

**Enquête sur la santé dans les collectivités canadiennes**  
**Cycle 1.2, 2002**  
**Fichier fictif pour le télé-accès au fichier maître**

Janvier 2005

**Avertissement**

*Le fichier fictif ne devrait pas servir à d'autres buts que le développement et la vérification des programmes informatiques devant être soumis par télé-accès. Le fichier fictif comprend des données modifiées et ne doit en aucun cas être utilisé à des fins d'analyse.*

## 1. Introduction et contexte

L'Enquête sur la santé dans les collectivités canadiennes (ESCC) est une enquête transversale qui vise à recueillir des renseignements sur l'état de santé, l'utilisation des services de santé et les déterminants de la santé de la population canadienne. Le cycle de collecte des données de l'ESCC s'étend sur deux années. La première année du cycle, indiquée par la notation « .1 », correspond à une enquête générale sur la santé de la population réalisée auprès d'un grand échantillon et conçue pour fournir des estimations fiables pour la région socio-sanitaire. La deuxième année du cycle, représentée par la notation « .2 », correspond à une enquête de moins grande portée conçue pour fournir des données à l'échelle provinciale sur des sujets particuliers ayant trait à la santé.

Afin de permettre un accès élargi aux données, un fichier de microdonnées à grande diffusion (FMGD) a été produit pour le cycle 1.2. La création d'un FMGD nécessite l'application de procédures rigoureuses afin d'assurer la confidentialité des données. Ainsi, certaines variables de l'enquête doivent être soit agrégées, soit assujetties à des valeurs maximales ou minimales, soit simplement supprimées.

Dans les cas où le FMGD ne peut suffire à ces besoins, un chercheur a comme alternative d'exécuter ses analyses sur le fichier maître de l'ESCC. Afin justement d'offrir aux chercheurs la possibilité d'avoir accès à ce fichier maître, un service de télé-accès a été instauré. Ce service permet au chercheur de soumettre des programmes informatiques par courriel à une adresse électronique dédiée (cchs-escc@statcan.ca) et d'en recevoir les résultats par retour de courriel. Pour se prévaloir du service de télé-accès, le chercheur doit nécessairement obtenir l'approbation préalable de la Division de la statistique de la santé. Il doit, à cette fin, soumettre une demande d'accès à l'adresse électronique susmentionnée dans laquelle sont clairement indiqués les points suivants :

- l'affiliation du chercheur,
- les noms de tous les chercheurs impliqués dans le projet de recherche,
- le titre du projet de recherche,
- un résumé du projet de recherche,
- les objectifs de la recherche,
- les données requises (enquête, cycle),
- pourquoi la recherche nécessite le recours au fichier maître plutôt que le FMGD,
- pourquoi le service de télé-accès est choisi au lieu de l'accès sur place dans un centre de données de recherche (CDR),
- les résultats projetés, et
- la date de complétion prévue du projet.

Pour obtenir plus de renseignements sur le processus, veuillez communiquer avec l'équipe de l'ESCC à l'adresse électronique mentionnée précédemment ou par téléphone au (613) 951-1653. Une fois l'approbation au télé-accès obtenue, le chercheur peut soumettre ses programmes informatiques à l'équipe de l'ESCC pour le traitement des données du fichier maître. L'équipe vérifie les sorties des programmes par rapport aux normes de confidentialité et les transmet au chercheur. Cependant, le bon fonctionnement et l'exactitude de chacun des programmes demeurent en tout temps l'entière responsabilité du chercheur.

Le fichier fictif qui est fourni sur ce CD-ROM permet au chercheur de développer et de vérifier ses programmes informatiques avant de les soumettre. Le fichier fictif du cycle 1.2 est basé sur les données recueillies auprès de 36 984 répondants. Bien que certaines variables à caractère administratif ne représentant aucun intérêt analytique aient été recodées à « blanc » ou à « 9 », le fichier fictif présente la même structure que le fichier maître; il renferme les mêmes variables et a le même cliché d'article. Cependant, les valeurs des variables ont été modifiées afin de préserver la confidentialité des répondants. Le fichier fictif du cycle 1.2 de l'ESCC contient 36 984 enregistrements.

Le CD-ROM contient également des programmes SAS et SPSS pour le calcul de la variance des estimations. Un fichier fictif de poids bootstrap associé aux poids du fichier maître étant présent sur le fichier fictif, est également inclus afin de permettre au chercheur de développer et vérifier ses programmes de calcul de variance. Ce fichier fictif de poids bootstrap simule le fichier original, utilise le même cliché d'article et contient les mêmes variables que celui-ci, mais les valeurs des poids sont modifiées par rapport à celles présentes dans le fichier original. On retrouve également sur le CD-ROM, des fichiers périphériques d'énoncés SAS et SPSS pour la lecture des fichiers, ainsi que la documentation technique et méthodologique accompagnant normalement le fichier maître. Veuillez consulter la section 3 du présent document pour la liste complète des fichiers de ce CD-ROM.

Les prochaines sections de ce document décrivent de façon plus détaillée les étapes menant à la création du fichier fictif.

## **2. Création du fichier fictif**

### **2.1 Catégories de variables**

Les enregistrements du fichier maître sont d'abord classés selon l'âge, le sexe. Un des objectifs consiste à former des catégories regroupant des enregistrements pour lesquels le cheminement à travers le questionnaire est le même, de sorte que, après permutation aléatoire des données à l'intérieur des catégories, chaque enregistrement artificiel résultant contienne des données respectant une certaine cohérence interne.

### **2.2 Échantillonnage et restriction de valeurs des variables**

Comme mesure additionnelle, le fichier fictif a été créé en sous-échantillonnant 50 % des enregistrements du fichier maître. Ce sous-échantillon de 50% est toutefois dupliqué avant de perturber les données pour faire en sorte que le fichier fictif ait une taille identique à celle du fichier maître. De plus, les variables quantitatives ont été tronquées par le haut et par le bas. Finalement, certaines variables jugées sensibles à la confidentialité ou n'ayant aucune valeur analytique ont quant à elles été tout simplement recodées à « blanc » ou « 9 » (voir sous-section 2.4). Toutes ces modifications font en sorte qu'on ne retrouve pas nécessairement toutes les modalités possibles pour chaque variable sur le fichier fictif. Pour obtenir la liste complète des modalités pour chaque variable, les utilisateurs sont priés de se référer à la documentation du fichier maître. Finalement, les poids d'échantillonnage et les poids bootstrap ont été repoststratifiés selon les mêmes méthodes utilisées lors de la poststratification des poids du fichier maître. Cependant, il est à noter que ces poids demeurent des poids artificiels. Les autres détails relatifs à la création de ce fichier ne sont pas rapportés car ils pourraient fournir trop de renseignements sur la méthode utilisée pour créer les données fictives et mettraient la confidentialité en péril.

### **2.3 Blocs de variables**

La troisième étape consiste à créer des blocs de variables. On produit les enregistrements artificiels pour le fichier fictif en permutant de façon aléatoire des blocs de variables entre les enregistrements d'une même catégorie. On regroupe d'abord les variables en blocs élémentaires correspondant aux diverses sections (modules) du questionnaire. Dans le cas où le contenu d'un bloc élémentaire dépend des réponses obtenues à des questions d'un bloc précédent, les blocs élémentaires respectifs ont été intégrés ensemble afin d'assurer la cohérence interne des enregistrements.

### **2.4 Variables recodées à « blanc » ou « 9 »**

Les variables énumérées ci-après ont été recodées à « blanc » ou « 9 » lors de la création du fichier fictif :

PERSONID SAMB\_ACP SAMB\_OCP ADMB\_ENT ADMB\_DAT ADMB\_STA ADMBDN09 ADMB\_N10  
ADMB\_LHH ADMB\_N12 DHHB\_YOB DHHB\_MOB DHHB\_DOB FILLR001 FILLR002 FILLR004 FILLROO1

## **3. Contenu du CD-ROM**

LISEZ_MOI.PDF	Documentation en français sur le téléaccès, la création du fichier fictif, le contenu du CD-ROM.
READ_ME.PDF	Documentation en anglais sur le téléaccès, la création du fichier fictif, le contenu du CD-ROM.

*Fichier fictif :*

DATA\  
HSI.TXT                      Fichier fictif du cycle 1.2, 36 984 enregistrements

*Clichés d'enregistrement, énoncés :*

LAYOUT\  
HSI\_FMT.SAS                Énoncé SAS Format.  
HSI\_I.SAS                  Énoncés SAS Infile et Input.  
HSI\_LBF.SAS                Énoncé SAS Label en français.  
HSI\_LBE.SAS                Énoncé SAS Label en anglais.  
HSI\_PFF.SAS                Énoncé SAS Proc Format en français.  
HSI\_PFE.SAS                Énoncé SAS Proc Format en anglais.  
  
HSI\_I.SPS                    Énoncé SPSS Infile.  
HSIMISS.SPS                Énoncé SPSS Missing Values.  
HSIVALF.SPS                Énoncé SPSS Value Labels en français.  
HSIVALE.SPS                Énoncé SPSS Value Labels en anglais.  
HSIVARF.SPS                Énoncé SPSS Variable Labels en français.  
HSIVARE.SPS                Énoncé SPSS Variable Labels en anglais.  
  
READFILE.SAS              Énoncés SAS pour lire, étiqueter et formater le fichier ASCII.  
READFILE.SPS              Énoncés SPSS pour lire, étiqueter et formater le fichier ASCII.

*Documentation:*

DOC\PDF\_E\  
DD\_E.PDF                    Dictionnaire de données en anglais.  
INDEX\_A\_E.PDF              Index alphabétique des variables en anglais.  
INDEX\_T\_E.PDF              Index des variables par sujet en anglais.  
LAYOUT\_E.PDF               Cliché d'enregistrement en anglais.  
QUEST\_E.PDF                Questionnaire en anglais.  
QUEST\_HHLD\_E.PDF          Questionnaire ménage en anglais.  
DVDOC\_E.PDF                Documentation sur les variables dérivées en anglais.  
GUIDE\_E.PDF                Guide du Fichier Maître en anglais.

DOC\PDF\_F\  
DD\_F.PDF                    Dictionnaire de données en français.  
INDEX\_A\_F.PDF              Index alphabétique des variables en français.  
INDEX\_T\_F.PDF              Index des variables par sujet en français.  
LAYOUT\_F.PDF               Cliché d'enregistrement en français.  
QUEST\_F.PDF                Questionnaire en français.  
QUEST\_HHLD\_F.PDF          Questionnaire ménage en français.  
DVDOC\_F.PDF                Documentation sur les variables dérivées en français.  
GUIDE\_F.PDF                Guide du Fichier Maître en français.

*Fichiers relatifs à l'utilisation des poids bootstrap pour le calcul de la variance :*

BOOTSTRP\DATA\  
B5.TXT                      Fichier de poids bootstrap en format ASCII.

BOOTSTRP\SAS\DOC\  
SASBOOTDOC\_F.PDF          Documentation en français du programme BOOTVARF\_V30.SAS.  
SASBOOTDOC\_E.PDF          Documentation en anglais du programme BOOTVARE\_V30.SAS.  
APPENDIXC\_HEALTH.PDF      Noms des variables et des fichiers à utiliser avec le programme BOOTVAR.  
ANNEXEC\_SANTE.PDF         Noms des variables et des fichiers à utiliser avec le programme BOOTVAR.

BOOTSTRP\SAS\PGM\

BOOTVARF_V30.SAS	Programme SAS de calcul de la variance utilisant le bootstrap, avec commentaires en français.
BOOTVARE_V30.SAS	Programme SAS de calcul de la variance utilisant le bootstrap, avec commentaires en anglais.
MACROF_V30.SAS	Programme de macros SAS avec commentaires en français.
MACROE_V30.SAS	Programme de macros SAS avec commentaires en anglais.
ETAPE1.SAS	Exemple, avec commentaires en français, du programme SAS créant le fichier d'analyse.
STEP1.SAS	Exemple, avec commentaires en anglais, du programme SAS créant le fichier d'analyse.

BOOTSTRP\SPSS\DOC\

SPSSBOOTDOC_F.PDF	Documentation en français du programme BOOTVARF_V30.SPS.
SPSSBOOTDOC_E.PDF	Documentation en anglais du programme BOOTVARE_V30.SPS.
APPENDIXC_HEALTH.PDF	Noms des variables et des fichiers à utiliser avec le programme BOOTVAR.
ANNEXEC_SANTE.PDF	Noms des variables et des fichiers à utiliser avec le programme BOOTVAR.

BOOTSTRP\SPSS\PGM\

BOOTVARF_V30.SPS	Programme SPSS de calcul de la variance utilisant le bootstrap, avec commentaires en français.
BOOTVARE_V30.SPS	Programme SPSS de calcul de la variance utilisant le bootstrap, avec commentaires en anglais.
MACROF_V30.SPS	Programme de macros SPSS avec commentaires en français.
MACROE_V30.SPS	Programme de macros SPSS avec commentaires en anglais.
ETAPE1.SPS	Exemple, avec commentaires en français, du programme SPSS créant le fichier d'analyse.
STEP1.SPS	Exemple, avec commentaires en anglais, du programme SPSS créant le fichier d'analyse.

BOOTSTRP\LAYOUT\

B5_I.SAS	Cliché d'article SAS pour le fichier de poids bootstrap B5.TXT
B5_I.SPS	Cliché d'article SPSS pour le fichier de poids bootstrap B5.TXT