Estimation of the Variance Using Bootstrap Weights
User's Guide for the BOOTVARE_V30.SPS Program
(VERSION 3.0)

1. Introduction

This guide is for users of the SPSS program BOOTVARE_V30.SPS which was created to estimate the variance using the bootstrap method.

Section 2 of this guide briefly explains the bootstrap resampling method used to estimate the variance. Section 3 gives detailed instructions for using the BOOTVARE_V30.SPS program, as well as a description of the preliminary steps that are required. The programs are given in Appendix A. Appendix B contains a complete example (programs and results). Finally, the survey-specific parameters required for executing the programs (file names, identification variables, etc.) are provided in the document "AppendixC_*XYZ*", where *XYZ* identifies the survey.

Changes from the Previous Version:

The biggest change made to the program is that a single version of the program supports all the Statistics Canada surveys that use the BOOTVAR program. Users only need to specify a few parameters (see the document AppendixC_*XYZ*) in section 1 of the program.

Please note that the program was tested and works with SPSS versions 10.0 and 11.0.

2. Bootstrap Method

The sampling designs for Statistics Canada's surveys are generally complex. Since the variance for such designs cannot be estimated with simple formulas, resampling methods are often used to estimate the variance.

The bootstrap method consists of subsampling the initial sample. Within each stratum, a simple random sample (SRS) is selected, with replacement, from n-1 clusters within the n clusters of the stratum. The process is repeated 'B' times, creating B new samples (or replicates). Weights are recalculated for each of the B samples – the B weights are called the bootstrap weights. The bootstrap weights are used to calculate B estimates which are then used to estimate the variance.

The bootstrap weights have already been generated, and are available with the data. The BOOTVARE_V30.SPS program uses these bootstrap weights to estimate the variance for simple statistics such as totals and ratios, as well as for more complex analyses like regressions. These estimates of the variance should be used to derive quality indicators and to apply the survey's rules for releasing the estimates.

Here are the main steps for estimating the variance of a particular estimate using the bootstrap method:

A) Calculate an estimate (total, ratio, etc.) using the final weight included in the data file. This estimate is the point estimate.

B) Calculate the same estimate, this time using each of the B bootstrap weights contained in the bootstrap file. B estimates (total, ratio, etc) are then obtained.

C) Finally, calculate the variance of the B estimates. This variance is the estimate of the variance of the point estimate calculated in A.

3. Variance Estimation with the BOOTVARE_V30.SPS Program

The BOOTVARE_V30.SPS program calculates estimates of the variance of totals, ratios, differences between ratios, and linear or logistic regression parameters.

Variance estimation is performed in *two steps* and involves the use of three SPSS programs. The *first step* consists of creating a data file containing the variables required for the analysis (first program). The *second step* involves using BOOTVARE_V30.SPS (and MACROE_V30.SPS) to estimate the variances.

Step 1: Creation of the Analysis File

The user needs to create a SPSS data file which will be used as the input file for the program estimating the variance in step 2. The following tasks must be done in this step:

1. Reading of the input file
2. Creation of the variables required for the analysis

1 - <u>Reading of the input file</u>: The analysis file is created from the survey data file. The file layout must be provided in order to read in the variables contained in the file. See AppendixC_*XYZ* for the file and variable names.

2 - <u>Creation of the variables required for the analysis</u>: Variables derived from the input variables should be created in this step. It may be necessary to create dichotomous variables (1 or 0) which identify records that have a characteristic of interest – such variables will take a value of 1 for records that have the characteristic and a value of 0 otherwise. The total of a dichotomous variable will sum the weights of the records with the characteristic. See the example in Appendix B for more details.

The analysis file must contain:

- The necessary variables for the analysis (derived variables including dichotomous variables, and input variables that do not need to be modified). To reduce the runtime of the program, DO NOT keep unnecessary variables.
- The identification variable(s) of the respondents.
- If needed, the breakdown variable(s), identifying the groups for which a separate analysis is desired (ex.: province, sex, etc...).

- If the analysis is only of interest for a certain subgroup (for example, a province or an age group), keep only the records that are part of this subgroup.

REMARKS:

- It is recommended that point estimates be calculated at this step to be sure that the desired estimate is being calculated correctly, and that the program BOOTVARE_V30.SPS correctly calculates the same estimate. In this case, it is necessary to keep the weight variable when creating the analysis file.

- Means are estimated using the ratio macro in Step 2. Dichotomous variables identifying the records that are part of the group of interest must be created for the denominator.

The users must create their own program to prepare the analysis file containing the necessary variables for the analysis. An example of a program that creates this file is included in Appendix A (the program STEP1.SPS).

Step 2: Variance Calculation Using the BOOTVARE_V30.SPS Program

Once the new SPSS data file is created in step 1, the next step consists of running the BOOTVARE_V30.SPS program. Before running it, the desired parameters and analyses must be specified. This program calls the MACROE_V30.SPS program.

MACROE_V30.SPS contains the program code of the various macros. *For standard use of the variance estimation program, no modification of the MACROE_V30.SPS program by the user is necessary.* Changes may be required in certain cases, as explained later.

The BOOTVARE_V30.SPS program is included in Appendix A. The parts that are to be changed by the user are given in ***bold type.*** The rest of the program does not need to be changed. The program is divided into ***two sections***. The ***first section*** is for specifying the required parameters, and the ***second section*** is for listing the desired analyses.

Section 1:

In this section, the user must specify:

- The name of the directory of the analysis file created in step 1 and of the output file containing the results
- The name of the data file (analysis file created in step 1)
- The name and directory of the bootstrap weights file
- The breakdown variable(s) to specify that the analysis is to be performed separately for specific sub-groups (ex.: provinces, sex)
- The identification variable(s) of the respondents, the weight variables and the number of bootstrap weights
- The name of the directory where the program MACROE_V30.SPS is located

N.B. AppendixC_*XYZ* contains survey-specific information (file names, certain variable names, number of bootstrap weights).

Section 2:

In this section, the user lists the analyses for which estimates of the variance are desired. The following types of analyses are supported:

- Totals
- Ratios (including means)
- Differences between ratios
- Regression models (linear or logistic)

For means: To estimate the variance of a mean, the macro for ratios can be used. The numerator is the variable of interest and the denominator is a dichotomious variable that identifies the population of interest.

For differences between ratios: To calculate the difference between ratios, it may be necessary for the user to modify the macro *diff_rat* in the MACROE_V30.SPS program to suit their needs. See the notes included in the BOOTVARE_V30.SPS program for more details.

For regressions: Categorical variables will be treated as continuous variables. Dichotomous variables must be created in step 1 for each possible value (except one) of the categorical variable in order to treat this variable properly.

Modification to the program for testing purposes: Running the program could take long, especially for complex model analyses. It is possible to reduce the number of bootstrap weights used in order to test the program. **(However, to obtain the final estimates of the variance, it is important to use all of the bootstrap weights provided.)** To test the program, all that is required is to modify the parameter that specifies the number of bootstrap weights in the first part of the BOOTVARE_V30.SPS program.

Results Obtained with BOOTVARE_V30.SPS

The following results are obtained after running BOOTVARE_V30.SPS for totals, ratios and differences between ratios. See Appendix B for interpretation of the results:

TYPE:	Type of estimate (total, ratio,diff_ratio)
VAR1 & VAR2:	Variables used to calculate the estimates. For a total, VAR2 = None.
VAR3 & VAR4:	Variables used to calculate the difference between ratios.
YHAT:	Parameter estimate
BS_SD:	Standard deviation
BS_CV:	Coefficient of variation
CIL95:	Lower limit of the 95% confidence interval
CIU95:	Upper limit of the 95% confidence interval

The following results are obtained after running BOOTVARE_V30.SPS for linear and logistic regressions.   See Appendix B for interpretation of the results:

PARAM:	Parameter to estimate
BETA:	Parameter estimate
ODDS:	Odds ratio (logistic regression only)
WALD:	Wald statistic (logistic regression only)
PVALUE:	P-value of the Wald statistic (logistic regression only)
BSVAR:	Variance of the parameter estimate
BS_SD:	Standard deviation of the parameter
BS_CV:	Coefficient of variation for the parameter estimate
CIL95:	Lower limit of the 95% confidence interval
CIU95:	Upper limit of the 95% confidence interval

Appendix A contains the BOOTVARE_V30.SPS program, preceded by an example of a program which prepares the analysis file (STEP1.SPS).   Appendix B contains a complete example (programs and results).   Finally, AppendixC_*XYZ* contains survey-specific information (file names, names of certain variables, number of bootstrap weights).

APPENDIX A - Programs to Run

STEP 1.SPS Program
(Used as an example; the users can use their own program or use the Data Editor window)

The parts in ***bold*** need to be changed.

```
***********************************************************************************
*                          STEP1.SPS                                             *
*                                                                                *
*  This program creates the SPSS datafile containing the necessary variables for the  *
*  BOOTVARE_V30.SPS program.                                                     *
*                                                                                *
*  IT IS NOT NECESSARY TO USE THIS PROGRAM.  Dichotomous variables can           *
*  be computed directly in the Data Editor window and the analysis data file can be  *
*  created from there.                                                           *
***********************************************************************************.

SET COM NO HEA ON MES ON ERR ON RES ON JOU ON LEN NONE WID 132 PRI NO MPR NO.


* Creation of the SPSS "sav" file containing the variables and cases required for the analysis. Note
* that this file should be as small as possible (containing only necessary variables and cases) in
* order to reduce time and memory requirements especially if regression type analysis are to be done.



DEFINE !Go()
!LET !Layout = ' name_and_location_of_layout '
!LET !DataIn = ' name_and_location_of_source_file '
!LET !DataOut = ' name_of_resulting_file_in_SPSS_format '              /* <- (analysis file to be used in BOOTVARE_V30.SPS)


* Do not modify the following block of instructions.
FILE HANDLE INFILE/NAME = !QUOTE(!DataIn).
INCLUDE FILE=!QUOTE(!Layout).
SET PRI ON.


***********************************************************************************************
* Creation of Dichotomous Variables ( -> REPLACE all following instructions (up to "SAVE"  )    *
* (examples are presented below using NPHS cycle 3 variables)                                  *
***********************************************************************************************.

NUMERIC diab males females mdiab fdiab.

IF (ccc8_1j=1)   diab=1.
IF (ccc8_1j~=1) diab=0.

IF (sex=1) males=1.
IF (sex~=1) males=0.
IF (sex=2) females=1.
IF (sex~=2) females=0.

compute mdiab = males*diab.
compute fdiab  =  females*diab.

execute.


***********************************************************************************************.

SAVE OUTFILE !QUOTE(!DataOut)
          /KEEP = list of variables to keep .

          * It is recommended that only the necessary variables be kept    *
          * in order to reduce the runtime of BOOTVARE_V30.SPS.           *
          * IMPORTANT:  the identification variables and, if necessary,    *
          * the breakdown variable(s) (ex: province, sex) must be kept. The *
          * weight variable also must be kept if point estimates are       *
          * calculated at this step                                       *.


* Replace "SAVE" by "XSAVE" in previous command if point estimates are desired at this time.

***   DO NOT FORGET TO RUN THE LAST LINES OF THE PROGRAM (AFTER THE EXAMPLE)  ***.


********************************************
* Calculation of point estimates     *
* Suggested, but not required...      *
********************************************.
```

*** Examples: ***.

WEIGHT BY **weight_variable.**

FREQ VAR = **variable list.**
LOGISTIC REG VAR = **var** WITH **variable list** /PRI SUM
        /CRI BCO(0.0001) LCO(0) ITE(25) EPS(10E-12).

WEIGHT OFF.

* DO NOT DELETE the following instructions
***.

SET PRI NO.
!ENDDEFINE /* !Go.

SET MPR ON.
!Go.
SET MPR NO.

BOOTVARE_V30.SPS Program

The parts in *bold* need to be changed.

. . .
. . .
. . .

```
********************************************************************
***                       SECTION 1                    ***
********************************************************************
***                                                    ***
***   This section lets the user specify the different parameters of   ***
***   interest (variable names, directory names, file names, etc.)     ***
***                                                    ***
********************************************************************.

DEFINE !Bootvar()

PRESERVE.
SET COM NO HEA ON MES ON ERR ON RES ON JOU ON LEN NONE WID 132 PRI NO MPR NO MITERATE=3000.

!Let !Tot=!Null
!Let !Reg=!Null


*********************************************************************************************
* SPECIFY THE ANALYSIS DATA FILE TO BE USED (created in step 1) (Folder and file) :    *
*********************************************************************************************.

!LET !Mfile='Folder_and_file_name_of_the_analysis_file_(step 1)'                    /*  <- (ex: c:\data\analysis.sav).



***********************************************************************
*   SPECIFY THE BOOTSTRAP WEIGHTS FOLDER AND FILE NAMES :        *
*         NB:  Only run one of the two following series of commands      *
*         (comment the other one out, or erase it):                 *
***********************************************************************.

* EXECUTE THIS PART IF THE BOOTSTRAP WEIGHTS ARE IN SPSS FORMAT (remove the " * ") *
***********************************************************************************************.

*!LET !BWsav=' Folder_and_file_names_of_the_bootstrap_weights_in_SPSS_format_(with_extension) '


* EXECUTE THIS PART IF THE BOOTSTRAP WEIGHTS ARE IN ASCII (.TXT) FORMAT (remove the " * ") *
***********************************************************************************************.

*!LET !BWTXT=' Folder_and_file_names_of_the_bootstrap_weights_in_text_format_(with_extension) '
*!LET !Bwlay=' Folder_and_file_names_of_the_bootstrap_layout_file_in_SPSS_syntax_(with_extension) '
*!LET !Bwsav='
Folder_and_file_names_of_the_bootstrap_weights_in_SPSS_format_(created_in_this_section)_with_extension '

*FILE HANDLE BOOTLAY/NAME = !QUOTE(!BWTXT).
*INCLUDE FILE=!QUOTE(!BWlay).
*SAVE OUT=!QUOTE(!BWsav) /UNC.


*****************************************************************************************************
* SPECIFY, IF DESIRED, THE BREAKDOWN VARIABLE(S) (EG: PROVINCE, SEX, ETC...):              *
*      Write the name of the breakdown variable(s) below.                                 *
*                                                                                         *
*      - If the analysis includes all of the data in the file created in step 1, put two single quotes (%LET !Classes = ' ' )   *
*      - If more than one variable, leave a space between each variable  ( %LET !Classes = ' var1 var2 ' )    *
*      - DO NOT ERASE OR COMMENT OUT THIS COMMAND                                          *
*****************************************************************************************************.

!LET !Classes = ' breakdown_variable(s)_or_empty_(keep_the_2_single_quotes) '
```

```
***********************************************************************************************************
* SPECIFY THE FOLLOWING INFORMATION (SPECIFIC TO THE SURVEY YOU ARE USING):          *
*    You must specify:                                                               *
*        1- The unique identifier variable(s) (within quotes, separated by a space)  *
*        2- The Final Weight (variable included in the bootstrap weight file)        *
*        3- The prefix of the boostraps weight variables                             *
*        4- The number of bootstrap weights to use (note: For testing, B must be >= 2. *
*           IT IS NECESSARY TO USE ALL THE BOOTSTRAP WEIGHTS WHEN PERFORMING THE FINAL *
*           ANALYSIS.  THE COMPLETE BOOTVARE_V30.SPS PROGRAM MUST THEN BE RUN.)       *
*                                                                                    *
*        - Refer to AppendixC_XYZ  to obtain this information                         *
***********************************************************************************************************.

!LET !ident  = ' unique identifier variable(s) '
!LET !fwgt   = final_weight
!LET !prefix = prefix_of_bootstrap_weight_variables
!LET !B      = number_of_weights_to_use


***********************************************************************************************************
* SPECIFY THE DIRECTORY AND THE NAME OF THE FILE THAT CONTAINS THE MACROS            *
* (THE PROGRAM MACROE_V30.SPS IF NO MODIFICATIONS HAVE BEEN MADE BY THE USER)        *
***********************************************************************************************************.

INCLUDE FILE='directory_name_of_macroe_v30.sps\MACROE_V30.SPS'.



* DO NOT MODIFY THE NEXT LINE. (GO TO SECTION 2)
***********************************************************.
!Prepare ident=!ident /fwgt=!fwgt /prefix=!prefix /B=!B /Classes=!Classes /mfile=!Mfile /bwsav=!bwsav.




*************************************************************************************
***                            SECTION 2                              ***
*************************************************************************************
***                                                                   ***
***       This section lets the user specify the different analyses of interest.   ***
***                                                                   ***
*************************************************************************************.
...
...
...

* TO OBTAIN VARIANCE ESTIMATES OF A TOTAL, RUN:
*************************************************************.

*!Total !B !Classes/ Var = name_of_the_variable .
*!Let !Tot=1.


* TO OBTAIN VARIANCE ESTIMATES OF A RATIO, RUN:
*************************************************************.

*!Ratio !B !Classes/ Varlist = name_of_numerator   name_of_denominator .
*!Let !Tot=1.


* TO OBTAIN VARIANCE ESTIMATES OF A DIFFERENCE OF RATIO, RUN:
*************************************************************************.
    ****    NOTE: see the comment at the beginning of section 2 ...   ****

*!Diffrat !B !Classes/ Varlist = variable1   variable2   variable3   variable4 .
*!Let !Tot=1.

                * where: variable1  : the numerator variable of the first ratio    *
                *        variable2  : the denominator variable of the first ratio   *
                *        variable3  : the numerator variable of the second ratio    *
                *        variable4  : the denominator variable of the second ratio  *.


* TO OBTAIN VARIANCE ESTIMATES OF REGRESSION PARAMETERS, RUN:
*************************************************************************.
```

```
*!Regress  !B !Classes/   Dep = dependant_variable  Indep = Independant_variable(s)
*!Let !Reg=1.


* TO OBTAIN VARIANCE ESTIMATES OF LOGISTIC REGRESSION PARAMETERS, RUN:
********************************************************************************************************.

*!Log_Reg !B !Classes/ Cri=SAS Dep = dependant_variable  Indep = Independant_variable(s)
*!Let !Reg=1.

        * NOTE: Default iteration criteria differs between SAS and SPSS.
        *     Use Cri=SAS to use SAS default values.
        *     Indep= must be specified LAST because of the unknown number of independent variables.




**********************************************************************
*           DO NOT MODIFY THE REMAINING LINES            *
**********************************************************************.

SET PRI NO MPR NO.

**** Printing of results ****.
!IF (!Tot=1) !THEN
+ !Print_T !Classes.   /* Printing of Totals, Ratios and Differences of Ratios results
!IFEND
!IF (!Reg=1) !THEN
+ !Print_R !Classes.   /* Printing of Linear and Logistic Regressions results
!IFEND
******************************.

!Stop !NBlocks.
!ENDDEFINE /* !Bootvar.

SET MPR ON.
!Bootvar.
SET MPR NO.


*** END OF BOOTVARE_V30.SPS PROGRAM  ***.
```

APPENDIX B

This is a complete example showing how to use the program BOOTVARE_V30.SPS. First, the analysis data file is created (step 1). Then, BOOTVARE_V30.SPS is adapted to obtain the desired analysis. The results that are produced follow the programs.

Example:

This example uses the cycle 3 (1998) cross-sectional file of the National Population Health Survey, general component. This example:

1-   Computes the total number and the proportion of diabetics in the population and for the men, for each province (only four provinces will be kept).

2-   Studies the relationship between diabetes, sex and type of interview (proxy or not), for each province.

The different parameters needed in the program (specified in AppendixC_Health) are:

NPHS - Household Component							
	Name of data file	**Name of bootstrap weights file**	**Identification variables**	**Weight variable** *(on the data file)*	**Weight variable** *(on the bootstrap weights file)*	**Prefixe of the bootstrap weights**	**# of weights**
Cycle 3 - General component	H35.txt	B5H35	REALUKEY PERSONID	WT58 (M) WT58_S (S)	FWGT	BSW	500

Step 1:

```
*************************************************************************************************
*                                    STEP1.SPS                                    *
*                                                                                 *
*  This program creates the SPSS datafile containing the necessary variables for the  *
*  BOOTVARE_V30.SPS program.                                                      *
*                                                                                 *
*  IT IS NOT NECESSARY TO USE THIS PROGRAM.  Dichotomous variables can            *
*  be computed directly in the Data Editor window and the analysis data file can be  *
*  created from there.                                                            *
*************************************************************************************************.

SET COM NO HEA ON MES ON ERR ON RES ON JOU ON LEN NONE WID 132 PRI NO MPR NO.


* Creation of the SPSS "sav" file containing the variables and cases required for the analysis. Note
* that this file should be as small as possible (containing only necessary variables and cases) in
* order to reduce time and memory requirements especially if regression type analysis are to be done.


DEFINE !Go()
!LET !Layout = 'D:\LAYOUT\h35_i.sps'
!LET !DataIn = ' D:\DATA\h35.txt'
!LET !DataOut = ' C:\BOOTVAR\diabetes.sav'          /* <- (analysis file to be used in BOOTVARE_V30.SPS)


* Do not modify the following block of instructions.
FILE HANDLE INFILE/NAME = !QUOTE(!DataIn).
INCLUDE FILE=!QUOTE(!Layout).
SET PRI ON.


*************************************************************************************************************
* Creation of Dichotomous Variables ( -> REPLACE all following instructions (up to "SAVE"  )    *
* (examples are presented below using NPHS cycle 3 variables)                                    *
*************************************************************************************************************.


NUMERIC diab nonproxy total males females mdiab fdiab.

IF (ccc8_1j=1)  diab=1.
IF (ccc8_1j~=1) diab=0.

compute nonproxy = 0.
IF (am58_pxy>2)   nonproxy=$SYSMIS.
IF (am58_pxy=2)   nonproxy=1.

compute total = 1.

IF (dhc8_sex=1) males=1.
IF (dhc8_sex~=1) males=0.
IF (dhc8_sex=2) females=1.
IF (dhc8_sex~=2) females=0.

compute mdiab = males*diab.
compute fdiab  = females*diab.

SELECT IF (prc8_cur = 10 or prc8_cur = 24 or prc8_cur = 35 or prc8_cur = 59).

execute.

*******************************************************************************************.

SAVE OUTFILE !QUOTE(!DataOut)
        /KEEP =  diab total males females mdiab fdiab nonproxy wt58 realukey personid prc8_cur.

        * It is recommended that only the necessary variables be kept     *
        * in order to reduce the runtime of BOOTVARE_V30.SPS.             *
        * IMPORTANT:  the identification variables and, if necessary,     *
        * the breakdown variable(s) (ex: province, sex) must be kept. The *
        * weight variable also must be kept if point estimates are        *
```

    * calculated at this step   *.

* Replace "SAVE" by "XSAVE" in previous command if point estimates are desired at this time.

*** DO NOT FORGET TO RUN THE LAST LINES OF THE PROGRAM (AFTER THE EXAMPLE) ***.


```
*********************************************
* Calculation of point estimates       *
* Suggested, but not required...        *
*********************************************.

*** Examples: ***.

SORT CASES BY prc8_cur.
SPLIT FILE BY prc8_cur.

WEIGHT BY wt58.

FREQ VAR=diab mdiab fdiab.
LOGISTIC REG VAR=diab WITH nonproxy females /PRI SUM
         /CRI BCO(0.0001) LCO(0) ITE(25) EPS(10E-12).

WEIGHT OFF.
SPLIT FILE OFF.


* DO NOT DELETE the following instructions
****************************************************.

SET PRI NO.
!ENDDEFINE /* !Go.

SET MPR ON.
!Go.
SET MPR NO.
```

Step 2 - BOOTVARE_V30.SPS_program:

```
. . .
. . .
. . .


*****************************************************************************
***                      SECTION 1                         ***
*****************************************************************************
***                                                        ***
***   This section lets the user specify the different parameters of   ***
***   interest (variable names, directory names, file names, etc.)     ***
***                                                        ***
*****************************************************************************.


DEFINE !Bootvar()

PRESERVE.
SET COM NO HEA ON MES ON ERR ON RES ON JOU ON LEN NONE WID 132 PRI NO MPR NO MITERATE=3000.

!Let !Tot=!Null
!Let !Reg=!Null



***********************************************************************************************
* SPECIFY THE ANALYSIS DATA FILE TO BE USED (created in step 1) (Folder and file) :    *
***********************************************************************************************.

!LET !Mfile='C:\BOOTVAR\diabetes.sav'                 /*  <- (ex: c:\data\analysis.sav).



***********************************************************************************************
*   SPECIFY THE BOOTSTRAP WEIGHTS FOLDER AND FILE NAMES :            *
*         NB:  Only run one of the two following series of commands        *
*            (comment the other one out, or erase it):                *
***********************************************************************************************.

* EXECUTE THIS PART IF THE BOOTSTRAP WEIGHTS ARE IN SPSS FORMAT (remove the " * ") *
***********************************************************************************************.

!LET !BWsav='D:\Bootstrp\data\b5h35.sav'


* EXECUTE THIS PART IF THE BOOTSTRAP WEIGHTS ARE IN ASCII (.TXT) FORMAT (remove the " * ") *
***********************************************************************************************.

*!LET !BWTXT=' Folder_and_file_names_of_the_bootstrap_weights_in_text_format_(with_extension) '
*!LET !Bwlay=' Folder_and_file_names_of_the_bootstrap_layout_file_in_SPSS_syntax_(with_extension) '
*!LET !Bwsav=' Folder_and_file_names_of_the_bootstrap_weights_in_SPSS_format_(created_in_this_section)_with_extension '

*FILE HANDLE BOOTLAY/NAME = !QUOTE(!BWTXT).
*INCLUDE FILE=!QUOTE(!BWlay).
*SAVE OUT=!QUOTE(!BWsav) /UNC.



*****************************************************************************************************
* SPECIFY, IF DESIRED, THE BREAKDOWN VARIABLE(S) (EG: PROVINCE, SEX, ETC...):           *
*     Write the name of the breakdown variable(s) below.                    *
*                                                                *
*     - If the analysis includes all of the data in the file created in step 1, put two single quotes (%LET !Classes = ' ' )   *
*     - If more than one variable, leave a space between each variable  ( %LET !Classes = ' var1 var2 ' )   *
*     - DO NOT ERASE OR COMMENT OUT THIS COMMAND               *
*****************************************************************************************************.


!LET !Classes = 'prc8_cur'
```

```
**********************************************************************************************
* SPECIFY THE FOLLOWING INFORMATION (SPECIFIC TO THE SURVEY YOU ARE USING):      *
*   You must specify:                                                            *
*       1- The unique identifier variable(s) (within quotes, separated by a space)   *
*       2- The Final Weight (variable included in the bootstrap weight file)         *
*       3- The prefix of the boostraps weight variables                              *
*       4- The number of bootstrap weights to use (note: For testing, B must be >= 2.  *
*          IT IS NECESSARY TO USE ALL THE BOOTSTRAP WEIGHTS WHEN PERFORMING THE FINAL *
*          ANALYSIS.  THE COMPLETE BOOTVARE_V30.SPS PROGRAM MUST THEN BE RUN.)        *
*                                                                                    *
*       - Refer to AppendixC_XYZ  to obtain this information                          *
**********************************************************************************************.

!LET !ident  = ' realukey personid '
!LET !fwgt   = fwgt
!LET !prefix = bsw
!LET !B      = 500


**********************************************************************************************
* SPECIFY THE DIRECTORY AND THE NAME OF THE FILE THAT CONTAINS THE MACROS         *
* (THE PROGRAM MACROE_V30.SPS IF NO MODIFICATIONS HAVE BEEN MADE BY THE USER)     *
**********************************************************************************************.

INCLUDE FILE='directory_name_of_macroe_v30.sps\MACROE_V30.SPS'.


* DO NOT MODIFY THE NEXT LINE. (GO TO SECTION 2)
*****************************************************************.
!Prepare ident=!ident /fwgt=!fwgt /prefix=!prefix /B=!B /Classes=!Classes /mfile=!Mfile /bwsav=!bwsav.


***********************************************************************************
***                          SECTION 2                          ***
***********************************************************************************
***                                                             ***
***      This section lets the user specify the different analyses of interest.    ***
***                                                             ***
***********************************************************************************.

...
...
...

* TO OBTAIN VARIANCE ESTIMATES OF A TOTAL, RUN:
*****************************************************************.

*!Total !B !Classes/ Var = name_of_the_variable .
*!Let !Tot=1.

!Total !B !Classes/ Var = diab.
!Total !B !Classes/ Var = mdiab.
!Let !Tot=1.

* TO OBTAIN VARIANCE ESTIMATES OF A RATIO, RUN:
*****************************************************************.

*!Ratio !B !Classes/ Varlist = name_of_numerator   name_of_denominator .
*!Let !Tot=1.

!Ratio !B !Classes/ Varlist = diab total.
!Ratio !B !Classes/ Varlist = mdiab males.
!Let !Tot=1.


* TO OBTAIN VARIANCE ESTIMATES OF A DIFFERENCE OF RATIO, RUN:
*****************************************************************.
    ****    NOTE: see the comment at the beginning of section 2 ...  ****

*!Diffrat !B !Classes/ Varlist =  variable1   variable2   variable3   variable4 .
*!Let !Tot=1.

            * where:  variable1   : the numerator variable of the first ratio        *
```

```
*            variable2  : the denominator variable of the first ratio     *
*            variable3  : the numerator variable of the second ratio      *
*            variable4  : the denominator variable of the second ratio  *.


* TO OBTAIN VARIANCE ESTIMATES OF REGRESSION PARAMETERS, RUN:
******************************************************************************************.

*!Regress  !B !Classes/    Dep = dependant_variable  Indep = Independant_variable(s)
*!Let !Reg=1.


* TO OBTAIN VARIANCE ESTIMATES OF LOGISTIC REGRESSION PARAMETERS, RUN:
*************************************************************************************************.

*!Log_Reg !B !Classes/ Cri=SAS Dep = dependant_variable  Indep = Independant_variable(s)
*!Let !Reg=1.

!Log_Reg !B !Classes/ Cri=SAS Dep = diab     Indep = nonproxy females
!Let !Reg=1.

    * NOTE: Default iteration criteria differs between SAS and SPSS.
    *     Use Cri=SAS to use SAS default values.
    *     Indep= must be specified LAST because of the unknown number of independent variables.




***********************************************************************
*        DO NOT MODIFY THE REMAINING LINES           *
***********************************************************************.

SET PRI NO MPR NO.

**** Printing of results ****.
!IF (!Tot=1) !THEN
+ !Print_T !Classes.   /* Printing of Totals, Ratios and Differences of Ratios results
!IFEND
!IF (!Reg=1) !THEN
+ !Print_R !Classes.   /* Printing of Linear and Logistic Regressions results
!IFEND
*****************************.

!Stop !NBlocks.
!ENDDEFINE /* !Bootvar.

SET MPR ON.
!Bootvar.
SET MPR NO.


*** END OF BOOTVARE_V30.SPS PROGRAM  ***.
```

Results and interpretation:

The tables on the next page present the results of the analyses done using the programs from the example. Results for the totals and ratios are presented in the first table. For example, if we want the ratio of the number of diabetic males to the total number of males, in Ontario, we look at the twelfth line. The region 35 corresponds to the province of Ontario (see the data dictionary document included on the CD-ROM for the codes associated with each province) and the variable Type indicates the type of analysis, in this case a ratio. We find the variables *mdiab* as the numerator of the ratio and *males* as the denominator (VARIABLES). The estimate of the ratio is 3.57% (YHAT) with a standard deviation of 0.29 (BS_SD) and a coefficient of variation of 8.26 (BS_CV). The 95% confidence interval for this estimate is (2.99%, 4.14%) (CIL95, CIU95).

Results from the logistic regression are shown in the second table. For example, the estimate of the parameter for the variable female in Ontario (ninth line) is -0.34507 (BHAT) and the odds ratio is 0.70817 (ODDS). The Wald's statistic for this parameter and its associated p-value are 7.9964 (WALD) and p=0.004687 (PVALUE) respectively. The estimates of the standard deviation for the parameter estimate is 0.12203 (BS_SD) and the coefficient of variation is 35.36 (BS_CV). Finally, the confidence interval for the odds ratio is (0.55753, 0.89952) (CIL95,CIU95).

Following the results, is the execution time of this program.

21 Mar 02   Bootstrap [500 weights] Variance Estimate

PRC8_CUR	TYPE.	VARIABLE	YHAT	BS_SD	BS_CV	CIL95	CIU95
10	Total	DIAB	20741.32	1778.61	8.58	17255.24	24227.40
10	Total	MDIAB	7029.13	1356.31	19.30	4370.75	9687.51
10	Ratio	DIAB TOTAL	3.85	.33	8.58	3.21	4.50
10	Ratio	MDIAB MALES	2.63	.51	19.30	1.64	3.63
24	Total	DIAB	205292.27	16330.49	7.95	173284.50	237300.04
24	Total	MDIAB	110452.80	10818.85	9.79	89247.85	131657.75
24	Ratio	DIAB TOTAL	2.87	.23	7.95	2.42	3.32
24	Ratio	MDIAB MALES	3.12	.31	9.80	2.52	3.72
35	Total	DIAB	362439.60	20692.51	5.71	321882.28	402996.92
35	Total	MDIAB	198237.67	16369.18	8.26	166154.08	230321.26
35	Ratio	DIAB TOTAL	3.22	.18	5.71	2.86	3.58
35	Ratio	MDIAB MALES	3.57	.29	8.26	2.99	4.14
59	Total	DIAB	110375.39	10661.17	9.66	89479.50	131271.28
59	Total	MDIAB	62808.64	8301.60	13.22	46537.50	79079.78
59	Ratio	DIAB TOTAL	2.83	.27	9.66	2.29	3.37
59	Ratio	MDIAB MALES	3.24	.43	13.22	2.40	4.08

21 Mar 02   Bootstrap [500 weights] Variance Estimate

PRC8_CUR	TYPE	BETA	BHAT	ODDS	WALD	PVALUE	BS_SD	BS_CV	CIL95	CIU95
10	Logistic Regression	Const_	-4.00379	.01825	255.3510	.00000000	.25055	6.26	.01117	.02982
10	Logistic Regression	NONPROXY	.85788	2.35815	11.8177	.00058671	.24955	29.09	1.44594	3.84587
10	Logistic Regression	FEMALES	.46627	1.59403	2.8588	.09087612	.27577	59.14	.92845	2.73674
24	Logistic Regression	Const_	-3.89894	.02026	852.2376	.00000000	.13356	3.43	.01560	.02633
24	Logistic Regression	NONPROXY	.89123	2.43812	34.0970	.00000001	.15263	17.13	1.80774	3.28831
24	Logistic Regression	FEMALES	-.36991	.69080	5.7831	.01618064	.15382	41.58	.51100	.93387
35	Logistic Regression	Const_	-3.57575	.02799	961.1974	.00000000	.11533	3.23	.02233	.03510
35	Logistic Regression	NONPROXY	.60946	1.83943	23.5198	.00000124	.12567	20.62	1.43785	2.35318
35	Logistic Regression	FEMALES	-.34507	.70817	7.9964	.00468716	.12203	35.36	.55753	.89952
59	Logistic Regression	Const_	-3.99966	.01832	301.8469	.00000000	.23021	5.76	.01167	.02877
59	Logistic Regression	NONPROXY	1.05766	2.87963	18.8915	.00001384	.24334	23.01	1.78732	4.63951
59	Logistic Regression	FEMALES	-.46337	.62916	4.5912	.03213705	.21625	46.67	.41180	.96126

```
-------------------------------------------------
Actual Time:            21-MAR-2002 14:35:20.700
Starting Time:          21-MAR-2002 13:45:22.048
Total Elapsed Time to run Program:   0:49:58.652
-------------------------------------------------
```