

**Canadian Community Health Survey
Cycle 1.1, 2000-2001
Dummy File for Remote Access to the Integrated Master File**

April 2003

Notice

The Dummy File should not be used for purposes other than to develop and test the computer programs that are to be submitted by remote access. The Dummy File contains modified data, and must never be used for analytical purposes.

1. Background and Overview

The Canadian Community Health Survey (CCHS) is a new survey conducted by Statistics Canada in collaboration with its partners to collect information related to health, health status and health care utilization in the Canadian population at the sub-provincial level. The CCHS collects data on a biennial cycle composed of two distinct surveys: a regional survey in the first year and a provincial survey in the second. At this point, the first cycle has been completed.

In order to provide extensive access to the data, a public use microdata file (PUMF) was produced for Cycle 1.1. The creation of a PUMF involves the application of rigorous procedures to ensure data confidentiality. Consequently, survey variables may have to be grouped, capped, or simply suppressed.

As an alternative, in cases where the PUMF does not meet their needs, researchers can perform their analyses on the CCHS Integrated Master File. In order to provide researchers with a means to access the Integrated Master File, a remote access facility has been implemented. Remote access provides researchers with the possibility to submit computer programs via e-mail to a dedicated address (cchs-escc@statcan.ca), and to receive the results by return e-mail. To obtain remote access privileges, researchers must necessarily obtain advance approval from the Health Statistics Division. Requests must be submitted to the aforementioned e-mail address and must provide the following, clearly itemised information:

- the researcher's affiliation,
- the name of all researchers involved in the project,
- the title of the research project,
- an abstract of the project,
- the goals of the research,
- the data to which access is required (survey, cycle),
- why the project requires the access to the master data rather than the PUMF,
- why the Remote Access service is chosen rather than the on-site access in a Research Data Centre (RDC),
- the expected results, and
- the project's expected completion date.

Further information is available by contacting the CCHS team at the above e-mail address or by phone at (613) 951-1653. Once the request for remote access has been approved, the researcher can submit his/her computer programs to the CCHS team for processing on the Integrated Master File(s). The computer output is reviewed by the team for confidentiality concerns and returned to the researcher. However, the correctness and accuracy of each program submission remains at all times the sole responsibility of the researcher.

With the Dummy File supplied on this CD-ROM, the researcher can develop and test his/her computer programs before submitting them to the CCHS team. The Cycle 1.1 Dummy File applies to the data file containing all 131,535 respondents. While certain administrative variables, which are of no analytical interest, have been recoded to "blank" or "9", the Dummy File simulates the Integrated Master File perfectly. It contains the same variables and has the same record layout. The data values and the Dummy File size however, have been modified in order to protect the confidentiality of respondents. The CCHS Cycle 1.1 Dummy File contains 65,101 records, approximately 50% of the master file size. Users should note that the "Nord du Québec" health region is excluded from the PUMF for confidentiality reasons but is included in the Dummy File.

This CD-ROM also includes SAS and SPSS macro programs for calculating the variance of estimates. Four dummy bootstrap weight files, each one corresponding to one of the four Integrated Master File weights present on the Dummy File, are also included to help develop and test the variance calculation programs. These dummy bootstrap weight files simulate the original files, have the same record layouts and contain the same variables as those in the original files, but the weight values have been modified. Since the number of records in the Dummy File has been reduced, the number of records in the Dummy Bootstrap File has consequently been reduced. The CD-ROM also contains peripheral SAS and SPSS layout statement files, as well as the technical and methodological documentation usually accompanying the Integrated Master Files. Please refer to Section 3 of the present document for the complete list of files on this CD-ROM.

The following sections of this document describe in more detail the steps leading to the creation of the Dummy File.

2. Creation of the Dummy File

2.1 Classes of Variables

The Integrated Master file records are first classified by age, sex and health region (for optional modules only). The purpose for this is to create classes of records having the same questionnaire flow so that, after random permutation of the data within the classes, each dummy record contain data respecting a certain internal consistency.

2.2 Block of variables

The second step consists of creating blocks of variables. Dummy records are produced by randomly swapping blocks of variables for the records in the same class. At first, elementary blocks of variables are grouped following the various modules of the questionnaire. If the content of the elementary block depends on the answers obtained in a previous block, the respective elementary block is amalgamated with the other one to preserve the internal consistency of the records.

2.3 Sampling and Restrictions of Variable Values

As an additional measure of preserving data confidentiality in the Integrated Master File, the Dummy File was created using a sub-sample composed of 50% of the Integrated Master File records (once it has been perturbed). In addition, quantitative variables were top-coded. Variables considered as sensitive or without analytical value have been recoded to “blank” or to “9” (see section 2.4). All these modifications mean that we do not necessarily find, on the Dummy File, all the possible categories for each variable. To obtain complete information on a specific variable, users are asked to consult the Integrated Master File documentation. Finally, all the weight variables have been repoststratified according to the same methods used to poststratify the Integrated Master File weights. However, it should be noticed that the weights remain “dummy” weights. Other details related to the creation of this file are not available because they would disclose too much information on the method used for creating the dummy data, hence putting confidentiality at risk.

2.4 Recoded Variables to “blank” and to “9”

The following variables have been recoded to “blank” or to “9” during the creation of the Dummy file:

ADMA_RNO, PERSONID, SAMA_TYP, SAMA_CP, ADMA_PRX, ADMA_ENT, ADMA_IMP, ADMA_N09, ADMA_N10, ADMA_N11, ADMA_N12, ADMA_DAT, ADMA_STA, ADMA_LHH, DHHA_DB, SAMADNUM, ADMAFFE, ADMA_PH, INCAFDN et FINAFDN.

3. Content of CD-ROM

LISEZ_MOI.PDF	French documentation about remote-access, creation of the Dummy File and contents of the CD-ROM
READ_ME.PDF	English documentation about remote-access, creation of the Dummy File and contents of the CD-ROM

Files related to bootstrap weights for variance calculation:

BOOTSTRP\ LISEZ_MOI.TXT READ_ME.TXT	French documentation of the files related to the use of the Bootstrap program. English documentation of the files related to the use of the Bootstrap program.
BOOTSTRP\DATA\ B5.TXT B5_PEI.TXT B5_Q4.TXT B5_BC.TXT	Bootstrap weights file in ASCII format. Special bootstrap weights file in ASCII format for Prince Edward Island. Special bootstrap weights file in ASCII format for the 4 th Quarter. Special bootstrap weights file in ASCII format where British Columbia is subdivided into 16 health regions.
BOOTSTRP_LAYOUT\ B5_I.SAS B5_I.SPS B5_PEI_I.SAS B5_PEI_I.SPS B5_Q4_I.SAS B5_Q4_I.SPS B5_BC_I.SAS B5_BC_I.SPS	SAS Record Layout for the bootstrap weight file B5.TXT SPSS Record Layout for the bootstrap weight file B5.TXT SAS Record Layout for the bootstrap weight file B5_PEI.TXT SPSS Record Layout for the bootstrap weight file B5_PEI.TXT SAS Record Layout for the bootstrap weight file B5_Q4.TXT SPSS Record Layout for the bootstrap weight file B5_Q4.TXT SAS Record Layout for the bootstrap weight file B5_BC.TXT SPSS Record Layout for the bootstrap weight file B5_BC.TXT
BOOTSTRP\SAS\DOC\ SASBOOTDOC_F.PDF SASBOOTDOC_E.PDF	French documentation of the BOOTVARF_V20.SAS program. English documentation of the BOOTVARE_V20.SAS program.
BOOTSTRP\SAS\PGM\ BOOTVARF_V20.SAS BOOTVARE_V20.SAS MACROF_V20.SAS MACROE_V20.SAS ETAPE1.SAS STEP1.SAS	SAS bootstrap variance program with French comments. SAS bootstrap variance program with English comments. SAS Macro program with French comments. SAS Macro program with English comments. Example, with French comments, of the SAS program creating the data analysis file. Example, with English comments, of the SAS program creating the data analysis file.
BOOTSTRP\SPSS\DOC\ SPSSBOOTDOC_F.PDF SPSSBOOTDOC_E.PDF	French documentation of the BOOTVARF_V21.SPS program. English documentation of the BOOTVARE_V21.SPS program.
BOOTSTRP\SPSS\PGM\ BOOTVARF_V21.SPS BOOTVARE_V21.SPS MACROF_V21.SPS MACROE_V21.SPS ETAPE1.SPS STEP1.SPS	SPSS bootstrap variance program with French comments. SPSS bootstrap variance program with English comments SPSS Macro program with French comments. SPSS Macro program with English comments. Example, with French comments, of the SPSS program creating the data analysis file. Example, with English comments, of the SPSS program creating the data analysis file.

Dummy File :

DATA\ DUMY_HSI.TXT	Dummy file of Cycle 1.1, 65,101 records
-----------------------	---

Documentation:

DOC\PDF_E\ DD_E.PDF INDEX_A_E.PDF INDEX_T_E.PDF LAYOUT_E.PDF QUEST_E.PDF DVDOC_E.PDF	Data Dictionary in English. English alpha index. English topical index. English record layout. English Questionnaire. English derived variables documentation.
DOC\PDF_F\ DD_F.PDF INDEX_A_F.PDF INDEX_T_F.PDF LAYOUT_F.PDF QUEST_F.PDF DVDOC_F.PDF	Data Dictionary in French. French alpha index. French topical index. French record layout. French Questionnaire. French Derived variables documentation.

Record Layouts, Statements :

LAYOUT\ HSI_FMT.SAS HSI_I.SAS HSI_LBF.SAS HSI_LBE.SAS HSI_PFF.SAS HSI_PFE.SAS HSI_I.SPS HSIMISS.SPS HSIVALF.SPS HSIVALE.SPS HSIVARF.SPS HSIVARE.SPS	SAS Format statement. SAS Infile and Input statements. SAS Label statement in French. SAS Label statement in English. SAS Proc Format statement in French. SAS Proc Format statement in English. SPSS Infile statement. SPSS Missing Values statement. SPSS Value Labels statement in French. SPSS Value Labels statement in English. SPSS Variable Labels statement in French. SPSS Variable Labels statement in English.
---	---

