

Enquête nationale sur la santé de la population
Composante des ménages, 2000-2001
Fichier fictif pour le téléaccès au fichier maître

Juillet 2002

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 téléaccès. 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 nationale sur la santé de la population (ENSP), menée par la Division de la statistique de la santé à Statistique Canada, recueille de l'information sur la santé et les caractéristiques socio-économiques de la population canadienne. À ce jour, quatre cycles de collecte ont été complétés : l'ENSP Cycle 1 (1994-1995), l'ENSP Cycle 2 (1996-1997), l'ENSP Cycle 3 (1998-1999) et l'ENSP Cycle 4 (2000-2001). Les quatre cycles contiennent une composante des ménages et une composante des établissements de santé. Les cycles 1, 2 et 3 ont aussi une composante du Nord.

Les cycles précédents ont produit des fichiers de microdonnées à grande diffusion (FMGD) pour la composante des ménages (cycles 1, 2 et 3), ainsi que pour la composante des établissements de santé (cycles 1, 2) afin de permettre un accès élargi aux données. Un seul fichier a été créé pour la composante des ménages du cycle 4. Il s'agit du fichier longitudinal carré qui contient tous les 17 276 membres du panel. La création d'un FMGD nécessite l'application de procédures rigoureuses afin d'assurer la confidentialité des données. Ainsi, certaines des variables de l'enquête doivent être soit agrégées, soit assujetties à des valeurs maximales ou minimales, soit simplement supprimées. Puisque les règles de confidentialité limitent la création de FMGD de nature longitudinale et que l'ENSP est purement longitudinale commençant au cycle 4, un FMGD n'a pas été créé pour le cycle 4.

Afin d'offrir aux chercheurs la possibilité d'avoir accès à ces fichiers, un service de téléaccès a été instauré. Ce service permet au chercheur de soumettre des programmes informatiques par courriel à une adresse électronique dédiée (nphs-ensp@statcan.ca) et d'en recevoir les résultats par retour de courriel. Pour se prévaloir du service de téléaccès, 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 aux fichiers maîtres plutôt que les FMGD (pour les cycles où un FMGD existe)
- pourquoi le service de téléaccès 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'ENSP à l'adresse électronique mentionnée précédemment ou par téléphone au (613) 951-1653. Lorsque l'approbation pour le téléaccès a été obtenue, le chercheur peut soumettre ses programmes informatiques à l'équipe de l'ENSP pour le traitement des données du (des) fichier(s) maître(s). 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 longitudinal fictif du cycle 4 porte sur les données recueillies au cycle 4 auprès des 17 276 membres du panel longitudinal. Bien que certaines variables à caractère administratif, ne représentant aucun intérêt analytique, ont été recodées à « blanc » ou « 9 », le fichier fictif simule parfaitement le fichier maître. Il renferme les mêmes variables et a le même cliché d'article. Cependant, les valeurs des variables du fichier fictif ont été modifiées afin de préserver la confidentialité des répondants.

Le CD-ROM contient également des programmes SAS et SPSS pour le calcul de la variance des estimations. Trois fichiers fictifs de poids bootstrap associés au fichier fictif sont aussi inclus afin de permettre au chercheur de développer et vérifier ses programmes de calcul de variance. Ces fichiers fictifs de poids bootstrap simulent le fichier original, utilisent les mêmes clichés d'article et contiennent les mêmes variables que celui-ci, mais les valeurs des poids sont modifiées par rapport au fichier maître. 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 les fichiers maîtres. 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 d'enregistrements

Les enregistrements du fichier maître sont d'abord classés selon l'âge, le sexe et le profil de réponse longitudinale. 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 cohérentes.

Les catégories utilisées pour créer le fichier longitudinal fictif du cycle 4 se fondent sur l'âge au moment de l'interview en 1994-1995, en 1996-1997, en 1998-1999 et en 2000-2001. Afin d'assurer un nombre minimum d'enregistrements dans chaque catégorie, on ne tient pas compte des caractéristiques géographiques pour la classification des enregistrements.

Le contenu des questions supplémentaires, posées en Alberta et au Manitoba en 1994-1995, et en Alberta en 1996-1997, est rajouté aux enregistrements du fichier fictif à une étape ultérieure. À ce stade, le problème des catégories trop petites est moindre, puisque le nombre de catégories selon l'âge établi pour le contenu spécial est assez faible. Lorsque le problème se pose malgré tout, on y remédie en utilisant des classes de taille suffisamment large et en corrigeant toute incohérence ayant pu être introduite.

2.2 Blocs de variables

La deuxième 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 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.

Le principe est de former des blocs ayant une signification analytique tout en étant suffisamment petits pour être conformes aux normes de confidentialité des données. Les variables qui, regroupées, permettent d'identifier des personnes sont donc placées dans des blocs différents. Toutes les modalités de chaque variable ne se trouvent pas nécessairement sur le fichier fictif. Pour obtenir de l'information sur toutes les modalités possibles pour une variable, les utilisateurs sont priés de se référer à la documentation du fichier maître. Les autres détails relatifs à la création des blocs 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. Bien que les efforts soient faits afin d'assurer la cohérence entre les blocs, cette cohérence n'est pas nécessairement toujours présente.

2.3 Variables recodées à « blanc » ou « 9 »

Pour fins de confidentialité, les variables énumérées ci-après ont été recodées à « blanc » ou « 9 » lors de la création du fichier fictif :

DOD, SP34_CPS, SP34_MET, STRATUM, REPLICAT, CYCLE, SUBCYCLE, AM68_SHA, AM34_SRC, AM34_LNG, SP34_CPA, AM54_BMM, AM54_BDD, AM54_BYE, AM54_SRC, AM54_LNG, AM54_EMM, AM54_EDD, AM54_EYY, AM64_STA, AM64_BMM, AM64_BDD, AM64_BYE, AM64_SRC, AM64FR, AM64_LNG, AM64_EMM, AM64_EDD, AM64_EYY, SP36FOUT, SP36_CPA, AM36_SRC, AM36_LNG, SP36_STA, AM56_STA, AM56_BMM, AM56_BDD, AM56_BYE, AM56_SRC, AM56_LNG, AM56_EMM, AM56_EDD, AM56_EYY, AM66LDUR, AM66_STA, AM66_BMM, AM66_BDD, AM66_BYE, AM66_SRC, AM66FR, AM66_LNG, AM66_EMM, AM66_EDD, AM66_EYY, AM68LDUR, SP38FOUT, SP38_CPA, AM38_LP, AM38_SRC, AM38_LNG, SP38_STA, AM58_STA, AM68_STA, AM58_BDD, AM58_BMM, AM58_BYE, AM58_LNG, AM58_EMM, AM58_EDD, AM58_EYY, AM68_BMM, AM68_BDD, AM68_BYE, AM68_SRC, AM68FR, AM68_LNK, AM68_LNG, AM68_EMM, AM68_EDD, AM68_EYY, AM60LDUR, SP30FOUT, SP30_CPA, AM30_PL, SP30_STA, AM30_TEL, AM60_STA, AM60_BMM, AM60_BDD, AM60_BYE, AM60_SRC, AM60FR, AM60_LNK, AM60_SHA, AM60_TEL, AM60_LNG

2.4 Sous-échantillonnage

Finalement, dans un dernier effort afin de préserver la confidentialité, un sous-échantillonnage du fichier a été fait. Le sous-échantillonnage consiste à enlever une proportion des 17 276 enregistrements, et de doubler une partie des enregistrements qui restent afin de garder le nombre d'enregistrements égal à 17 276. Le nombre de répondants dans le sous-ensemble complet et le sous-ensemble complet C1 et C4 ont aussi été respectés dans ce sous-échantillonnage.

Les poids d'échantillonnage et les poids bootstrap ont ensuite été ajustés conformément aux estimations de la population de 1994 par province et groupe d'âge-sexe. Cependant, il est à noter que ces poids demeurent des poids artificiels.

3. Contenu du CD-ROM

LisezMoi.pdf	Documentation en français sur le télé-accès, la création du fichier fictif, le contenu du CD-ROM.
ReadMe.pdf	Documentation en anglais sur le télé-accès, la création du fichier fictif, le contenu du CD-ROM.

Fichier fictif :

DATA\ Dumylong.txt	Fichier fictif du cycle 4, 17 276 enregistrements
-----------------------	---

Clichés d'article, énoncés :

LAYOUT\ LONG_fmt.sas LONG_i.sas LONG_lbf.sas LONG_lbe.sas LONG_pff.sas LONG_pfe.sas LONG_i.sps LONGmiss.sps LONGvalf.sps LONGvale.sps LONGvarf.sps LONGvare.sps	Cycle 4 SAS FORMAT Cycle 4 SAS INFILE et INPUT Cycle 4 SAS LABEL en français Cycle 4 SAS LABEL en anglais Cycle 4 SAS PROC FORMAT en français Cycle 4 SAS PROC FORMAT en anglais Cycle 4 SPSS DATA LIST Cycle 4 SPSS MISSING VALUES Cycle 4 SPSS VALUE LABELS en français Cycle 4 SPSS VALUE LABELS en anglais Cycle 4 SPSS VARIABLE LABELS en français Cycle 4 SPSS VARIABLE LABELS en anglais
---	--

Documentation:

DOC\PDF_F\ Dd_lngf_f.pdf Dd_lngfe_f.pdf Dd_long_f.pdf Dv_f.pdf Index_a_f.pdf Index_t_f.pdf Layout_f.pdf Long00_f.pdf Quest_f.pdf	Dictionnaire de données en français, fichier maître longitudinal réponses complètes du cycle 4 Dictionnaire de données en français, fichier maître longitudinal réponses complètes au C1 et C4 Dictionnaire de données en français, fichier maître longitudinal du cycle 4 Documentation des variables dérivées en français Index alphabétique en français, fichier maître longitudinal du cycle 4 Index par sujet en français, fichier maître longitudinal du cycle 4 Cliché d'article en français, fichier maître longitudinal du cycle 4 Documentation en français du volet ménage du cycle 4 Questionnaire du cycle 4 en français
DOC\PDF_E\ Dd_lngf_e.pdf Dd_lngfe_e.pdf Dd_long_e.pdf Dv_e.pdf Index_a_e.pdf Index_t_e.pdf Layout_e.pdf	Dictionnaire de données en anglais, fichier maître longitudinal réponses complètes du cycle 4 Dictionnaire de données en anglais, fichier maître longitudinal réponses complètes au C1 et C4 Dictionnaire de données en anglais, fichier maître longitudinal du cycle 4 Documentation des variables dérivées en anglais Index alphabétique en anglais, fichier maître longitudinal du cycle 4 Index par sujet en anglais, fichier maître longitudinal du cycle 4 Cliché d'article en anglais, fichier maître longitudinal du cycle 4

Long00_e.pdf	Documentation en anglais du volet ménage du cycle 4
Quest_e.pdf	Questionnaire du cycle 4 en anglais

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

BOOTSTRP\	
Lisez_moi.txt	Documentation en français des fichiers relatifs à l'utilisation du programme de variance bootstrap
Read_me.txt	Documentation en anglais des fichiers relatifs à l'utilisation du programme de variance bootstrap
BOOTSTRP\DATA\	
Bd5long.txt	Fichier fictif de poids bootstrap carrés en format ASCII
Bd5lngf.txt	Fichier fictif de poids bootstrap réponses complètes en format ASCII
Bd5lngfe.txt	Fichier fictif de poids bootstrap réponses complètes au C1 et C4 en format ASCII
BOOTSTRP\SAS\DOC\	
SASBootdoc_f.pdf	Documentation en français du programme BOOTVARF_V20.SAS
SASBootdoc_e.pdf	Documentation en anglais du programme BOOTVARE_V20.SAS.
BOOTSTRP\SAS\PGM\	
Bootvarf_v20.sas	Programme SAS de calcul de la variance utilisant le bootstrap, avec commentaires en français
Bootvare_v20.sas	Programme SAS de calcul de la variance utilisant le bootstrap, avec commentaires en anglais
Etape1.sas	Exemple, avec commentaires en français, du programme SAS créant le fichier d'analyse
Macrof_v20.sas	Programme de macros SAS avec commentaires en français
Macroe_v20.sas	Programme de macros SAS avec commentaires en anglais
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_V21.SPS
SPSSBootdoc_e.pdf	Documentation en anglais du programme BOOTVARE_V21.SPS
BOOTSTRP\SPSS\PGM\	
Bootvarf_v21.sps	Programme SPSS de calcul de la variance utilisant le bootstrap, avec commentaires en français
Bootvare_v21.sps	Programme SPSS de calcul de la variance utilisant le bootstrap, avec commentaires en anglais
Etape1.sps	Exemple, avec commentaires en français, du programme SPSS créant le fichier d'analyse
Macrof_v21.sps	Programme de macros SPSS avec commentaires en français
Macroe_v21.sps	Programme de macros SPSS avec commentaires en anglais
Step1.sps	Exemple, avec commentaires en anglais, du programme SPSS créant le fichier d'analyse
BOOTSTRP\LAYOUT\	
b5long_i.sas	Cliché d'article SAS pour les fichiers fictifs de poids bootstrap BD5LONG.TXT, BD5LNGF.TXT, and BD5LNGFE.TXT
b5long_i.sps	Cliché d'article SPSS pour les fichiers fictifs de poids bootstrap BD5LONG.TXT, BD5LNGF.TXT, and BD5LNGFE.TXT