PV4 PV5=M1PV5));
            LENGTH CONTRAST $5;
            CONTRAST="&CAT1.-&CAT2";
            IF (&COMPARE=&CAT1);
            KEEP &BYVAR L CONTRAST M1PV1-M1PV5 &compare;

        RUN;

        DATA BRR_DIF2 ;
            SET BRR_TEMP0 (RENAME=(PV1=M2PV1 PV2=M2PV2 PV3=M2PV3
PV4=M2PV4 PV5=M2PV5));
            IF (&COMPARE=&CAT2);
            KEEP &BYVAR L M2PV1-M2PV5 &compare;

        RUN;

        PROC SORT DATA=BRR_DIF1;
            BY &BYVAR L;

        RUN;

        PROC SORT DATA=BRR_DIF2;
            BY &BYVAR L;

        RUN;

        DATA BRR_TEMP_DIF;
            MERGE BRR_DIF1 BRR_DIF2;
            BY &BYVAR L;
            IF (M1PV1 EQ . OR M2PV1 EQ .) THEN DELETE;
            PV1=M1PV1-M2PV1;
            PV2=M1PV2-M2PV2;
            PV3=M1PV3-M2PV3;
            PV4=M1PV4-M2PV4;
            PV5=M1PV5-M2PV5;
            ORDRE=&ORDRE;
            KEEP &BYVAR CONTRAST L PV1-PV5 ORDRE;

        RUN;

        PROC APPEND BASE = BRR_TEMP1 DATA=BRR_TEMP_DIF;
        RUN;

        %LET SUIVANT=%EVAL(&SUIVANT+1);
        %LET COMPTE=%EVAL(&COMPTE+1);
        %LET ORDRE=%EVAL(&ORDRE+1);

    %END;

%LET DEBUT=%EVAL(&DEBUT+1);
%LET SUIVANT=%EVAL(&DEBUT+1);

%END;

```



Box 17.10 [3/3] SAS® macro of PROC_DIF_PV.sas

```

DATA BRR_TEMP2 (DROP=STAT FIN1-FIN5 MESVAR) BRR_TEMP3 (KEEP=&BYVAR CONTRAST ORDRE
STAT FIN1-FIN5 MESVAR);
  SET BRR_TEMP1;
  IF L > 0 THEN OUTPUT BRR_TEMP2;
  ELSE DO;
    STAT = (PV1+PV2+PV3+PV4+PV5) / 5;
    FIN1=PV1;
    FIN2=PV2;
    FIN3=PV3;
    FIN4=PV4;
    FIN5=PV5;
    MESVAR= ( ((STAT-FIN1)**2) + ((STAT-FIN2)**2) + ((STAT-FIN3)**2) + ((STAT-
FIN4)**2) + ((STAT-FIN5)**2) ) / 4;
    OUTPUT BRR_TEMP3;
  END;
RUN;

PROC SORT DATA=BRR_TEMP2;
  BY &BYVAR CONTRAST;
RUN;

PROC SORT DATA=BRR_TEMP3;
  BY &BYVAR CONTRAST;
RUN;

DATA BRR_TEMP4;
  MERGE BRR_TEMP2 BRR_TEMP3;
  BY &BYVAR CONTRAST;
  ARRAY A (5)
    PV1-PV5;
  ARRAY B (5)
    FIN1-FIN5;
  ARRAY C (5)
    VAR1-VAR5;
  DO I=1 TO 5;
    C(I) = (1/20) * ((A(I) - B(I)) ** 2);
  END;
RUN;

PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
  VAR VAR1 VAR2 VAR3 VAR4 VAR5;
  BY &BYVAR CONTRAST;
  OUTPUT OUT=BRR_TEMP5 SUM=SS1 SS2 SS3 SS4 SS5;
RUN;

DATA BRR_TEMP6;
  MERGE BRR_TEMP3 BRR_TEMP5;
  BY &BYVAR CONTRAST;
  SAMP = (SS1+SS2+SS3+SS4+SS5) / 5;
  FINVAR = (SAMP + (1.2 * MESVAR));
  SESTAT = (FINVAR) ** 0.5;
  KEEP &BYVAR CONTRAST ORDRE STAT SESTAT;
RUN;

PROC SORT DATA=BRR_TEMP6 OUT=&OUTFILE (DROP=ORDRE);
  BY &BYVAR ORDRE;
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP0 BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5 BRR_TEMP6
  MEAN_TEMP BRRDATA BRR_DIF1 BRR_DIF2 BRR_TEMP_DIF;
RUN;

OPTIONS NOTES;

%MEND BRR_PROCMEAN_DIF_PV;

```



Box 17.11 [1/6] SAS® macro of QUARTILE_PV.sas

```

%MACRO QUARTILE_PV(INFILE=,
                   REPLI_ROOT=,
                   BYVAR = ,
                   PV_ROOT =,
                   INDEX =,
                   LIMIT=,
                   LIMIT_CRITERIA=,
                   ID_SCHOOL=,
                   OUTFILE =);

/*

MEANING OF THE MACRO ARGUMENTS

INFILE= INPUT DATA FILE.
REPLI_ROOT = ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL WEIGHT
VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
BYVAR = BREAKDOWN VARIABLES
PV_ROOT = ROOT OF THE 5 PLAUSIBLE VALUES VARIABLES NAMES
INDEX = VARIABLE NAME USED TO CREATE THE FOUR QUARTERS
LIMIT = FLAGGING YES OR NO
LIMIT_CRITERIA = 1) NUMBER OF STUDENTS 2) NUMBER OF SCHOOLS 3) PERCENTAGE OF STUDENTS
AND 4) NUMBER OF VARIABLES FROM THE BYVAR ARGUMENT FOR DEFINING THE POPULATION OF
REFERENCE.
ID_SCHOOL = VARIABLE NAME FOR THE SCHOOL IDENTIFICATION
OUTFILE = FILE WITH THE STATISTIC ESTIMATES AND THEIR STANDARD ERROR ESTIMATES.

*/

OPTIONS NONOTES;

DATA QUARTILE_TEMP1 (KEEP=&BYVAR &REPLI_ROOT.0-&REPLI_ROOT.80 &PV_ROOT.1-&PV_ROOT.5
INDEX1-INDE5 &ID_SCHOOL &INDEX NB_MISS) ;
  SET &INFILE;
  NB_MISS=0;
  ARRAY A (6) &INDEX &PV_ROOT.1-&PV_ROOT.5;
  DO I=1 TO 6;
    IF (A(I) IN (.,.M,.N,.I)) THEN NB_MISS=NB_MISS+1;
  END;
  INDEX1=&INDEX + (0.01*normal(-01));
  INDEX2=&INDEX + (0.01*normal(-23));
  INDEX3=&INDEX + (0.01*normal(-45));
  INDEX4=&INDEX + (0.01*normal(-67));
  INDEX5=&INDEX + (0.01*normal(-89));
RUN;

PROC SORT DATA=QUARTILE_TEMP1;
  BY &BYVAR;
RUN;

PROC MEANS DATA=QUARTILE_TEMP1 NOPRINT;
  VAR INDEX1-INDE5;
  BY &BYVAR;
  WEIGHT &REPLI_ROOT.0;
  WHERE (NB_MISS=0);
  OUTPUT OUT=QUARTILE_TEMP2 P25=Q1_1-Q1_5 P50=Q2_1-Q2_5 P75=Q3_1-Q3_5;
RUN;

```



Box 17.11 [2/6] SAS® macro of QUARTILE_PV.sas

```

DATA
  QUARTILE_TEMP3 (KEEP = &BYVAR &REPLI_ROOT.0-&REPLI_ROOT.80 &PV_ROOT.1-&PV_ROOT.5
    INDEX1-INDEX5 CAT1-CAT5 &ID SCHOOL &INDEX NB MISS)
  QUARTILE_TEMP4 (KEEP = &BYVAR CAT1-CAT5 &REPLI_ROOT.0 NB_MISS &ID SCHOOL
&INDEX);
  MERGE QUARTILE_TEMP1 QUARTILE_TEMP2;
  BY &BYVAR;
  ARRAY A1 (5) INDEX1-INDEX5;
  ARRAY A2 (5) Q1_1-Q1_5;
  ARRAY A3 (5) Q2_1-Q2_5;
  ARRAY A4 (5) Q3_1-Q3_5;
  ARRAY A5 (5) CAT1-CAT5;
  DO I=1 TO 5;
    IF (A1(i) <= A2(i)) THEN A5(i)=1;
    IF (A1(i) > A2(i) AND A1(i) <= A3(i)) THEN A5(i)=2;
    IF (A1(i) > A3(i) AND A1(i) <= A4(i)) THEN A5(i)=3;
    IF (A1(i) > A4(i)) THEN A5(i)=4;
    IF (A1(i) IN (.,.M,.N,.I)) THEN A5(i) = .;
  END;
  IF (NB_MISS=0) THEN OUTPUT QUARTILE_TEMP3;
  OUTPUT QUARTILE_TEMP4;
RUN;

%DO QUARTILE=1 %TO 5;

%SUBQUARTILE1(INFILE =QUARTILE_TEMP3,
  REPLI_ROOT =&REPLI_ROOT,
  SUB_BY_VAR = &BYVAR CAT&QUARTILE ,
  VAR =&INDEX,
  OUTFILE =INDEX&QUARTILE);

RUN;

DATA INDEX&QUARTILE;
  SET INDEX&QUARTILE;
  STAT&QUARTILE=STAT;
  SESTAT&QUARTILE=SESTAT;
  CAT=CAT&QUARTILE;
  KEEP &BYVAR CAT STAT&QUARTILE SESTAT&QUARTILE ;

RUN;

%END;

DATA QUARTILE_TEMP5;
  MERGE INDEX1 INDEX2 INDEX3 INDEX4 INDEX5;
  BY &BYVAR CAT;
  STAT=(STAT1+STAT2+STAT3+STAT4+STAT5)/5;
  VAR_IMP=(((STAT1-STAT)**2)+((STAT2-STAT)**2)+((STAT3-STAT)**2)+((STAT4-
STAT)**2)+((STAT5-STAT)**2))/4;
  VAR_SAMP=((SESTAT1**2)+(SESTAT2**2)+(SESTAT3**2)+(SESTAT4**2)+(SESTAT5**2))/5;
  SESTAT=(VAR_SAMP+(1.2*VAR_IMP)**0.5);
  KEEP &BYVAR CAT STAT SESTAT;

RUN;

%DO QUARTILE=1 %TO 5;

%SUBQUARTILE1(INFILE =QUARTILE_TEMP3,
  REPLI_ROOT =&REPLI_ROOT,
  SUB_BY_VAR = &BYVAR CAT&QUARTILE ,
  VAR =&PV_ROOT&QUARTILE,
  OUTFILE =RESULT&QUARTILE);

RUN;

DATA RESULT&QUARTILE;
  SET RESULT&QUARTILE;
  STAT&QUARTILE=STAT;
  SESTAT&QUARTILE=SESTAT;
  CAT=CAT&QUARTILE;
  KEEP &BYVAR CAT STAT&QUARTILE SESTAT&QUARTILE ;

RUN;

%END;

```

Box 17.11 [3/6] SAS® macro of QUARTILE_PV.sas

```

DATA QUARTILE_TEMP6;
  MERGE RESULT1 RESULT2 RESULT3 RESULT4 RESULT5;
  BY &BYVAR CAT;
  STAT=(STAT1+STAT2+STAT3+STAT4+STAT5)/5;
  VAR_IMP=(( (STAT1-STAT)**2)+((STAT2-STAT)**2)+((STAT3-STAT)**2)+((STAT4-
STAT)**2)+((STAT5-STAT)**2))/4;
  VAR_SAMP=((SESTAT1**2)+(SESTAT2**2)+(SESTAT3**2)+(SESTAT4**2)+(SESTAT5**2))/5;
  SESTAT=(VAR_SAMP+(1.2*VAR_IMP))*0.5;
  KEEP &BYVAR CAT STAT SESTAT;
RUN;

DATA QUARTILE_TEMP7;
  MERGE QUARTILE_TEMP5 (RENAME=(STAT=INDEX_STAT SESTAT=INDEX_SESTAT))
        QUARTILE_TEMP6 (RENAME=(STAT=PV_STAT SESTAT=PV_SESTAT)) ;
  BY &BYVAR CAT;
RUN;

%IF (%UPCASE(&LIMIT)=NO) %THEN %DO;

  DATA &OUTFILE;
    SET QUARTILE_TEMP7;;
  RUN;

%END;
%ELSE %DO;
  PROC FREQ DATA=QUARTILE_TEMP3 NOPRINT;
    TABLE &BYVAR * CAT1 /OUT=QUARTILE_TEMP8_1;
    TABLE &BYVAR * CAT2 /OUT=QUARTILE_TEMP8_2;
    TABLE &BYVAR * CAT3 /OUT=QUARTILE_TEMP8_3;
    TABLE &BYVAR * CAT4 /OUT=QUARTILE_TEMP8_4;
    TABLE &BYVAR * CAT5 /OUT=QUARTILE_TEMP8_5;
  RUN;
  %LET FLAG_STUD=%SCAN(&LIMIT_CRITERIA,1);
  DATA QUARTILE_TEMP8;
    MERGE
      QUARTILE_TEMP8_1 (RENAME=(COUNT=C1 CAT1=CAT))
      QUARTILE_TEMP8_2 (RENAME=(COUNT=C2 CAT2=CAT))
      QUARTILE_TEMP8_3 (RENAME=(COUNT=C3 CAT3=CAT))
      QUARTILE_TEMP8_4 (RENAME=(COUNT=C4 CAT4=CAT))
      QUARTILE_TEMP8_5 (RENAME=(COUNT=C5 CAT5=CAT));
    BY &BYVAR ;
    C=(C1+C2+C3+C4+C5)/5;
    IF (C < &FLAG_STUD) THEN FLAG_STUD=1;
    ELSE FLAG_STUD=0;
    KEEP &BYVAR FLAG_STUD;
  RUN;

  %DO M=1 %TO 5;
  PROC FREQ DATA=QUARTILE_TEMP3 NOPRINT;
    TABLE &BYVAR * CAT&M * &ID_SCHOOL /OUT=QUARTILE_TEMP9_&M;
  RUN;
  PROC FREQ DATA=QUARTILE_TEMP9_&M NOPRINT;
    TABLE &BYVAR * CAT&M /OUT=QUARTILE_TEMP10_&M;
  RUN;
  %END;
  %LET FLAG_SCH=%SCAN(&LIMIT_CRITERIA,2);
  DATA QUARTILE_TEMP10;
    MERGE
      QUARTILE_TEMP10_1 (RENAME=(COUNT=C1 CAT1=CAT))
      QUARTILE_TEMP10_2 (RENAME=(COUNT=C2 CAT2=CAT))
      QUARTILE_TEMP10_3 (RENAME=(COUNT=C3 CAT3=CAT))
      QUARTILE_TEMP10_4 (RENAME=(COUNT=C4 CAT4=CAT))
      QUARTILE_TEMP10_5 (RENAME=(COUNT=C5 CAT5=CAT));
    BY &BYVAR CAT;
    C=(C1+C2+C3+C4+C5)/5;
    IF (C < &FLAG_SCH) THEN FLAG_SCH=1;
    ELSE FLAG_SCH=0;
    KEEP &BYVAR FLAG_SCH;
  RUN;

```



Box 17.11 [4/6] SAS® macro of QUARTILE_PV.sas

```

%LET K=1;
%LET POPREF=;
%LET NBVAR=%SCAN(&LIMIT_CRITERIA,4);
%DO %WHILE(&K <= &NBVAR);
%LET POPREF_ADD=%SCAN(&BYVAR,&K);
%LET POPREF=&POPREF &POPREF_ADD;
%LET K=%EVAL(&K+1);
%END;
%LET FLAG_PCT=%SCAN(&LIMIT_CRITERIA,3);
PROC FREQ DATA=QUARTILE_TEMP4 NOPRINT;
TABLE &BYVAR * CAT1 * NB_MISS / OUT=QUARTILE_TEMP11_1;
TABLE &BYVAR * CAT2 * NB_MISS / OUT=QUARTILE_TEMP11_2;
TABLE &BYVAR * CAT3 * NB_MISS / OUT=QUARTILE_TEMP11_3;
TABLE &BYVAR * CAT4 * NB_MISS / OUT=QUARTILE_TEMP11_4;
TABLE &BYVAR * CAT5 * NB_MISS / OUT=QUARTILE_TEMP11_5;
WEIGHT W_FSTR0;
RUN;
DATA QUARTILE_TEMP11;
MERGE
    QUARTILE_TEMP11_1 (RENAME=(COUNT=C1 CAT1=CAT))
    QUARTILE_TEMP11_2 (RENAME=(COUNT=C2 CAT2=CAT))
    QUARTILE_TEMP11_3 (RENAME=(COUNT=C3 CAT3=CAT))
    QUARTILE_TEMP11_4 (RENAME=(COUNT=C4 CAT4=CAT))
    QUARTILE_TEMP11_5 (RENAME=(COUNT=C5 CAT5=CAT));
BY &BYVAR CAT NB_MISS;
RUN;
PROC MEANS DATA=QUARTILE_TEMP11 NOPRINT;
VAR C1-C5;
BY &BYVAR;
OUTPUT OUT=QUARTILE_TEMP12 SUM=SOMWGT1-SOMWGT5;
RUN;

DATA QUARTILE_TEMP13;
MERGE QUARTILE_TEMP11 QUARTILE_TEMP12;
BY &BYVAR;
IF (NB_MISS=0);
ARRAY A1 (5) C1-C5;
ARRAY A2 (5) SOMWGT1-SOMWGT5;
ARRAY A3 (5) PCT1-PCT5;
DO I=1 TO 5;
    A3(I)=(A1(I)/A2(I))*100;
END;
PCT=(PCT1+PCT2+PCT3+PCT4+PCT5)/5;
IF (PCT < &FLAG_PCT) THEN FLAG_PCT=1;
ELSE FLAG_PCT=0;
KEEP &BYVAR FLAG_PCT;
RUN;
DATA &OUTFILE;
MERGE QUARTILE_TEMP7 QUARTILE_TEMP8 QUARTILE_TEMP10 QUARTILE_TEMP13;
BY &BYVAR;
RUN;
PROC DATASETS LIBRARY=WORK NOLIST;
DELETE
    QUARTILE_TEMP8_1 QUARTILE_TEMP8_2 QUARTILE_TEMP8_3 QUARTILE_TEMP8_4
QUARTILE_TEMP8_5
    QUARTILE_TEMP8
    QUARTILE_TEMP9_1 QUARTILE_TEMP9_2 QUARTILE_TEMP9_3 QUARTILE_TEMP9_4
QUARTILE_TEMP9_5
    QUARTILE_TEMP10_1 QUARTILE_TEMP10_2 QUARTILE_TEMP10_3 QUARTILE_TEMP10_4
QUARTILE_TEMP10_5
    QUARTILE_TEMP10
    QUARTILE_TEMP11_1 QUARTILE_TEMP11_2 QUARTILE_TEMP11_3 QUARTILE_TEMP11_4
QUARTILE_TEMP11_5
    QUARTILE_TEMP11 QUARTILE_TEMP12 QUARTILE_TEMP13;
RUN;
%END;

```



Box 17.11 [5/6] SAS® macro of QUARTILE_PV.sas

```

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE QUARTILE_TEMP1 QUARTILE_TEMP2 QUARTILE_TEMP3 QUARTILE_TEMP4 QUARTILE_
TEMP5 QUARTILE_TEMP6 QUARTILE_TEMP7
        INDEX1 INDEX2 INDEX3 INDEX4 INDEX5
        RESULT1 RESULT2 RESULT3 RESULT4 RESULT5;
RUN;

OPTION NOTES;

%MEND ;

%MACRO SUBQUARTILE1(INFILE =,
                   REPLI_ROOT =,
                   SUB_BY_VAR =,
                   VAR =,
                   OUTFILE =);

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP1;
RUN;

PROC SORT DATA=&INFILE OUT=BRRDATA(KEEP=&REPLI_ROOT.0-&REPLI_ROOT.80 &SUB_BY_VAR
&VAR);
  BY &SUB_BY_VAR;
RUN;

%DO I = 0 %TO 80;
  PROC MEANS DATA=BRRDATA VARDEF=WGT NOPRINT;
    VAR &VAR ;
    BY &SUB_BY_VAR;
    WEIGHT &REPLI_ROOT&I;
    OUTPUT OUT=MEAN_TEMP MEAN=PV;
  RUN;

  DATA MEAN_TEMP;
    SET MEAN_TEMP;
    L=&I;
  RUN;

  PROC APPEND BASE = BRR_TEMP1 DATA=MEAN_TEMP;
  RUN;
%END;

DATA BRR_TEMP2(KEEP=&SUB_BY_VAR PV) BRR_TEMP3(KEEP=&SUB_BY_VAR STAT);
  SET BRR_TEMP1;
  IF L > 0 THEN OUTPUT BRR_TEMP2;
  ELSE DO;
    STAT =PV;
    OUTPUT BRR_TEMP3;
  END;
RUN;

PROC SORT DATA=BRR_TEMP2;
  BY &SUB_BY_VAR;
RUN;

PROC SORT DATA=BRR_TEMP3;
  BY &SUB_BY_VAR;
RUN;

DATA BRR_TEMP4;
  MERGE BRR_TEMP2 BRR_TEMP3;
  BY &SUB_BY_VAR;
  VARI=((PV-STAT)**2)*(1/20);
RUN;

```




Box 17.11 [6/6] SAS® macro of QUARTILE_PV.sas

```
PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
  VAR VARI;
  BY &SUB_BY_VAR;
  OUTPUT OUT=BRR_TEMP5 SUM=SS;
RUN;

DATA &OUTFILE;
  MERGE BRR_TEMP3 BRR_TEMP5;
  BY &SUB_BY_VAR;
  SESTAT=(SS)**0.5;
  FORMAT STAT F10.2;
  FORMAT SESTAT F10.2;
  KEEP &SUB_BY_VAR STAT SESTAT;
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5 BRR_TEMP6 MEAN_TEMP
BRRDATA;
RUN;

%MEND SUBQUARTILE1;
```



Box 17.12 [1/3] SAS® macro of RELATIVE_RISK_NO_PV.sas

```

%macro BRR_RR(INFILE=,
              REPLI_ROOT=,
              BYVAR=,
              ANTECEDENT=,
              OUTCOME=,
              LIMIT=,
              LIMIT_CRITERIA=,
              ID_SCHOOL=,
              OUTFILE=);

/*
MEANING OF THE MACRO ARGUMENTS

INFILE =          INPUT DATA FILE.
REPLI_ROOT =      ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL
WEIGHT VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
BYVAR = BREAKDOWN VARIABLES
ANTECEDENT = ANTECEDENT VARIABLES (0,1)
OUTCOME =        CONSEQUENCE VARIABLES (0,1)
LIMIT = FLAGGING YES OR NO
LIMIT_CRITERIA = 1) NUMBER OF STUDENTS 2) NUMBER OF SCHOOLS 3) PERCENTAGE OF
STUDENTS AND 4) NUMBER OF VARIABLES FROM THE BYVAR ARGUMENT FOR DEFINING THE
POPULATION OF REFERENCE.
ID_SCHOOL        = VARIABLE NAME FOR THE SCHOOL IDENTIFICATION
OUTFILE = FILE WITH THE STATISTIC ESTIMATES AND THEIR STANDARD ERROR ESTIMATES.

*/

OPTIONS NONOTES;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP;
RUN;

PROC SORT DATA=&INFILE
  OUT=BRRDATA(KEEP=&REPLI_ROOT.0-&REPLI_ROOT.80 &BYVAR &ANTECEDENT &OUTCOME
&ID_SCHOOL);
  BY &BYVAR &ANTECEDENT;
RUN;

%DO I = 0 %TO 80;
  PROC FREQ DATA=BRRDATA NOPRINT;
    TABLE &OUTCOME /OUT=FREQ_TEMP ;
    BY &BYVAR &ANTECEDENT ;
    WEIGHT &REPLI_ROOT&I;
    WHERE (&ANTECEDENT IN (0,1) AND &OUTCOME IN (0,1));

  RUN;

  DATA FREQ_TEMP;
    SET FREQ_TEMP;
    L=&I;
    IF (&OUTCOME=1);
    KEEP &BYVAR L &ANTECEDENT PERCENT;

  RUN;

  PROC APPEND BASE = BRR_TEMP DATA=FREQ_TEMP;
  RUN;

%END;

PROC SORT DATA=BRR_TEMP OUT=BRR_TEMP0;
  BY &BYVAR L &ANTECEDENT ;
RUN;

PROC TRANSPOSE DATA=BRR_TEMP0 OUT=BRR_TEMP1;
  VAR PERCENT;
  BY &BYVAR L;
  ID &ANTECEDENT ;
RUN;

```



Box 17.12 [2/3] SAS® macro of RELATIVE_RISK_NO_PV.sas

```

DATA BRR_TEMP1;
    SET BRR_TEMP1;
    PV=(_1/_0);
RUN;

DATA BRR_TEMP2(KEEP=&BYVAR PV)BRR_TEMP3(KEEP=&BYVAR STAT);
    SET BRR_TEMP1;
    IF L > 0 THEN OUTPUT BRR_TEMP2;
    ELSE DO;
        STAT =PV;
        OUTPUT BRR_TEMP3;
    END;
RUN;

PROC SORT DATA=BRR_TEMP2;
    BY &BYVAR;
RUN;

PROC SORT DATA=BRR_TEMP3;
    BY &BYVAR;
RUN;

DATA BRR_TEMP4;
    MERGE BRR_TEMP2 BRR_TEMP3;
    BY &BYVAR;
    VARI=((PV-STAT)**2)*(1/20);
RUN;

PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
    VAR VARI;
    BY &BYVAR;
    OUTPUT OUT=BRR_TEMP5 SUM=SS;
RUN;

DATA BRR_TEMP6;
    MERGE BRR_TEMP3 BRR_TEMP5;
    BY &BYVAR;
    SESTAT=(SS)**0.5;
    FORMAT STAT F10.2;
    FORMAT SESTAT F10.2;
    KEEP &BYVAR STAT SESTAT;
RUN;

%IF (%UPCASE(&LIMIT)=NO) %THEN %DO;

    DATA &OUTFILE;
        SET BRR_TEMP6;
    RUN;

%END;
%ELSE %DO;
    DATA BRRDATA;
        SET BRRDATA;
        MIS=0;
        IF (&ANTECEDENT NOT IN (0,1) OR &OUTCOME NOT IN (0,1)) THEN MIS=1;
    RUN;
    PROC FREQ DATA=BRRDATA NOPRINT;
        TABLE MIS /OUT=BRR_TEMP7;
        BY &BYVAR;
        WHERE (MIS=0);
    RUN;
    %LET FLAG_STUD=%SCAN(&LIMIT_CRITERIA,1);
    DATA BRR_TEMP7;
        SET BRR_TEMP7;
        IF (COUNT < &FLAG_STUD) THEN FLAG_STUD=1;
        ELSE FLAG_STUD=0;
        KEEP &BYVAR FLAG_STUD;
    RUN;

```

Box 17.12 [3/3] SAS® macro of RELATIVE_RISK_NO_PV.sas

```

PROC SORT DATA=BRRDATA;
    BY &BYVAR &ID_SCHOOL;
RUN;
PROC FREQ DATA=BRRDATA NOPRINT;
    TABLE MIS /OUT=BRR_TEMP8;
    BY &BYVAR &ID_SCHOOL;
    WHERE (MIS=0);
RUN;
PROC FREQ DATA=BRR_TEMP8 NOPRINT;
    TABLE MIS /OUT=BRR_TEMP9;
    BY &BYVAR;
RUN;
%LET FLAG_SCH=%SCAN(&LIMIT_CRITERIA,2);
DATA BRR_TEMP9;
    SET BRR_TEMP9;
    IF (COUNT < &FLAG_SCH) THEN FLAG_SCH=1;
    ELSE FLAG_SCH=0;
    KEEP &BYVAR FLAG_SCH ;
RUN;

PROC FREQ DATA=BRRDATA NOPRINT;
    TABLE MIS/OUT=BRR_TEMP10;
    BY &BYVAR;
    WEIGHT &REPLI_ROOT.0;
RUN;
%LET K=1;
%LET POPREF=;
%LET NBVAR=%SCAN(&LIMIT_CRITERIA,4);
%DO %WHILE(&K <= &NBVAR);
    %LET POPREF_ADD=%SCAN(&BYVAR,&K);
    %LET POPREF=&POPREF &POPREF_ADD;
    %LET K=%EVAL(&K+1);
%END;

PROC MEANS DATA=BRR_TEMP10 NOPRINT;
    VAR COUNT;
    BY &POPREF;
    OUTPUT OUT=BRR_TEMP11 SUM=SOMWGT;
RUN;
%LET FLAG_PCT=%SCAN(&LIMIT_CRITERIA,3);
DATA BRR_TEMP12;
    MERGE BRR_TEMP10 BRR_TEMP11;
    BY &POPREF;
    PCT=(COUNT/SOMWGT)*100;
    IF (PCT < &FLAG_PCT) THEN FLAG_PCT=1;
    ELSE FLAG_PCT=0;
    IF (MIS=0);
    KEEP &BYVAR FLAG_PCT;
RUN;
DATA &OUTFILE;
    MERGE BRR_TEMP6 BRR_TEMP7 BRR_TEMP9 BRR_TEMP12;
    BY &BYVAR;
    IF (STAT NE .);
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP7 BRR_TEMP8 BRR_TEMP9 BRR_TEMP10 BRR_TEMP11 BRR_
TEMP12;
RUN;
%END;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP BRR_TEMP0 BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5
BRR_TEMP6 FREQ_TEMP BRRDATA;
RUN;

OPTIONS NOTES;

%MEND BRR_RR;

```



Box 17.13 [1/5] SAS® macro of RELATIVE_RISK_PV.sas

```

%macro BRR_RR_PV(INFILE=,
                 REPLI_ROOT=,
                 BYVAR=,
                 ANTECEDENT_ROOT=,
                 OUTCOME_ROOT=,
                 LIMIT=,
                 LIMIT_CRITERIA=,
                 ID_SCHOOL=,
                 OUTFILE=);

/*

MEANING OF THE MACRO ARGUMENTS

INFILE = INPUT DATA FILE.
REPLI_ROOT = ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES.
FINAL WEIGHT VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
BYVAR = BREAKDOWN VARIABLES
ANTECEDENT_ROOT = ROOT OF THE 5 ANTECEDENT VARIABLES (0,1)
OUTCOME_ROOT = ROOT OF THE 5 CONSEQUENCE VARIABLES (0,1)
LIMIT = FLAGGING YES OR NO
LIMIT_CRITERIA = 1) NUMBER OF STUDENTS 2) NUMBER OF SCHOOLS 3) PERCENTAGE OF
STUDENTS AND 4) NUMBER OF VARIABLES FROM THE BYVAR ARGUMENT FOR DEFINING THE
POPULATION OF REFERENCE.
ID_SCHOOL = VARIABLE NAME FOR THE SCHOOL IDENTIFICATION
OUTFILE = FILE WITH THE STATISTIC ESTIMATES AND THEIR STANDARD ERROR ESTIMATES.

*/

OPTIONS NONOTES;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP;
RUN;

%DO I = 1 %TO 5;

PROC SORT DATA=&INFILE
          OUT=BRRDATA(KEEP=          &REPLI_ROOT.0-&REPLI_ROOT.80 &BYVAR &ANTECEDENT_
          ROOT.1-&ANTECEDENT_ROOT.5
          &OUTCOME_ROOT.1-&OUTCOME_ROOT.5 &ID_
          SCHOOL);
    BY &BYVAR &ANTECEDENT_ROOT&I;
RUN;

%DO J = 0 %TO 80;
    PROC FREQ DATA=BRRDATA NOPRINT;
        TABLE &OUTCOME_ROOT&I /OUT=FREQ_TEMP ;
        BY &BYVAR &ANTECEDENT_ROOT&I ;
        WEIGHT &REPLI_ROOT&J;
        WHERE (&ANTECEDENT_ROOT&I IN (0,1) AND &OUTCOME_ROOT&I IN (0,1));
    RUN;

    DATA FREQ_TEMP;
        SET FREQ_TEMP;
        K=&I;
        L=&J;
        &ANTECEDENT_ROOT=&ANTECEDENT_ROOT&I;
        IF (&OUTCOME_ROOT&I=1);
        KEEP &BYVAR K L &ANTECEDENT_ROOT PERCENT;
    RUN;

    PROC APPEND BASE = BRR_TEMP DATA=FREQ_TEMP;
    RUN;

%END;

```



Box 17.13 [2/5] **SAS® macro of RELATIVE_RISK_PV.sas**

```

PROC SORT DATA=BRR_TEMP OUT=BRR_PREP1;
    BY &BYVAR L K &ANTECEDENT_ROOT ;
RUN;
PROC TRANSPOSE DATA=BRR_PREP1 OUT=BRR_PREP2;
    VAR PERCENT;
    BY &BYVAR L K;
    ID &ANTECEDENT_ROOT ;
RUN;
DATA BRR_PREP2;
    SET BRR_PREP2;
    PV=( _1/_0 );
    KEEP &BYVAR L K PV;
RUN;
PROC TRANSPOSE DATA=BRR_PREP2 OUT=BRR_PREP3 PREFIX=PV;
    VAR PV;
    BY &BYVAR L;
    ID K ;
RUN;
DATA BRR_TEMP1;
    SET BRR_PREP3;
    DROP _NAME_ ;
RUN;

DATA BRR_TEMP2 (DROP=STAT FIN1-FIN5 MESVAR) BRR_TEMP3 (KEEP=&BYVAR STAT FIN1-FIN5
MESVAR) ;
    SET BRR_TEMP1;
    IF L > 0 THEN OUTPUT BRR_TEMP2;
    ELSE DO;
        STAT = (PV1+PV2+PV3+PV4+PV5) /5;
        FIN1=PV1;
        FIN2=PV2;
        FIN3=PV3;
        FIN4=PV4;
        FIN5=PV5;
        MESVAR= ( ( (STAT-FIN1)**2) + ( (STAT-FIN2)**2) + ( (STAT-FIN3)**2) + ( (STAT-
FIN4)**2) + ( (STAT-FIN5)**2) ) /4;
        OUTPUT BRR_TEMP3;
    END;
RUN;

PROC SORT DATA=BRR_TEMP2;
    BY &BYVAR;
RUN;

PROC SORT DATA=BRR_TEMP3;
    BY &BYVAR ;
RUN;

DATA BRR_TEMP4;
    MERGE BRR_TEMP2 BRR_TEMP3;
    BY &BYVAR ;
    ARRAY A (5)
        PV1-PV5;
    ARRAY B (5)
        FIN1-FIN5;
    ARRAY C(5)
        VAR1-VAR5;
    DO I=1 TO 5;
        C(I) = (1/20) * ( (A(I) -B(I))**2) ;
    END;
RUN;

PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
    VAR VAR1 VAR2 VAR3 VAR4 VAR5;
    BY &BYVAR ;
    OUTPUT OUT=BRR_TEMP5 SUM=SS1 SS2 SS3 SS4 SS5;
RUN;

```



Box 17.13 [3/5] SAS® macro of RELATIVE_RISK_PV.sas

```

DATA BRR_TEMP6;
  MERGE BRR_TEMP3 BRR_TEMP5;
  BY &BYVAR ;
  SAMP=(SS1+SS2+SS3+SS4+SS5)/5;
  FINVAR=(SAMP+(1.2*MESVAR));
  SESTAT=(FINVAR)**0.5;
  FORMAT STAT F10.2;
  FORMAT SESTAT F10.2;
  KEEP &BYVAR STAT SESTAT;
RUN;

%IF (%UPCASE(&LIMIT)=NO) %THEN %DO;

  DATA &OUTFILE;
    SET BRR_TEMP6;
  RUN;

%END;
%ELSE %DO;
  DATA BRRDATA;
    SET BRRDATA;
    MIS=1;
    IF (&ANTECEDENT_ROOT.1 IN (0,1) AND &OUTCOME_ROOT.1 IN (0,1)) THEN
MIS=0;
  RUN;
  %DO M=1 %TO 5;
  PROC FREQ DATA=BRRDATA NOPRINT;
    TABLE MIS/OUT=BRR_TEMP7;
    BY &BYVAR;
    WHERE (MIS=0);
  RUN;
  DATA BRR_TEMP8;
    SET BRR_TEMP7;
    COUNT&M=COUNT;
    KEEP &BYVAR COUNT&M;
  RUN;
  %IF (&M=1) %THEN %DO;
    DATA BRR_TEMP9;
      SET BRR_TEMP8;
    RUN;
  %END;
  %ELSE %DO;
    DATA BRR_TEMP9;
      MERGE BRR_TEMP9 BRR_TEMP8;
      BY &BYVAR ;
    RUN;
  %END;
  %END;
  %LET FLAG_STUD=%SCAN(&LIMIT_CRITERIA,1);
  DATA BRR_TEMP9;
    SET BRR_TEMP9;
    COUNT=(COUNT1+COUNT2+COUNT3+COUNT4+COUNT5)/5;
    IF (COUNT < &FLAG_STUD) THEN FLAG_STUD=1;
    ELSE FLAG_STUD=0;
    KEEP &BYVAR FLAG_STUD;
  RUN;

  PROC SORT DATA=BRRDATA;
    BY &BYVAR &ID_SCHOOL;
  RUN;

  %DO M=1 %TO 5;
  PROC FREQ DATA=BRRDATA NOPRINT;
    TABLE MIS /OUT=BRR_TEMP10;
    BY &BYVAR &ID_SCHOOL;
    WHERE (MIS=0);
  RUN;
  PROC FREQ DATA=BRR_TEMP10 NOPRINT;
    TABLE MIS /OUT=BRR_TEMP11;
    BY &BYVAR;
  RUN;

```



Box 17.13 [4/5] SAS® macro of RELATIVE_RISK_PV.sas

```

DATA BRR_TEMP12;
    SET BRR_TEMP11;
    COUNT&M=COUNT;
    KEEP &BYVAR COUNT&M;
RUN;
%IF (&M=1) %THEN %DO;
    DATA BRR_TEMP13;
        SET BRR_TEMP12;
    RUN;
%END;
%ELSE %DO;
    DATA BRR_TEMP13;
        MERGE BRR_TEMP13 BRR_TEMP12;
        BY &BYVAR;
    RUN;
%END;
%END;
%LET FLAG_SCH=%SCAN(&LIMIT_CRITERIA,2);
DATA BRR_TEMP13;
    SET BRR_TEMP13;
    COUNT=(COUNT1+COUNT2+COUNT3+COUNT4+COUNT5)/5;
    IF (COUNT < &FLAG_SCH) THEN FLAG_SCH=1;
    ELSE FLAG_SCH=0;
    KEEP &BYVAR FLAG_SCH;
RUN;

%DO M=1 %TO 5;
PROC FREQ DATA=BRRDATA NOPRINT;
    TABLE MIS/OUT=BRR_TEMP14;
    BY &BYVAR;
    WEIGHT &REPLI_ROOT.0;
RUN;
DATA BRR_TEMP15;
    SET BRR_TEMP14;
    COUNT&M=COUNT;
    KEEP &BYVAR MIS COUNT&M;
RUN;
%IF (&M=1) %THEN %DO;
    DATA BRR_TEMP16;
        SET BRR_TEMP15;
    RUN;
%END;
%ELSE %DO;
    DATA BRR_TEMP16;
        MERGE BRR_TEMP16 BRR_TEMP15;
        BY &BYVAR MIS;
    RUN;
%END;
%END;

%LET K=1;
%LET POPREF=;
%LET NBVAR=%SCAN(&LIMIT_CRITERIA,4);
%DO %WHILE(&K <= &NBVAR);
    %LET POPREF_ADD=%SCAN(&BYVAR,&K);
    %LET POPREF=&POPREF &POPREF_ADD;
    %LET K=%EVAL(&K+1);
%END;

PROC MEANS DATA=BRR_TEMP16 NOPRINT;
    VAR COUNT1-COUNT5 ;
    BY &POPREF;
    OUTPUT OUT=BRR_TEMP17 SUM=SOMWGT1-SOMWGT5;
RUN;
  
```




Box 17.13 [5/5] SAS® macro of RELATIVE_RISK_PV.sas

```

%LET FLAG_PCT=%SCAN(&LIMIT_CRITERIA,3);
DATA BRR_TEMP18;
    MERGE BRR_TEMP16 BRR_TEMP17;
    BY &POPREF;
    IF (MIS=0);
    ARRAY B1 (5) PCT1-PCT5;
    ARRAY B2 (5) COUNT1-COUNT5;
    ARRAY B3 (5) SOMWGT1-SOMWGT5;
    DO A=1 TO 5;
        B1(A) = (B2(A)/B3(A)) *100;
    END;
    PCT=(PCT1+PCT2+PCT3+PCT4+PCT5)/5;
    IF (PCT < &FLAG_PCT) THEN FLAG_PCT=1;
    ELSE FLAG_PCT=0;
    KEEP &BYVAR FLAG_PCT;
RUN;

DATA &OUTFILE;
    MERGE BRR_TEMP6 BRR_TEMP9 BRR_TEMP13 BRR_TEMP18;
    BY &BYVAR ;
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP7 BRR_TEMP8 BRR_TEMP9 BRR_TEMP10 BRR_TEMP11 BRR_
TEMP12 BRR_TEMP13
                                BRR_TEMP14 BRR_TEMP15 BRR_TEMP16 BRR_TEMP17 BRR_
TEMP18;
RUN;

%END;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_PREP1 BRR_PREP2 BRR_PREP3 BRR_TEMP BRR_TEMP1 BRR_TEMP2 BRR_TEMP3
BRR_TEMP4 BRR_TEMP5 FREQ_TEMP BRR_TEMP6 MEAN_TEMP BRRDATA;
RUN;

OPTIONS NOTES;

%MEND BRR_RR_PV;

```



Box 17.14 [1/2] SAS® macro of EFFECT_SIZE_NO_PV.sas

```

%MACRO BRR_EFFECT (INFILE=,
                   REPLI_ROOT=,
                   BYVAR=,
                   VAR=,
                   EFFECT=,
                   OUTFILE=);

/*
MEANING OF THE MACRO ARGUMENTS

INFILE = INPUT DATA FILE.
REPLI_ROOT = ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL
WEIGHT VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
BYVAR = BREAKDOWN VARIABLES.
VAR = DEPENDENT VARIABLES.
EFFECT = THIS ARGUMENT HAS THREE COMPONENTS: THE VARIABLE NAME AND THE TWO
CATEGORIES ON WHICH THE EFFECT SIZE WILL BE COMPUTED. THE EFFECT SIZE WILL BE THE
SECOND COMPONENT MINUS THE THIRD COMPONENT.
OUTFILE = FILE WITH THE STATISTIC ESTIMATES AND THEIR STANDARD ERROR ESTIMATES.

*/

%LET GROUP=%SCAN(&EFFECT,1);
%LET CAT1=%SCAN(&EFFECT,2);
%LET CAT2=%SCAN(&EFFECT,3);

OPTIONS NONOTES;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP1;
RUN;

DATA BRRDATA;
  SET &INFILE;
  IF (&GROUP=&CAT1) THEN VAR_GROUP=1;
  IF (&GROUP=&CAT2) THEN VAR_GROUP=2;
  IF (VAR_GROUP IN (1,2));
  KEEP &REPLI_ROOT.0-&REPLI_ROOT.80 &BYVAR &VAR VAR_GROUP;
RUN;

PROC SORT DATA=BRRDATA;
  BY &BYVAR VAR_GROUP;
RUN;

%DO I = 0 %TO 80;
  PROC MEANS DATA=BRRDATA VARDEF=WGT NOPRINT;
    VAR &VAR ;
    BY &BYVAR VAR_GROUP;
    WEIGHT &REPLI_ROOT&I;
    OUTPUT OUT=MEAN_TEMP MEAN=STAT1 VAR=STAT2;
  RUN;

  DATA MEAN_TEMP;
    SET MEAN_TEMP;
    L=&I;
  RUN;

  PROC APPEND BASE = BRR_TEMP1 DATA=MEAN_TEMP;
  RUN;

%END;

PROC SORT DATA=BRR_TEMP1;
  BY &BYVAR L VAR_GROUP;
RUN;

```



Box 17.14 [2/2] SAS® macro of EFFECT_SIZE_NO_PV.sas

```

PROC TRANSPOSE DATA=BRR_TEMP1 OUT=BRR_TEMP2 PREFIX=M;
    VAR STAT1 ;
    BY &BYVAR L;
    ID VAR_GROUP;
RUN;
PROC TRANSPOSE DATA=BRR_TEMP1 OUT=BRR_TEMP3 PREFIX=V;
    VAR STAT2 ;
    BY &BYVAR L;
    ID VAR_GROUP;
RUN;
DATA BRR_TEMP4;
    MERGE BRR_TEMP2 BRR_TEMP3;
    BY &BYVAR L;
    PV=(M1-M2)/(((V1+V2)/2)**0.5);
    KEEP &BYVAR L PV;
RUN;

DATA BRR_TEMP5 (KEEP=&BYVAR PV) BRR_TEMP6 (KEEP=&BYVAR STAT);
    SET BRR_TEMP4;
    IF L > 0 THEN OUTPUT BRR_TEMP5;
    ELSE DO;
        STAT =PV;
        OUTPUT BRR_TEMP6;
    END;
RUN;

PROC SORT DATA=BRR_TEMP5;
    BY &BYVAR;
RUN;

PROC SORT DATA=BRR_TEMP6;
    BY &BYVAR;
RUN;

DATA BRR_TEMP7;
    MERGE BRR_TEMP5 BRR_TEMP6;
    BY &BYVAR;
    VARI=((PV-STAT)**2)*(1/20);
RUN;

PROC UNIVARIATE DATA=BRR_TEMP7 NOPRINT;
    VAR VARI;
    BY &BYVAR;
    OUTPUT OUT=BRR_TEMP8 SUM=SS;
RUN;

DATA &OUTFILE;
    MERGE BRR_TEMP6 BRR_TEMP8;
    BY &BYVAR;
    SESTAT=(SS)**0.5;
    FORMAT STAT F10.2;
    FORMAT SESTAT F10.2;
    KEEP &BYVAR STAT SESTAT;
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRRDATA MEAN_TEMP BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5
    BRR_TEMP6 BRR_TEMP7 BRR_TEMP8;
RUN;

OPTIONS NOTES;

%MEND;

```



Box 17.15 [1/3] SAS® macro of EFFECT_SIZE_PV.sas

```

%MACRO BRR_EFFECT_PV(INFILE=,
                     REPLI_ROOT=,
                     BYVAR=,
                     PV_ROOT=,
                     EFFECT=,
                     OUTFILE=);

/*

MEANING OF THE MACRO ARGUMENTS

INFILE = INPUT DATA FILE.
REPLI_ROOT = ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL
WEIGHT VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
BYVAR = BREAKDOWN VARIABLES.
PV_ROOT = ROOT OF THE 5 PLAUSIBLE VALUES VARIABLES NAMES.
EFFECT = THIS ARGUMENT HAS THREE COMPONENTS: THE VARIABLE NAME AND THE TWO
CATEGORIES ON WHICH THE EFFECT SIZE WILL BE COMPUTED. THE EFFECT SIZE WILL
BE THE SECOND COMPONENT MINUS THE THIRD COMPONENT.
OUTFILE = FILE WITH THE STATISTIC ESTIMATES AND THEIR STANDARD ERROR ESTIMATES.

*/

%LET GROUP=%SCAN(&EFFECT,1);
%LET CAT1=%SCAN(&EFFECT,2);
%LET CAT2=%SCAN(&EFFECT,3);

OPTIONS NONOTES;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP1;
RUN;

DATA BRRDATA;
    SET &INFILE;
    IF (&GROUP=&CAT1) THEN VAR_GROUP=1;
    IF (&GROUP=&CAT2) THEN VAR_GROUP=2;
    IF (VAR_GROUP IN (1,2));
    KEEP &REPLI_ROOT.0-&REPLI_ROOT.80 &PV_ROOT.1-&PV_ROOT.5 &BYVAR VAR_GROUP;
RUN;

PROC SORT DATA=BRRDATA;
    BY &BYVAR VAR_GROUP;
RUN;

%DO I = 0 %TO 80;
    PROC MEANS DATA=BRRDATA VARDEF=WGT NOPRINT;
        VAR &PV_ROOT.1 - &PV_ROOT.5 ;
        BY &BYVAR VAR_GROUP;
        WEIGHT &REPLI_ROOT&I;
        OUTPUT OUT=MEAN_TEMP MEAN=MOY1-MOY5 VAR=VARI1-VARI5;
    RUN;

    DATA MEAN_TEMP;
        SET MEAN_TEMP;
        L=&I;
    RUN;

    PROC APPEND BASE = BRR_TEMP1 DATA=MEAN_TEMP;
    RUN;

%END;

PROC SORT DATA=BRR_TEMP1;
    BY &BYVAR L VAR_GROUP;
RUN;

```



Box 17.15 [2/3] SAS® macro of EFFECT_SIZE_PV.sas

```

PROC TRANSPOSE DATA=BRR_TEMP1 (RENAME=(MOY1=V1 MOY2=V2 MOY3=V3 MOY4=V4 MOY5=V5))
OUT=BRR_TEMP2 PREFIX=M;
  VAR V1-V5
  ;
  BY &BYVAR L;
  ID VAR_GROUP;
RUN;
PROC TRANSPOSE DATA=BRR_TEMP1 (RENAME=(VARI1=V1 VARI2=V2 VARI3=V3 VARI4=V4
VARI5=V5))OUT=BRR_TEMP3 PREFIX=V;
  VAR V1-V5
  ;
  BY &BYVAR L;
  ID VAR_GROUP;
RUN;
DATA BRR_TEMP4;
  MERGE BRR_TEMP2 BRR_TEMP3;
  BY &BYVAR L _NAME_;
  PV=(M1-M2)/(((V1+V2)/2)**0.5);
  KEEP &BYVAR L PV _NAME_;
RUN;
PROC TRANSPOSE DATA=BRR_TEMP4 OUT=BRR_TEMP5 (DROP=_NAME_) PREFIX=P;
  VAR PV;
  BY &BYVAR L;
  ID _NAME_;
RUN;

DATA BRR_TEMP6 (DROP=STAT FIN1-FIN5 MESVAR) BRR_TEMP7 (KEEP=&BYVAR STAT FIN1-FIN5
MESVAR);
  SET BRR_TEMP5;
  IF L > 0 THEN OUTPUT BRR_TEMP6;
  ELSE DO;
    STAT = (PV1+PV2+PV3+PV4+PV5)/5;
    FIN1=PV1;
    FIN2=PV2;
    FIN3=PV3;
    FIN4=PV4;
    FIN5=PV5;
    MESVAR=(((STAT-FIN1)**2)+((STAT-FIN2)**2)+((STAT-FIN3)**2)+((STAT-
FIN4)**2)+((STAT-FIN5)**2))/4;
    OUTPUT BRR_TEMP7;
  END;
RUN;

PROC SORT DATA=BRR_TEMP6;
  BY &BYVAR;
RUN;

PROC SORT DATA=BRR_TEMP7;
  BY &BYVAR;
RUN;

DATA BRR_TEMP8;
  MERGE BRR_TEMP6 BRR_TEMP7;
  BY &BYVAR;
  ARRAY A (5)
    PV1-PV5;
  ARRAY B (5)
    FIN1-FIN5;
  ARRAY C (5)
    VAR1-VAR5;
  DO I=1 TO 5;
    C(I)=(1/20)*((A(I)-B(I))**2);
  END;
RUN;

PROC UNIVARIATE DATA=BRR_TEMP8 NOPRINT;
  VAR VAR1 VAR2 VAR3 VAR4 VAR5;
  BY &BYVAR;
  OUTPUT OUT=BRR_TEMP9 SUM=SS1 SS2 SS3 SS4 SS5;
RUN;

```



Box 17.15 [3/3] SAS® macro of EFFECT_SIZE_PV.sas

```
DATA &OUTFILE;
  MERGE BRR_TEMP7 BRR_TEMP9;
  BY &BYVAR;
  SAMP=(SS1+SS2+SS3+SS4+SS5)/5;
  FINVAR=(SAMP+(1.2*MESVAR));
  SESTAT=(FINVAR)**0.5;
  FORMAT STAT F10.2;
  FORMAT SESTAT F10.2;
  KEEP &BYVAR STAT SESTAT ;
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRRDATA MEAN_TEMP BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5
  BRR_TEMP6 BRR_TEMP7 BRR_TEMP8 BRR_TEMP9;
RUN;

OPTIONS NOTES;

%MEND;
```



Box 17.16 [1/5] SAS® macro of PROC_MIXED_NO_PV.sas

```

%macro BRR_MIXED(INFILE=,
                 REPLI_ROOT=,
                 VARDEP=,
                 FIXEF=,
                 RANEF=,
                 BYVAR=,
                 LEVEL2=,
                 OUTSCREEN=,
                 OUTFILE=);

/*
MEANING OF THE MACRO ARGUMENTS

INFILE =      INPUT DATA FILE.
REPLI_ROOT =  ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL
WEIGHT VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
VARDEP =      DEPENDENT VARIABLES.
FIXEF = LIST OF INDEPENDENT VARIABLES WITH FIXED EFFECT.
RANEF = LIST OF INDEPENDENT VARIABLES WITH RANDOM EFFECT.
LEVEL2 =      LEVEL 2 IDENTIFICATION VARIABLE.
BYVAR = BREAKDOWN VARIABLES.
OUTSCREEN     =      DOS ADDRESS FOR EXPORTING LISTING.
OUTFILE =     FILE WITH THE STATISTIC ESTIMATES AND THEIR STANDARD ERROR
ESTIMATES.

*/

OPTION NONOTES CLEANUP;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP0;
RUN;

PROC SORT DATA=&INFILE
    OUT=BRR_PREP1(KEEP=&REPLI_ROOT.0-&REPLI_ROOT.80  &VARDEP &BYVAR &FIXEF
&RANEF &LEVEL2);
    BY &BYVAR;
RUN;

%LET I=1;
%DO %WHILE(%LENGTH(%SCAN(&FIXEF,&I)));
    %LET I=%EVAL(&I+1);
%END;
%LET NB_FIXE=%EVAL(&I-1);

%LET I=1;
%DO %WHILE(%LENGTH(%SCAN(&RANEF,&I)));
    %LET I=%EVAL(&I+1);
%END;
%LET NB_RAN=%EVAL(&I-1);
%LET NB_TOT=%EVAL(&NB_FIXE+&NB_RAN+1);

DATA BRR_PREP2;
    SET BRR_PREP1;
    NB_MISS=0;
    ARRAY LIST_VAR (&NB_TOT) &VARDEP &FIXEF &RANEF;
    DO K=1 TO &NB_TOT;
        IF (LIST_VAR(K) IN (.,.I,.M,.N)) THEN NB_MISS=NB_MISS+1;
    END;
    IF (NB_MISS>1) THEN NB_MISS=1;
RUN;

PROC FREQ DATA=BRR_PREP2 NOPRINT;
    TABLE NB_MISS/OUT=&OUTFILE._DELETION;
    BY &BYVAR;
    WEIGHT &REPLI_ROOT.0;
RUN;

```



Box 17.16 [2/5] SAS® macro of PROC_MIXED_NO_PV.sas

```

PROC UNIVARIATE DATA=BRR_PREP2 NOPRINT;
  VAR &REPLI_ROOT.0;
  BY &BYVAR;
  WHERE (NB_MISS=0);
  OUTPUT OUT=BRR_PREP3 N=NBRE SUM=SOMWGT;
RUN;

DATA BRRDATA;
  MERGE BRR_PREP2 BRR_PREP3;
  BY &BYVAR;
  ARRAY OLDWGT (81)&REPLI_ROOT.0-&REPLI_ROOT.80;
  ARRAY NEWWGT (81) STD_WGT0-STD_WGT80;
  DO I=1 TO 81;
    NEWWGT(I) = (OLDWGT(I) / SOMWGT) * NBRE;
  END;
RUN;

%DO J = 0 %TO 80;

  PROC MIXED DATA=BRRDATA NOCLPRINT NOITPRINT NOINFO METHOD=ML;
    CLASS &LEVEL2;
    MODEL &VARDEP=&FIXEF &RANEF /SOLUTION;
    RANDOM INTERCEPT &RANEF /SUBJECT=&LEVEL2;
    WEIGHT STD_WGT&J;
    BY &BYVAR;
    ODS OUTPUT SOLUTIONF=FIXE_TEMP COVPARMS=COV_TEMP;
  RUN;

  DATA FIXE_TEMP;
    SET FIXE_TEMP;
    L=&J;
    KEEP &BYVAR L EFFECT ESTIMATE ;
  RUN;
  PROC SORT DATA=FIXE_TEMP;
    by &BYVAR L EFFECT;
  RUN;
  DATA COV_TEMP;
    SET COV_TEMP;
    L=&J;
    KEEP &BYVAR L COVPARM ESTIMATE ;
  RUN;
  PROC SORT DATA=COV_TEMP;
    BY &BYVAR L COVPARM;
  RUN;

  %IF (&J=0) %then %do;
    DATA BRR_TEMP_FIXE;
      SET FIXE_TEMP;
    RUN;
    DATA BRR_TEMP_VAR;
      SET COV_TEMP;
    RUN;
  %END;

  DATA BRR_TEMP_FIXE;
    MERGE BRR_TEMP_FIXE FIXE_TEMP;
    BY &BYVAR L EFFECT;
  RUN;
  DATA BRR_TEMP_VAR;
    MERGE BRR_TEMP_VAR COV_TEMP;
    BY &BYVAR L COVPARM;
  RUN;

%END;

/* FIXE PARAMETERS */

PROC SORT DATA=BRR_TEMP_FIXE OUT=BRR_TEMP1;
  BY &BYVAR L EFFECT;
RUN;

```




Box 17.16 [3/5] SAS® macro of PROC_MIXED_NO_PV.sas

```

DATA BRR_TEMP2 (KEEP=&BYVAR EFFECT ESTIMATE) BRR_TEMP3 (KEEP=&BYVAR EFFECT STAT);
  SET BRR_TEMP1;
  IF L > 0 THEN OUTPUT BRR_TEMP2;
  ELSE DO;
    STAT =ESTIMATE;
    OUTPUT BRR_TEMP3;
  END;
RUN;

PROC SORT DATA=BRR_TEMP2;
  BY &BYVAR EFFECT;
RUN;

PROC SORT DATA=BRR_TEMP3;
  BY &BYVAR EFFECT;
RUN;

DATA BRR_TEMP4;
  MERGE BRR_TEMP2 BRR_TEMP3;
  BY &BYVAR EFFECT;
  VARI=((ESTIMATE-STAT)**2)*(1/20);
RUN;

PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
  VAR VARI;
  BY &BYVAR EFFECT;
  OUTPUT OUT=BRR_TEMP5 SUM=SS;
RUN;

DATA &OUTFILE._FIXE;
  MERGE BRR_TEMP3 BRR_TEMP5;
  BY &BYVAR EFFECT;
  SESTAT=(SS)**0.5;
  FORMAT STAT F10.2;
  FORMAT SESTAT F10.2;
  KEEP &BYVAR EFFECT STAT SESTAT;
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_PREP1 BRR_PREP2 BRR_PREP3 BRR_TEMP_FIXE FIXE_TEMP BRR_TEMP1
  BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5 MEAN_TEMP BRRDATA;
RUN;

/* RANDOM PARAMETERS */

PROC SORT DATA=BRR_TEMP VAR OUT=BRR_TEMP1;
  BY &BYVAR L COVPARM;
RUN;

DATA BRR_TEMP2 (KEEP=&BYVAR COVPARM ESTIMATE) BRR_TEMP3 (KEEP=&BYVAR COVPARM STAT);
  SET BRR_TEMP1;
  IF L > 0 THEN OUTPUT BRR_TEMP2;
  ELSE DO;
    STAT =ESTIMATE;
    OUTPUT BRR_TEMP3;
  END;
RUN;

PROC SORT DATA=BRR_TEMP2;
  BY &BYVAR COVPARM;
RUN;

PROC SORT DATA=BRR_TEMP3;
  BY &BYVAR COVPARM;
RUN;

DATA BRR_TEMP4;
  MERGE BRR_TEMP2 BRR_TEMP3;
  BY &BYVAR COVPARM;
  VARI=((ESTIMATE-STAT)**2)*(1/20);
RUN;

```



Box 17.16 [4/5] **SAS® macro of PROC_MIXED_NO_PV.sas**

```

PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
  VAR VARI;
  BY &BYVAR COVPARM;
  OUTPUT OUT=BRR_TEMP5 SUM=SS;
RUN;

DATA &OUTFILE._VARIANCE;
  MERGE BRR_TEMP3 BRR_TEMP5;
  BY &BYVAR COVPARM;
  SESTAT=(SS)**0.5;
  FORMAT STAT F10.2;
  FORMAT SESTAT F10.2;
  KEEP &BYVAR COVPARM STAT SESTAT;
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE COV_TEMP BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5 MEAN_
TEMP BRRDATA;
RUN;

/* INTRACLASS CORRELATION */

PROC SORT DATA=BRR_TEMP_VAR ;
  BY &BYVAR L COVPARM;
RUN;

PROC TRANSPOSE DATA=BRR_TEMP_VAR OUT=BRR_TEMP0;
  VAR ESTIMATE;
  BY &BYVAR L;
  ID COVPARM;
RUN;

DATA BRR_TEMP1;
  SET BRR_TEMP0;
  ESTIMATE=(INTERCEPT/(INTERCEPT+RESIDUAL));
  KEEP &BYVAR L ESTIMATE;
RUN;

DATA BRR_TEMP2 (KEEP=&BYVAR ESTIMATE) BRR_TEMP3 (KEEP=&BYVAR STAT);
  SET BRR_TEMP1;
  IF L > 0 THEN OUTPUT BRR_TEMP2;
  ELSE DO;
    STAT =ESTIMATE;
    OUTPUT BRR_TEMP3;
  END;
RUN;

PROC SORT DATA=BRR_TEMP2;
  BY &BYVAR;
RUN;

PROC SORT DATA=BRR_TEMP3;
  BY &BYVAR ;
RUN;

DATA BRR_TEMP4;
  MERGE BRR_TEMP2 BRR_TEMP3;
  BY &BYVAR ;
  VARI=((ESTIMATE-STAT)**2)*(1/20);
RUN;

PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
  VAR VARI;
  BY &BYVAR ;
  OUTPUT OUT=BRR_TEMP5 SUM=SS;
RUN;

```



Box 17.16 [5/5] SAS® macro of PROC_MIXED_NO_PV.sas

```
DATA &OUTFILE._INTRACLASS;
  MERGE BRR_TEMP3 BRR_TEMP5;
  BY &BYVAR ;
  SESTAT=(SS)**0.5;
  FORMAT STAT F10.2;
  FORMAT SESTAT F10.2;
  KEEP &BYVAR STAT SESTAT;
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE COV_TEMP BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5 MEAN_
TEMP BRRDATA;
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP_VAR COV_TEMP BRR_TEMP0 BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_
TEMP4 BRR_TEMP5 MEAN_TEMP BRRDATA;
RUN;

OPTIONS NOTES;

%MEND;
```



Box 17.17 [1/6] SAS® macro of PROC_MIXED_PV.sas

```

%macro BRR_MIXED_PV(INFILE=,
                    REPLI_ROOT=,
                    PV_ROOT=,
                    FIXEF=,
                    RANEF=,
                    BYVAR=,
                    LEVEL2=,
                    OUTSCREEN=,
                    OUTFILE=,);

/*
MEANING OF THE MACRO ARGUMENTS

INFILE = INPUT DATA FILE.
REPLI_ROOT = ROOT OF THE FINAL WEIGHT AND 80 REPLICATES VARIABLE NAMES. FINAL
WEIGHT VARIABLE NAME MUST BE THE REPLICATION ROOT FOLLOWED BY 0.
PV_ROOT = ROOT OF THE 5 PROFICIENCY LEVEL VARIABLES NAMES
FIXEF = LIST OF INDEPENDENT VARIABLES WITH FIXED EFFECT
RANEF = LIST OF INDEPENDENT VARIABLES WITH RANDOM EFFECT
LEVEL2 = LEVEL 2 IDENTIFICATION VARIABLE
BYVAR = BREAKDOWN VARIABLES
OUTSCREEN = DOS ADDRESS FOR EXPORTING LISTING
OUTFILE = FILE WITH THE STATISTIC ESTIMATES AND THEIR STANDARD ERROR ESTIMATES.

*/

OPTION NONOTES CLEANUP;

FILENAME MYOUTPUT &OUTSCREEN;
PROC PRINTTO PRINT=MYOUTPUT NEW;
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
    DELETE BRR_TEMP0;
RUN;

PROC SORT DATA=&INFILE
    OUT=BRR_PREP1(KEEP=&REPLI_ROOT.0-&REPLI_ROOT.80 &PV_ROOT.1-&PV_ROOT.5
&BYVAR &FIXEF &RANEF &LEVEL2);
    BY &BYVAR;
RUN;

%LET I=1;
%DO %WHILE(%LENGTH(%SCAN(&FIXEF,&I)));
    %LET I=%EVAL(&I+1);
%END;
%LET NB_FIXE=%EVAL(&I-1);

%LET I=1;
%DO %WHILE(%LENGTH(%SCAN(&RANEF,&I)));
    %LET I=%EVAL(&I+1);
%END;
%LET NB_RAN=%EVAL(&I-1);
%LET NB_TOT=%EVAL(&NB_FIXE+&NB_RAN+5);

DATA BRR_PREP2;
    SET BRR_PREP1;
    NB_MISS=0;
    ARRAY LIST_VAR (&NB_TOT) &PV_ROOT.1-&PV_ROOT.5 &FIXEF &RANEF;
    DO K=1 TO &NB_TOT;
        IF (LIST_VAR(K) IN (.,.I,.M,.N)) THEN NB_MISS=NB_MISS+1;
    END;
    IF (NB_MISS>1) THEN NB_MISS=1;
RUN;

PROC FREQ DATA=BRR_PREP2 NOPRINT;
    TABLE NB_MISS/OUT=&OUTFILE._DELETION;
    BY &BYVAR;
    WEIGHT &REPLI_ROOT.0;
RUN;

```



Box 17.17 [2/6] SAS® macro of PROC_MIXED_PV.sas

```

PROC UNIVARIATE DATA=BRR_PREP2 NOPRINT;
  VAR &REPLI_ROOT.0;
  BY &BYVAR;
  WHERE (NB_MISS=0);
  OUTPUT OUT=BRR_PREP3 N=NBRE SUM=SOMWGT;
RUN;

DATA BRRDATA;
  MERGE BRR_PREP2 BRR_PREP3;
  BY &BYVAR;
  ARRAY OLDWGT (81)&REPLI_ROOT.0-&REPLI_ROOT.80;
  ARRAY NEWWGT (81) STD_WGT0-STD_WGT80;
  DO I=1 TO 81;
    NEWWGT(I) = (OLDWGT(I) / SOMWGT) * NBRE;
  END;
RUN;

%DO I = 1 %TO 5;

  %DO J = 0 %TO 80;

    PROC MIXED DATA=BRRDATA NOCLPRINT NOITPRINT NOINFO METHOD=ML;
      CLASS &LEVEL2;
      MODEL &PV_ROOT&I=&FIXEF &RANEF /SOLUTION;
      RANDOM INTERCEPT &RANEF /SUBJECT=&LEVEL2;
      WEIGHT STD_WGT&J;
      BY &BYVAR;
      ODS OUTPUT SOLUTIONF=FIXE_TEMP COVPARMS=COV_TEMP;
    RUN;

    DATA FIXE_TEMP;
      SET FIXE_TEMP;
      L=&J;
      PV&I=ESTIMATE;
      KEEP &BYVAR L EFFECT PV&I ;
    RUN;
    PROC SORT DATA=FIXE_TEMP;
      by &BYVAR L EFFECT;
    RUN;
    DATA COV_TEMP;
      SET COV_TEMP;
      L=&J;
      PV&I=ESTIMATE;
      KEEP &BYVAR L COVPARAM PV&I ;
    RUN;
    PROC SORT DATA=COV_TEMP;
      BY &BYVAR L COVPARAM;
    RUN;

    %IF (&I=1 AND &J=0) %then %do;
      DATA BRR_TEMP_FIXE;
        SET FIXE_TEMP;
      RUN;
      DATA BRR_TEMP_VAR;
        SET COV_TEMP;
      RUN;
    %END;

    DATA BRR_TEMP_FIXE;
      MERGE BRR_TEMP_FIXE FIXE_TEMP;
      BY &BYVAR L EFFECT;
    RUN;
    DATA BRR_TEMP_VAR;
      MERGE BRR_TEMP_VAR COV_TEMP;
      BY &BYVAR L COVPARAM;
    RUN;

  %END;

%END;

```



Box 17.17 [3/6] SAS® macro of PROC_MIXED_PV.sas

```

PROC SORT DATA=BRR_TEMP_FIXE OUT=BRR_TEMP1;
  BY &BYVAR L EFFECT;
RUN;

DATA BRR_TEMP2 (DROP=STAT FIN1-FIN5 MESVAR) BRR_TEMP3 (KEEP=&BYVAR EFFECT STAT FIN1-
FIN5 MESVAR);
  SET BRR_TEMP1;
  IF L > 0 THEN OUTPUT BRR_TEMP2;
  ELSE DO;
    STAT = (PV1+PV2+PV3+PV4+PV5) / 5;
    FIN1=PV1;
    FIN2=PV2;
    FIN3=PV3;
    FIN4=PV4;
    FIN5=PV5;
    MESVAR=(((STAT-FIN1)**2)+((STAT-FIN2)**2)+((STAT-FIN3)**2)+((STAT-
FIN4)**2)+((STAT-FIN5)**2))/4;
    OUTPUT BRR_TEMP3;
  END;
RUN;

PROC SORT DATA=BRR_TEMP2;
  BY &BYVAR EFFECT L;
RUN;

PROC SORT DATA=BRR_TEMP3;
  BY &BYVAR EFFECT;
RUN;

DATA BRR_TEMP4;
  MERGE BRR_TEMP2 BRR_TEMP3;
  BY &BYVAR EFFECT;
  ARRAY A (5)
    PV1-PV5;
  ARRAY B (5)
    FIN1-FIN5;
  ARRAY C (5)
    VAR1-VAR5;
  DO I=1 TO 5;
    C(I) = (1/20)*((A(I)-B(I))**2);
  END;
RUN;

PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
  VAR VAR1 VAR2 VAR3 VAR4 VAR5;
  BY &BYVAR EFFECT;
  OUTPUT OUT=BRR_TEMP5 SUM=SS1 SS2 SS3 SS4 SS5;
RUN;

DATA &OUTFILE._FIXE;
  MERGE BRR_TEMP3 BRR_TEMP5;
  BY &BYVAR EFFECT;
  SAMP=(SS1+SS2+SS3+SS4+SS5)/5;
  FINVAR=(SAMP+(1.2*MESVAR));
  SESTAT=(FINVAR)**0.5;
  FORMAT STAT F10.1;
  FORMAT SESTAT F10.2;
  KEEP &BYVAR EFFECT STAT SESTAT;
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP_FIXE FIXE_TEMP BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4
BRR_TEMP5 MEAN_TEMP BRRDATA;
RUN;

PROC SORT DATA=BRR_TEMP_VAR OUT=BRR_TEMP1;
  BY &BYVAR L COVPARM;
RUN;

```



Box 17.17 [4/6] SAS® macro of PROC_MIXED_PV.sas

```

DATA BRR_TEMP2 (DROP=STAT FIN1-FIN5 MESVAR) BRR_TEMP3 (KEEP=&BYVAR COVPARM STAT FIN1-
FIN5 MESVAR);
  SET BRR_TEMP1;
  IF L > 0 THEN OUTPUT BRR_TEMP2;
  ELSE DO;
    STAT = (PV1+PV2+PV3+PV4+PV5) / 5;
    FIN1=PV1;
    FIN2=PV2;
    FIN3=PV3;
    FIN4=PV4;
    FIN5=PV5;
    MESVAR=(((STAT-FIN1)**2) + ((STAT-FIN2)**2) + ((STAT-FIN3)**2) + ((STAT-
FIN4)**2) + ((STAT-FIN5)**2)) / 4;
    OUTPUT BRR_TEMP3;
  END;
RUN;

PROC SORT DATA=BRR_TEMP2;
  BY &BYVAR COVPARM L;
RUN;

PROC SORT DATA=BRR_TEMP3;
  BY &BYVAR COVPARM;
RUN;

DATA BRR_TEMP4;
  MERGE BRR_TEMP2 BRR_TEMP3;
  BY &BYVAR COVPARM;
  ARRAY A (5)
    PV1-PV5;
  ARRAY B (5)
    FIN1-FIN5;
  ARRAY C(5)
    VAR1-VAR5;
  DO I=1 TO 5;
    C(I) = (1/20) * ((A(I) - B(I))**2);
  END;
RUN;

PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
  VAR VAR1 VAR2 VAR3 VAR4 VAR5;
  BY &BYVAR COVPARM;
  OUTPUT OUT=BRR_TEMP5 SUM=SS1 SS2 SS3 SS4 SS5;
RUN;

DATA &OUTFILE._VARIANCE;
  MERGE BRR_TEMP3 BRR_TEMP5;
  BY &BYVAR COVPARM;
  SAMP = (SS1+SS2+SS3+SS4+SS5) / 5;
  FINVAR = (SAMP + (1.2*MESVAR));
  SESTAT = (FINVAR)**0.5;
  FORMAT STAT F10.1;
  FORMAT SESTAT F10.2;
  KEEP &BYVAR COVPARM STAT SESTAT;
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_TEMP4 BRR_TEMP5 MEAN_TEMP
/*BRRDATA*/;
RUN;

PROC SORT DATA=BRR_TEMP_VAR ;
  BY &BYVAR L COVPARM;
RUN;

%DO K=1 %TO 5;

PROC TRANSPOSE DATA=BRR_TEMP_VAR OUT=BRR_TEMP0;
  VAR PV&K;
  BY &BYVAR L;
  ID COVPARM;
RUN;

```



Box 17.17 [5/6] SAS® macro of PROC_MIXED_PV.sas

```

DATA BRR_TEMP0;
  SET BRR_TEMP0;
  PV&K=(INTERCEPT/(INTERCEPT+RESIDUAL));
  KEEP &BYVAR L PV&K;
RUN;

%IF (&K=1) %THEN %DO;
  DATA BRR_TEMP1;
    SET BRR_TEMP0;
  RUN;
%END;
DATA BRR_TEMP1;
  MERGE BRR_TEMP1 BRR_TEMP0;
  BY &BYVAR L;
RUN;
%END;

DATA BRR_TEMP2 (DROP=STAT FIN1-FIN5 MESVAR) BRR_TEMP3 (KEEP=&BYVAR STAT FIN1-FIN5
MESVAR);
  SET BRR_TEMP1;
  IF L > 0 THEN OUTPUT BRR_TEMP2;
  ELSE DO;
    STAT = (PV1+PV2+PV3+PV4+PV5) /5;
    FIN1=PV1;
    FIN2=PV2;
    FIN3=PV3;
    FIN4=PV4;
    FIN5=PV5;
    MESVAR=(((STAT-FIN1)**2)+((STAT-FIN2)**2)+((STAT-FIN3)**2)+((STAT-
FIN4)**2)+((STAT-FIN5)**2))/4;
    OUTPUT BRR_TEMP3;
  END;
RUN;

PROC SORT DATA=BRR_TEMP2;
  BY &BYVAR L;
RUN;

PROC SORT DATA=BRR_TEMP3;
  BY &BYVAR ;
RUN;

DATA BRR_TEMP4;
  MERGE BRR_TEMP2 BRR_TEMP3;
  BY &BYVAR ;
  ARRAY A (5)
    PV1-PV5;
  ARRAY B (5)
    FIN1-FIN5;
  ARRAY C (5)
    VAR1-VAR5;
  DO I=1 TO 5;
    C(I) = (1/20) * ((A(I) -B(I)) **2);
  END;
RUN;

PROC UNIVARIATE DATA=BRR_TEMP4 NOPRINT;
  VAR VAR1 VAR2 VAR3 VAR4 VAR5;
  BY &BYVAR ;
  OUTPUT OUT=BRR_TEMP5 SUM=SS1 SS2 SS3 SS4 SS5;
RUN;

```




Box 17.17 [6/6] SAS® macro of PROC_MIXED_PV.sas

```
DATA &OUTFILE._INTRACLASS;
  MERGE BRR_TEMP3 BRR_TEMP5;
  BY &BYVAR ;
  SAMP=(SS1+SS2+SS3+SS4+SS5)/5;
  FINVAR=(SAMP+(1.2*MESVAR));
  SESTAT=(FINVAR)**0.5;
  FORMAT STAT F10.2;
  FORMAT SESTAT F10.2;
  KEEP &BYVAR STAT SESTAT;
RUN;

PROC DATASETS LIBRARY=WORK NOLIST;
  DELETE BRR_TEMP_VAR COV_TEMP BRR_TEMP0 BRR_TEMP1 BRR_TEMP2 BRR_TEMP3 BRR_
TEMP4 BRR_TEMP5 MEAN_TEMP
  BRR_PREP1 BRR_PREP2 BRR_PREP3;
RUN;

PROC PRINTTO PRINT=PRINT;
RUN;

OPTIONS NOTES;

%MEND;
```



References

- Beaton, A.E.** (1987), *The NAEP 1983-1984 Technical Report*, Educational Testing Service, Princeton.
- Beaton, A.E., et al.** (1996), *Mathematics Achievement in the Middle School Years, IEA's Third International Mathematics and Science Study*, Boston College, Chestnut Hill, MA.
- Bloom, B.S.** (1979), *Caractéristiques individuelles et apprentissage scolaire*, Éditions Labor, Brussels.
- Bressoux, P.** (2008), *Modélisation statistique appliquée aux sciences sociales*, De Boeck, Brussels.
- Bryk, A.S. and S.W. Raudenbush** (1992), *Hierarchical Linear Models for Social and Behavioural Research: Applications and Data Analysis Methods*, Sage Publications, Newbury Park, CA.
- Buchmann, C.** (2000), *Family structure, parental perceptions and child labor in Kenya: What factors determine who is enrolled in school?* *aSoc. Forces*, No. 78, pp. 1349-79.
- Cochran, W.G.** (1977), *Sampling Techniques*, J. Wiley and Sons, Inc., New York.
- Dunn, O.J.** (1961), "Multiple Comparisons among Menas", *Journal of the American Statistical Association*, Vol. 56, American Statistical Association, Alexandria, pp. 52-64.
- Kish, L.** (1995), *Survey Sampling*, J. Wiley and Sons, Inc., New York.
- Knighton, T. and P. Bussière** (2006), "Educational Outcomes at Age 19 Associated with Reading Ability at Age 15", Statistics Canada, Ottawa.
- Gonzalez, E. and A. Kennedy** (2003), *PIRLS 2001 User Guide for the International Database*, Boston College, Chestnut Hill, MA.
- Ganzeboom, H.B.G., P.M. De Graaf and D.J. Treiman** (1992), "A Standard International Socio-economic Index of Occupation Status", *Social Science Research* 21(1), Elsevier Ltd, pp 1-56.
- Goldstein, H.** (1995), *Multilevel Statistical Models*, 2nd Edition, Edward Arnold, London.
- Goldstein, H.** (1997), "Methods in School Effectiveness Research", *School Effectiveness and School Improvement* 8, Swets and Zeitlinger, Lisse, Netherlands, pp. 369-395.
- Hubin, J.P.** (ed.) (2007), *Les indicateurs de l'enseignement*, 2nd Edition, Ministère de la Communauté française, Brussels.
- Husen, T.** (1967), *International Study of Achievement in Mathematics: A Comparison of Twelve Countries*, Almqvist and Wiksells, Uppsala.
- International Labour Organisation (ILO)** (1990), *International Standard Classification of Occupations: ISCO-88*. Geneva: International Labour Office.
- Lafontaine, D. and C. Monseur** (forthcoming), "Impact of Test Characteristics on Gender Equity Indicators in the Assessment of Reading Comprehension", *European Educational Research Journal*, Special Issue on PISA and Gender.
- Lietz, P.** (2006), "A Meta-Analysis of Gender Differences in Reading Achievement at the Secondary Level", *Studies in Educational Evaluation* 32, pp. 317-344.
- Monseur, C. and M. Crahay** (forthcoming), "Composition académique et sociale des établissements, efficacité et inégalités scolaires : une comparaison internationale – Analyse secondaire des données PISA 2006", *Revue française de pédagogie*.
- OECD** (1998), *Education at a Glance – OECD Indicators*, OECD, Paris.
- OECD** (1999a), *Measuring Student Knowledge and Skills – A New Framework for Assessment*, OECD, Paris.
- OECD** (1999b), *Classifying Educational Programmes – Manual for ISCED-97 Implementation in OECD Countries*, OECD, Paris.
- OECD** (2001), *Knowledge and Skills for Life – First Results from PISA 2000*, OECD, Paris.
- OECD** (2002a), *Programme for International Student Assessment – Manual for the PISA 2000 Database*, OECD, Paris.



- OECD (2002b), *Sample Tasks from the PISA 2000 Assessment – Reading, Mathematical and Scientific Literacy*, OECD, Paris.
- OECD (2002c), *Programme for International Student Assessment – PISA 2000 Technical Report*, OECD, Paris.
- OECD (2002d), *Reading for Change: Performance and Engagement across Countries – Results from PISA 2000*, OECD, Paris.
- OECD (2003a), *Literacy Skills for the World of Tomorrow – Further Results from PISA 2000*, OECD, Paris.
- OECD (2003b), *The PISA 2003 Assessment Framework – Mathematics, Reading, Science and Problem Solving Knowledge and Skills*, OECD, Paris.
- OECD (2004a), *Learning for Tomorrow's World – First Results from PISA 2003*, OECD, Paris.
- OECD (2004b), *Problem Solving for Tomorrow's World – First Measures of Cross-Curricular Competencies from PISA 2003*, OECD, Paris.
- OECD (2005a), *PISA 2003 Technical Report*, OECD, Paris.
- OECD (2005b), *PISA 2003 Data Analysis Manual*, OECD, Paris.
- OECD (2006), *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006*, OECD, Paris.
- OECD (2007), *PISA 2006: Science Competencies for Tomorrow's World*, OECD, Paris.
- OECD (2009), *PISA 2006 Technical Report*, OECD, Paris.
- Peaker, G.F. (1975), *An Empirical Study of Education in Twenty-One Countries: A Technical report. International Studies in Evaluation VIII*, Wiley, New York and Almqvist and Wiksell, Stockholm.
- Rust, K.F. and J.N.K. Rao (1996), "Variance Estimation for Complex Surveys Using Replication Techniques", *Statistical Methods in Medical Research*, Vol. 5, Hodder Arnold, London, pp. 283-310.
- Rutter, M., et al. (2004), "Gender Differences in Reading Difficulties: Findings from Four Epidemiology Studies", *Journal of the American Medical Association* 291, pp. 2007-2012.
- Schulz, W. (2006), *Measuring the socio-economic background of students and its effect on achievement in PISA 2000 and PISA 2003*, Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.
- Wagemaker, H. (1996), *Are Girls Better Readers. Gender Differences in Reading Literacy in 32 Countries*, IEA, The Hague.
- Warm, T.A. (1989), "Weighted Likelihood Estimation of Ability in Item Response Theory", *Psychometrika*, Vol. 54(3), Psychometric Society, Williamsburg, VA., pp. 427-450.
- Wright, B.D. and M.H. Stone (1979), *Best Test Design: Rasch Measurement*, MESA Press, Chicago.



Table of contents

FOREWORD	3
USER'S GUIDE	17
CHAPTER 1 THE USEFULNESS OF PISA DATA FOR POLICY MAKERS, RESEARCHERS AND EXPERTS ON METHODOLOGY	19
PISA – an overview	20
▪ The PISA surveys.....	20
How can PISA contribute to educational policy, practice and research?	22
▪ Key results from PISA 2000, PISA 2003 and PISA 2006.....	23
Further analyses of PISA datasets	25
▪ Contextual framework of PISA 2006.....	28
▪ Influence of the methodology on outcomes.....	31
CHAPTER 2 EXPLORATORY ANALYSIS PROCEDURES	35
Introduction	36
Weights	36
Replicates for computing the standard error	39
Plausible values	43
Conclusion	46
CHAPTER 3 SAMPLE WEIGHTS	49
Introduction	50
Weights for simple random samples	51
Sampling designs for education surveys	53
Why do the PISA weights vary?	57
Conclusion	58
CHAPTER 4 REPLICATE WEIGHTS	59
Introduction	60
Sampling variance for simple random sampling	60
Sampling variance for two-stage sampling	65
Replication methods for simple random samples	70
Replication methods for two-stage samples	72
▪ The Jackknife for unstratified two-stage sample designs.....	72
▪ The Jackknife for stratified two-stage sample designs.....	73
▪ The Balanced Repeated Replication method.....	74
Other procedures for accounting for clustered samples	76
Conclusion	76

CHAPTER 5 THE RASCH MODEL	79
Introduction	80
How can the information be summarised?	80
The Rasch Model for dichotomous items	81
▪ Introduction to the Rasch Model.....	81
▪ Item calibration.....	85
▪ Computation of a student's score.....	87
▪ Computation of a student's score for incomplete designs.....	91
▪ Optimal conditions for linking items.....	92
▪ Extension of the Rasch Model.....	93
Other item response theory models	94
Conclusion	94
 CHAPTER 6 PLAUSIBLE VALUES	 95
Individual estimates versus population estimates	96
The meaning of plausible values (PVs)	96
Comparison of the efficiency of WLEs, EAP estimates and PVs for the estimation of some population statistics	99
How to perform analyses with plausible values	102
Conclusion	103
 CHAPTER 7 COMPUTATION OF STANDARD ERRORS	 105
Introduction	106
The standard error on univariate statistics for numerical variables	106
The SAS® macro for computing the standard error on a mean	109
The standard error on percentages	112
The standard error on regression coefficients	115
The standard error on correlation coefficients	117
Conclusion	117
 CHAPTER 8 ANALYSES WITH PLAUSIBLE VALUES	 119
Introduction	120
Univariate statistics on plausible values	120
The standard error on percentages with PVs	123
The standard error on regression coefficients with PVs	123
The standard error on correlation coefficients with PVs	126
Correlation between two sets of plausible values	126
A fatal error shortcut	130
An unbiased shortcut	131
Conclusion	133
 CHAPTER 9 USE OF PROFICIENCY LEVELS	 135
Introduction	136
Generation of the proficiency levels	136
Other analyses with proficiency levels	141
Conclusion	143



CHAPTER 10 ANALYSES WITH SCHOOL-LEVEL VARIABLES	145
Introduction	146
Limits of the PISA school samples	147
Merging the school and student data files	148
Analyses of the school variables	148
Conclusion	150
CHAPTER 11 STANDARD ERROR ON A DIFFERENCE	151
Introduction	152
Statistical issues and computing standard errors on differences	152
The standard error on a difference without plausible values	154
The standard error on a difference with plausible values	159
Multiple comparisons	163
Conclusion	164
CHAPTER 12 OECD TOTAL AND OECD AVERAGE	167
Introduction	168
Recoding of the database to estimate the pooled OECD total and the pooled OECD average	170
Duplication of the data to avoid running the procedure three times	172
Comparisons between the pooled OECD total or pooled OECD average estimates and a country estimate	173
Comparisons between the arithmetic OECD total or arithmetic OECD average estimates and a country estimate	175
Conclusion	175
CHAPTER 13 TRENDS	177
Introduction	178
The computation of the standard error for trend indicators on variables other than performance	179
The computation of the standard error for trend indicators on performance variables	181
Conclusion	185
CHAPTER 14 STUDYING THE RELATIONSHIP BETWEEN STUDENT PERFORMANCE AND INDICES DERIVED FROM CONTEXTUAL QUESTIONNAIRES	187
Introduction	188
Analyses by quarters	188
The concept of relative risk	190
▪ Instability of the relative risk	191
▪ Computation of the relative risk	192
Effect size	195
Linear regression and residual analysis	197
▪ Independence of errors	197
Statistical procedure	200
Conclusion	201



CHAPTER 15 MULTILEVEL ANALYSES	203
Introduction	204
Two-level modelling with SAS®	206
▪ Decomposition of the variance in the empty model.....	206
▪ Models with only random intercepts.....	209
▪ Shrinkage factor.....	213
▪ Models with random intercepts and fixed slopes.....	213
▪ Models with random intercepts and random slopes.....	215
▪ Models with Level 2 independent variables.....	220
▪ Computation of final estimates and their respective standard errors.....	223
Three-level modelling	225
Limitations of the multilevel model in the PISA context	227
Conclusion	228
CHAPTER 16 PISA AND POLICY RELEVANCE – THREE EXAMPLES OF ANALYSES	231
Introduction	232
Example 1: Gender differences in performance	232
Example 2: Promoting socio-economic diversity within school?	236
Example 3: The influence of an educational system on the expected occupational status of students at age 30	242
Conclusion	246
CHAPTER 17 SAS® MACRO	247
Introduction	248
Structure of the SAS® Macro	248
REFERENCES	313
APPENDICES	315
Appendix 1 Three-level regression analysis.....	316
Appendix 2 PISA 2006 International database.....	324
Appendix 3 PISA 2006 Student questionnaire.....	333
Appendix 4 PISA 2006 Information communication technology (ICT) Questionnaire.....	342
Appendix 5 PISA 2006 School questionnaire.....	344
Appendix 6 PISA 2006 Parent questionnaire.....	351
Appendix 7 Codebook for PISA 2006 student questionnaire data file.....	355
Appendix 8 Codebook for PISA 2006 non-scored cognitive and embedded attitude items.....	399
Appendix 9 Codebook for PISA 2006 scored cognitive and embedded attitude items.....	419
Appendix 10 Codebook for PISA 2006 school questionnaire data file.....	431
Appendix 11 Codebook for PISA 2006 parents questionnaire data file.....	442
Appendix 12 PISA 2006 questionnaire indices.....	448



LIST OF BOXES

Box 2.1	WEIGHT statement in the proc means procedure	37
<hr/>		
Box 7.1	SAS® syntax for computing 81 means (e.g. PISA 2003).....	106
Box 7.2	SAS® syntax for computing the mean of HISEI and its standard error (e.g. PISA 2003).....	109
Box 7.3	SAS® syntax for computing the standard deviation of HISEI and its standard error by gender (e.g. PISA 2003).....	112
Box 7.4	SAS® syntax for computing the percentages and their standard errors for gender (e.g. PISA 2003).....	112
Box 7.5	SAS® syntax for computing the percentages and its standard errors for grades by gender (e.g. PISA 2003).....	114
Box 7.6	SAS® syntax for computing regression coefficients, R ² and its respective standard errors: Model 1 (e.g. PISA 2003).....	115
Box 7.7	SAS® syntax for computing regression coefficients, R ² and its respective standard errors: Model 2 (e.g. PISA 2003).....	116
Box 7.8	SAS® syntax for computing correlation coefficients and its standard errors (e.g. PISA 2003).....	117
<hr/>		
Box 8.1	SAS® syntax for computing the mean on the science scale by using the PROC_MEANS_NO_PV macro (e.g. PISA 2006).....	121
Box 8.2	SAS® syntax for computing the mean and its standard error on PVs (e.g. PISA 2006).....	122
Box 8.3	SAS® syntax for computing the standard deviation and its standard error on PVs by gender (e.g. PISA 2006).....	123
Box 8.4	SAS® syntax for computing regression coefficients and their standard errors on PVs by using the PROC_REG_NO_PV macro (e.g. PISA 2006).....	124
Box 8.5	SAS® syntax for running the simple linear regression macro with PVs (e.g. PISA 2006).....	125
Box 8.6	SAS® syntax for running the correlation macro with PVs (e.g. PISA 2006).....	126
Box 8.7	SAS® syntax for the computation of the correlation between mathematics/quantity and mathematics/space and shape by using the PROC_CORR_NO_PV macro (e.g. PISA 2003).....	129
<hr/>		
Box 9.1	SAS® syntax for generating the proficiency levels in science (e.g. PISA 2006).....	137
Box 9.2	SAS® syntax for computing the percentages of students by proficiency level in science and its standard errors by using the PROC_FREQ_NO_PV macro (e.g. PISA 2006).....	138
Box 9.3	SAS® syntax for computing the percentage of students by proficiency level in science and its standard errors by using the PROC_FREQ_PV macro (e.g. PISA 2006).....	140
Box 9.4	SAS® syntax for computing the percentage of students by proficiency level and its standard errors by gender (e.g. PISA 2006).....	140
Box 9.5	SAS® syntax for generating the proficiency levels in mathematics (e.g. PISA 2003).....	141
Box 9.6	SAS® syntax for computing the mean of self-efficacy in mathematics and its standard errors by proficiency level (e.g. PISA 2003).....	142
<hr/>		
Box 10.1	SAS® syntax for merging the student and school data files (e.g. PISA 2006).....	148
Box 10.2	Question on school location in PISA 2006.....	149
Box 10.3	SAS® syntax for computing the percentage of students and the average performance in science, by school location (e.g. PISA 2006).....	149
<hr/>		
Box 11.1	SAS® syntax for computing the mean of job expectations by gender (e.g. PISA 2003).....	154
Box 11.2	SAS® macro for computing standard errors on differences (e.g. PISA 2003).....	157



Box 11.3	Alternative SAS® macro for computing the standard error on a difference for a dichotomous variable (e.g. PISA 2003).....	158
Box 11.4	SAS® syntax for computing standard errors on differences which involve PVs (e.g. PISA 2003).....	160
Box 11.5	SAS® syntax for computing standard errors on differences that involve PVs (e.g. PISA 2006).....	162
<hr/>		
Box 12.1	SAS® syntax for computing the pooled OECD total for the mathematics performance by gender (e.g. PISA 2003).....	170
Box 12.2	SAS® syntax for the pooled OECD average for the mathematics performance by gender (e.g. PISA 2003).....	171
Box 12.3	SAS® syntax for the creation of a larger dataset that will allow the computation of the pooled OECD total and the pooled OECD average in one run (e.g. PISA 2003).....	172
<hr/>		
Box 14.1	SAS® syntax for the quarter analysis (e.g. PISA 2006).....	189
Box 14.2	SAS® syntax for computing the relative risk with five antecedent variables and five outcome variables (e.g. PISA 2006).....	193
Box 14.3	SAS® syntax for computing the relative risk with one antecedent variable and one outcome variable (e.g. PISA 2006).....	194
Box 14.4	SAS® syntax for computing the relative risk with one antecedent variable and five outcome variables (e.g. PISA 2006).....	194
Box 14.5	SAS® syntax for computing effect size (e.g. PISA 2006).....	196
Box 14.6	SAS® syntax for residual analyses (e.g. PISA 2003).....	200
<hr/>		
Box 15.1	Normalisation of the final student weights (e.g. PISA 2006).....	207
Box 15.2	SAS® syntax for the decomposition of the variance in student performance in science (e.g. PISA 2006).....	208
Box 15.3	SAS® syntax for normalising PISA 2006 final student weights with deletion of cases with missing values and syntax for variance decomposition (e.g. PISA 2006).....	211
Box 15.4	SAS® syntax for a multilevel regression model with random intercepts and fixed slopes (e.g. PISA 2006).....	214
Box 15.5	SAS® output for the multilevel model in Box 15.4.....	214
Box 15.6	SAS® syntax for a multilevel regression model (e.g. PISA 2006).....	216
Box 15.7	SAS® output for the multilevel model in Box 15.6.....	217
Box 15.8	SAS® output for the multilevel model with covariance between random parameters.....	218
Box 15.9	Interpretation of the within-school regression coefficient.....	220
Box 15.10	SAS® syntax for a multilevel regression model with a school-level variable (e.g. PISA 2006).....	221
Box 15.11	SAS® syntax for a multilevel regression model with interaction (e.g. PISA 2006).....	222
Box 15.12	SAS® output for the multilevel model in Box 15.11.....	222
Box 15.13	SAS® syntax for using the multilevel regression macro (e.g. PISA 2006).....	224
Box 15.14	SAS® syntax for normalising the weights for a three-level model (e.g. PISA 2006).....	226
<hr/>		
Box 16.1	SAS® syntax for testing the gender difference in standard deviations of reading performance (e.g. PISA 2000).....	233
Box 16.2	SAS® syntax for testing the gender difference in the 5th percentile of the reading performance (e.g. PISA 2006).....	235
Box 16.3	SAS® syntax for preparing a data file for the multilevel analysis.....	238



Box 16.4	SAS® syntax for running a preliminary multilevel analysis with one PV	239
Box 16.5	SAS® output for fixed parameters in the multilevel model.....	239
Box 16.6	SAS® syntax for running multilevel models with the PROC_MIXED_PV macro	242
<hr/>		
Box 17.1	SAS® macro of PROC_MEANS_NO_PV.sas.....	250
Box 17.2	SAS® macro of PROC_MEANS_PV.sas.....	253
Box 17.3	SAS® macro of PROC_FREQ_NO_PV.sas.....	256
Box 17.4	SAS® macro of PROC_FREQ_PV.sas.....	259
Box 17.5	SAS® macro of PROC_REG_NO_PV.sas.....	263
Box 17.6	SAS® macro of PROC_REG_PV.sas.....	266
Box 17.7	SAS® macro of PROC_CORR_NO_PV.sas.....	270
Box 17.8	SAS® macro of PROC_CORR_PV.sas.....	273
Box 17.9	SAS® macro of PROC_DIF_NO_PV.sas	276
Box 17.10	SAS® macro of PROC_DIF_PV.sas	279
Box 17.11	SAS® macro of QUARTILE_PV.sas	282
Box 17.12	SAS® macro of RELATIVE_RISK_NO_PV.sas.....	288
Box 17.13	SAS® macro of RELATIVE_RISK_PV.sas.....	291
Box 17.14	SAS® macro of EFFECT_SIZE_NO_PV.sas	296
Box 17.15	SAS® macro of EFFECT_SIZE_PV.sas.....	298
Box 17.16	SAS® macro of PROC_MIXED_NO_PV.sas.....	301
Box 17.17	SAS® macro of PROC_MIXED_PV.sas	306
<hr/>		
Box A1.1	Descriptive statistics of background and explanatory variables.....	318
Box A1.2	Background model for student performance.....	319
Box A1.3	Final net combined model for student performance.....	320
Box A1.4	Background model for the impact of socio-economic background.....	321
Box A1.5	Model of the impact of socio-economic background: “school resources” module	322
Box A1.6	Model of the impact of socio-economic background: “accountability practices” module	323
Box A1.7	Final combined model for the impact of socio-economic background.....	323

LIST OF FIGURES

Figure 1.1	Relationship between social and academic segregations.....	27
Figure 1.2	Relationship between social segregation and the correlation between science performance and student HISEI	27
Figure 1.3	Conceptual grid of variable types.....	29
Figure 1.4	Two-dimensional matrix with examples of variables collected or available from other sources	30
<hr/>		
Figure 2.1	Science mean performance in OECD countries (PISA 2006).....	38
Figure 2.2	Gender differences in reading in OECD countries (PISA 2000).....	38
Figure 2.3	Regression coefficient of ESCS on mathematic performance in OECD countries (PISA 2003).....	39
Figure 2.4	Design effect on the country mean estimates for science performance and for ESCS in OECD countries (PISA 2006)	42
Figure 2.5	Simple random sample and unbiased standard errors of ESCS on science performance in OECD countries (PISA 2006)	43



Figure 4.1	Distribution of the results of 36 students.....	60
Figure 4.2	Sampling variance distribution of the mean.....	62
Figure 5.1	Probability of success for two high jumpers by height (dichotomous).....	82
Figure 5.2	Probability of success for two high jumpers by height (continuous).....	83
Figure 5.3	Probability of success to an item of difficulty zero as a function of student ability.....	83
Figure 5.4	Student score and item difficulty distributions on a Rasch continuum.....	86
Figure 5.5	Response pattern probabilities for the response pattern (1, 1, 0, 0).....	88
Figure 5.6	Response pattern probabilities for a raw score of 1.....	89
Figure 5.7	Response pattern probabilities for a raw score of 2.....	90
Figure 5.8	Response pattern probabilities for a raw score of 3.....	90
Figure 5.9	Response pattern likelihood for an easy test and a difficult test.....	91
Figure 5.10	Rasch item anchoring.....	92
Figure 6.1	Living room length expressed in integers.....	96
Figure 6.2	Real length per reported length.....	97
Figure 6.3	A posterior distribution on a test of six items.....	98
Figure 6.4	EAP estimators.....	99
Figure 8.1	A two-dimensional distribution.....	127
Figure 8.2	Axes for two-dimensional normal distributions.....	127
Figure 13.1	Trend indicators in PISA 2000, PISA 2003 and PISA 2006.....	179
Figure 14.1	Percentage of schools by three school groups (PISA 2003).....	198
Figure 15.1	Simple linear regression analysis versus multilevel regression analysis.....	205
Figure 15.2	Graphical representation of the between-school variance reduction.....	215
Figure 15.3	A random multilevel model.....	216
Figure 15.4	Change in the between-school residual variance for a fixed and a random model.....	218
Figure 16.1	Relationship between the segregation index of students' expected occupational status and the segregation index of student performance in reading (PISA 2000).....	244
Figure 16.2	Relationship between the segregation index of students' expected occupational status and the correlation between HISEI and students' expected occupational status.....	245

LIST OF TABLES

Table 1.1	Participating countries/economies in PISA 2000, PISA 2003, PISA 2006 and PISA 2009.....	21
Table 1.2	Assessment domains covered by PISA 2000, PISA 2003 and PISA 2006.....	22
Table 1.3	Correlation between social inequities and segregations at schools for OECD countries.....	28
Table 1.4	Distribution of students per grade and per ISCED level in OECD countries (PISA 2006).....	31
Table 2.1	Design effect and type I errors.....	41
Table 2.2	Mean estimates and standard errors.....	45



Table 2.3	Standard deviation estimates and standard errors.....	45
Table 2.4	Correlation estimates and standard errors.....	45
Table 2.5	ESCS regression coefficient estimates and standard errors.....	46
<hr/>		
Table 3.1	Height and weight of ten persons	52
Table 3.2	Weighted and unweighted standard deviation estimate	52
Table 3.3	School, within-school, and final probability of selection and corresponding weights for a two-stage, simple random sample with the first-stage units being schools of equal size.....	54
Table 3.4	School, within-school, and final probability of selection and corresponding weights for a two-stage, simple random sample with the first-stage units being schools of unequal size	54
Table 3.5	School, within-school, and final probability of selection and corresponding weights for a simple and random sample of schools of unequal size (smaller schools)	55
Table 3.6	School, within-school, and final probability of selection and corresponding weights for a simple and random sample of schools of unequal size (larger schools)	55
Table 3.7	School, within-school, and final probability of selection and corresponding weights for PPS sample of schools of unequal size	56
Table 3.8	Selection of schools according to a PPS and systematic procedure.....	57
<hr/>		
Table 4.1	Description of the 630 possible samples of 2 students selected from 36 students, according to their mean.....	61
Table 4.2	Distribution of all possible samples with a mean between 8.32 and 11.68.....	63
Table 4.3	Distribution of the mean of all possible samples of 4 students out of a population of 36 students.....	64
Table 4.4	Between-school and within-school variances on the mathematics scale in PISA 2003.....	67
Table 4.5	Current status of sampling errors.....	67
Table 4.6	Between-school and within-school variances, number of participating schools and students in Denmark and Germany in PISA 2003	68
Table 4.7	The Jackknives replicates and sample means.....	70
Table 4.8	Values on variables X and Y for a sample of ten students.....	71
Table 4.9	Regression coefficients for each replicate sample.....	71
Table 4.10	The Jackknife replicates for unstratified two-stage sample designs.....	72
Table 4.11	The Jackknife replicates for stratified two-stage sample designs.....	73
Table 4.12	Replicates with the Balanced Repeated Replication method.....	74
Table 4.13	The Fay replicates	75
<hr/>		
Table 5.1	Probability of success when student ability equals item difficulty.....	84
Table 5.2	Probability of success when student ability is less than the item difficulty by 1 unit.....	84
Table 5.3	Probability of success when student ability is greater than the item difficulty by 1 unit	84
Table 5.4	Probability of success when student ability is less than the item difficulty by 2 units	85
Table 5.5	Probability of success when student ability is greater than the item difficulty by 2 units.....	85
Table 5.6	Possible response pattern for a test of four items.....	87
Table 5.7	Probability for the response pattern (1, 1, 0, 0) for three student abilities.....	87
Table 5.8	Probability for the response pattern (1, 0) for two students of different ability in an incomplete test design.....	91
Table 5.9	PISA 2003 test design	93

Table 6.1	Structure of the simulated data.....	100
Table 6.2	Means and variances for the latent variables and the different student ability estimators.....	100
Table 6.3	Percentiles for the latent variables and the different student ability estimators.....	101
Table 6.4	Correlation between HISEI, gender and the latent variable, the different student ability estimators.....	101
Table 6.5	Between- and within-school variances.....	102
<hr/>		
Table 7.1	HISEI mean estimates.....	107
Table 7.2	Squared differences between replicate estimates and the final estimate.....	108
Table 7.3	Output data file exercise1 from Box 7.2.....	111
Table 7.4	Available statistics with the PROC_MEANS_NO_PV macro.....	111
Table 7.5	Output data file exercise2 from Box 7.3.....	112
Table 7.6	Output data file exercise3 from Box 7.4.....	112
Table 7.7	Percentage of girls for the final and replicate weights and squared differences.....	113
Table 7.8	Output data file exercise4 from Box 7.5.....	114
Table 7.9	Output data file exercise5 from Box 7.6.....	115
Table 7.10	Output data file exercise6 from Box 7.7.....	116
Table 7.11	Output data file exercise6_criteria from Box 7.7.....	117
Table 7.12	Output data file exercise7 from Box 7.8.....	117
<hr/>		
Table 8.1	The 405 mean estimates.....	120
Table 8.2	Mean estimates and their respective sampling variances on the science scale for Belgium (PISA 2006).....	121
Table 8.3	Output data file exercise6 from Box 8.2.....	123
Table 8.4	Output data file exercise7 from Box 8.3.....	123
Table 8.5	The 450 regression coefficient estimates.....	125
Table 8.6	HISEI regression coefficient estimates and their respective sampling variance on the science scale in Belgium after accounting for gender (PISA 2006).....	125
Table 8.7	Output data file exercise8 from Box 8.5.....	125
Table 8.8	Output data file exercise9 from Box 8.6.....	126
Table 8.9	Correlation between the five plausible values for each domain, mathematics/quantity and mathematics/space and shape.....	128
Table 8.10	The five correlation estimates between mathematics/quantity and mathematics/space and shape and their respective sampling variance.....	129
Table 8.11	Standard deviations for mathematics scale using the correct method (plausible values) and by averaging the plausible values at the student level (pseudo-EAP) (PISA 2003).....	131
Table 8.12	Unbiased shortcut for a population estimate and its standard error.....	132
Table 8.13	Standard errors from the full and shortcut computation (PISA 2006).....	132
<hr/>		
Table 9.1	The 405 percentage estimates for a particular proficiency level.....	138
Table 9.2	Estimates and sampling variances per proficiency level in science for Germany (PISA 2006).....	139
Table 9.3	Final estimates of the percentage of students, per proficiency level, in science and its standard errors for Germany (PISA 2006).....	139
Table 9.4	Output data file exercise6 from Box 9.3.....	140
Table 9.5	Output data file exercise7 from Box 9.4.....	140
Table 9.6	Mean estimates and standard errors for self-efficacy in mathematics per proficiency level (PISA 2003).....	143
Table 9.7	Output data file exercise8 from Box 9.6.....	143



Table 10.1	Percentage of students per grade and ISCED level, by country (PISA 2006)	146
Table 10.2	Output data file exercise1 from Box 10.3	150
Table 10.3	Output data file exercise2 from Box 10.3	150
Table 11.1	Output data file exercise1 from Box 11.1	155
Table 11.2	Mean estimates for the final and 80 replicate weights by gender (PISA 2003)	155
Table 11.3	Difference in estimates for the final weight and 80 replicate weights between females and males (PISA 2003)	157
Table 11.4	Output data file exercise2 from Box 11.2	158
Table 11.5	Output data file exercise3 from Box 11.3	159
Table 11.6	Gender difference estimates and their respective sampling variances on the mathematics scale (PISA 2003)	159
Table 11.7	Output data file exercise4 from Box 11.4	160
Table 11.8	Gender differences on the mathematics scale, unbiased standard errors and biased standard errors (PISA 2003)	161
Table 11.9	Gender differences in mean science performance and in standard deviation for science performance (PISA 2006)	161
Table 11.10	Regression coefficient of HISEI on the science performance for different models (PISA 2006)	163
Table 11.11	Cross tabulation of the different probabilities	163
Table 12.1	Regression coefficients of the index of instrumental motivation in mathematics on mathematic performance in OECD countries (PISA 2003)	169
Table 12.2	Output data file exercise1 from Box 12.1	170
Table 12.3	Output data file exercise2 from Box 12.2	171
Table 12.4	Difference between the country mean scores in mathematics and the OECD total and average (PISA 2003)	174
Table 13.1	Trend indicators between PISA 2000 and PISA 2003 for HISEI, by country	180
Table 13.2	Linking error estimates	182
Table 13.3	Mean performance in reading by gender in Germany	184
Table 14.1	Distribution of the questionnaire index of cultural possession at home in Luxembourg (PISA 2006)	188
Table 14.2	Output data file exercise1 from Box 14.1	190
Table 14.3	Labels used in a two-way table	190
Table 14.4	Distribution of 100 students by parents' marital status and grade repetition	191
Table 14.5	Probabilities by parents' marital status and grade repetition	191
Table 14.6	Relative risk for different cutpoints	191
Table 14.7	Output data file exercise2 from Box 14.2	193
Table 14.8	Mean and standard deviation for the student performance in reading by gender, gender difference and effect size (PISA 2006)	195
Table 14.9	Output data file exercise4 from Box 14.5	197
Table 14.10	Output data file exercise5 from Box 14.5	197
Table 14.11	Mean of the residuals in mathematics performance for the bottom and top quarters of the PISA index of economic, social and cultural status, by school group (PISA 2003)	199

Table 15.1	Between- and within-school variance estimates and intraclass correlation (PISA 2006).....	209
Table 15.2	Output data file “ranparm1” from Box 15.3.....	212
Table 15.3	Output data file “fixparm3” from Box 15.6.....	217
Table 15.4	Output data file “ranparm3” from Box 15.6.....	217
Table 15.5	Variance/covariance estimates before and after centering.....	219
Table 15.6	Output data file of the fixed parameters file.....	221
Table 15.7	Average performance and percentage of students by student immigrant status and by type of school.....	223
Table 15.8	Variables for the four groups of students.....	223
Table 15.9	Comparison of the regression coefficient estimates and their standard errors in Belgium (PISA 2006).....	224
Table 15.10	Comparison of the variance estimates and their respective standard errors in Belgium (PISA 2006).....	225
Table 15.11	Three-level regression analyses.....	226
Table 16.1	Differences between males and females in the standard deviation of student performance (PISA 2000).....	234
Table 16.2	Distribution of the gender differences (males – females) in the standard deviation of the student performance.....	234
Table 16.3	Gender difference on the PISA combined reading scale for the 5 th , 10 th , 90 th and 95 th percentiles (PISA 2000).....	235
Table 16.4	Gender difference in the standard deviation for the two different item format scales in reading (PISA 2000).....	236
Table 16.5	Random and fixed parameters in the multilevel model with student and school socio-economic background.....	237
Table 16.6	Random and fixed parameters in the multilevel model with socio-economic background and grade retention at the student and school levels.....	241
Table 16.7	Segregation indices and correlation coefficients by country (PISA 2000).....	243
Table 16.8	Segregation indices and correlation coefficients by country (PISA 2006).....	244
Table 16.9	Country correlations (PISA 2000).....	245
Table 16.10	Country correlations (PISA 2006).....	246
Table 17.1	Synthesis of the 17 SAS® macros.....	249
Table A2.1	Cluster rotation design used to form test booklets for PISA 2006.....	324
Table A12.1	Mapping of ISCED to accumulated years of education.....	449
Table A12.2	ISCO major group white-collar/blue-collar classification.....	451
Table A12.3	ISCO occupation categories classified as science-related occupations.....	451
Table A12.4	Household possessions and home background indices.....	455
Table A12.5	Factor loadings and internal consistency of ESCS 2006 in OECD countries.....	465
Table A12.6	Factor loadings and internal consistency of ESCS 2006 in partner countries/economies.....	466



User's Guide

Preparation of data files

All data files (in text format) and the SAS® control files are available on the PISA website (www.pisa.oecd.org).

SAS® users

By running the SAS® control files, the PISA data files are created in the SAS® format. Before starting analysis, assigning the folder in which the data files are saved as a SAS® library.

For example, if the PISA 2000 data files are saved in the folder of "c:\pisa2000\data\", the PISA 2003 data files are in "c:\pisa2003\data\", and the PISA 2006 data files are in "c:\pisa2006\data\", the following commands need to be run to create SAS® libraries:

```
libname PISA2000 "c:\pisa2000\data\" ;  
libname PISA2003 "c:\pisa2003\data\" ;  
libname PISA2006 "c:\pisa2006\data\" ;  
run;
```

SAS® syntax and macros

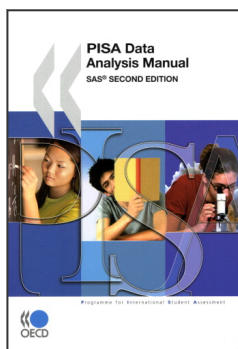
All syntaxes and macros in this manual can be copied from the PISA website (www.pisa.oecd.org). The 17 SAS® macros presented in Chapter 17 need to be saved under "c:\pisa\macro\", before starting analysis. Each chapter of the manual contains a complete set of syntaxes, which must be done sequentially, for all of them to run correctly, within the chapter.

Rounding of figures

In the tables and formulas, figures were rounded to a convenient number of decimal places, although calculations were always made with the full number of decimal places.

Country abbreviations used in this manual

AUS	Australia	FRA	France	MEX	Mexico
AUT	Austria	GBR	United Kingdom	NLD	Netherlands
BEL	Belgium	GRC	Greece	NOR	Norway
CAN	Canada	HUN	Hungary	NZL	New Zealand
CHE	Switzerland	IRL	Ireland	POL	Poland
CZE	Czech Republic	ISL	Iceland	PRT	Portugal
DEU	Germany	ITA	Italy	SVK	Slovak Republic
DNK	Denmark	JPN	Japan	SWE	Sweden
ESP	Spain	KOR	Korea	TUR	Turkey
FIN	Finland	LUX	Luxembourg	USA	United States



From:
PISA Data Analysis Manual: SAS, Second Edition

Access the complete publication at:
<https://doi.org/10.1787/9789264056251-en>

Please cite this chapter as:

OECD (2009), "SAS® Macro", in *PISA Data Analysis Manual: SAS, Second Edition*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/9789264056251-18-en>

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.