

Enquête nationale sur la santé de la population (ENSP)
Composante des ménages de 1996-1997
Téléaccès
Documentation sur les fichiers principaux fictifs

1. Contexte et vue d'ensemble

Le Programme de l'Enquête nationale sur la santé de la population comprend trois volets. La composante des ménages couvre les ménages des dix provinces, la composante des établissements de santé couvre les personnes faisant un séjour prolongé dans un établissement de soins et la composante du Nord couvre les ménages des trois territoires du Nord. Par souci de diffuser les données à grande échelle, on produit des fichiers de microdonnées à grande diffusion pour les deux premières composantes, mais non pour la troisième. Toutefois, la transformation d'un fichier principal en fichier à grande diffusion oblige à prendre des mesures pour assurer le respect des normes de confidentialité, si bien que les données détaillées sont souvent agrégées, assujetties à une valeur plafond ou supprimées. Par conséquent, les analystes qui se servent des produits à grande diffusion constatent souvent en cours d'étude qu'ils ont besoin de données plus détaillées pour achever leurs travaux. Dans le cas de l'ENSP, la fonction de téléaccès est une des options offertes à ces analystes pour leur permettre de consulter les fichiers principaux originaux.

La fonction de téléaccès permet aux utilisateurs de soumettre des programmes par la voie d'Internet en les transmettant par courrier électronique à l'adresse nphs@statcan.ca. Tout utilisateur qui souhaite se prévaloir de cette option doit obtenir l'approbation préalable de la Division de la statistique de la santé en envoyant par Internet, à l'adresse susmentionnée, un résumé décrivant ses travaux ou en appelant au (613) 951-1653 pour obtenir plus de renseignements sur le processus. Pour réaliser un essai logiciel des programmes personnalisés, il est conseillé aux clients de tester leurs applications en se servant du ou des fichiers principaux fictifs appropriés créés par Statistique Canada aux fins de diffusion. Un fichier fictif est une simulation de fichier principal qui possède la même disposition et contient des variables de même nom, mais des données modifiées, de façon à protéger l'anonymat des répondants. Une fois qu'ils fonctionnent correctement, les programmes peuvent être transmis par courrier électronique aux employés du programme de l'ENSP qui les appliqueront aux fichiers principaux pertinents. La responsabilité de l'exactitude logique du programme incombe uniquement au client. Le personnel de Statistique Canada ne vérifie les données de sortie que pour s'assurer qu'elles répondent aux normes de confidentialité, puis les envoie au client.

Pour 1996-1997, il existe quatre fichiers fictifs, à savoir les deux fichiers transversaux représentant les fichiers principaux Général (H35) et Santé (H356) produits antérieurement pour les données de 1994-1995, ainsi que deux nouveaux fichiers longitudinaux. Le premier fichier longitudinal fictif (LNGF) représente le fichier principal longitudinal des réponses complètes des deux années d'enquête et le deuxième (LNGP) représente le fichier principal longitudinal composé de toutes les combinaisons de réponses tant partielles que complètes des deux cycles. Pour le cycle de 1996-1997,

étant donné la taille des fichiers principaux originaux, les fichiers fictifs ne sont que des sous-échantillons du nombre original d'enregistrements. Le fichier Général fictif comprend un échantillon de 5 % des enregistrements du fichier principal Général (échantillonnage au 1/20) et les trois autres fichiers fictifs comprennent des échantillons de 10 % (échantillonnage au 1/10).

On a supprimé certaines variables administratives qui ne présentent aucun intérêt d'un point de vue analytique de tous les enregistrements et remplacé les codes par un « blanc » ou par des « 9 ».

Les fichiers contenant les dictionnaires de données d'accompagnement diffusés sur le CD-ROM sont ceux établis pour les fichiers principaux correspondants et contiennent des chiffres pondérés et des chiffres d'échantillon pour chaque valeur d'une variable. Ces fichiers peuvent être consultés au moyen du logiciel Adobe Acrobat installé sur le CD-ROM du produit de microdonnées à grande diffusion. Le contenu de chaque fichier est décrit plus bas.

Il convient de souligner que le calcul d'estimations de la population totale d'après ces fichiers produira des résultats dont la valeur ne correspond qu'à 5 % (fichier Général) ou 10 % (fichier Santé, LNGF et LNGP) environ de celle des résultats obtenus en se servant des fichiers principaux réels. Toute estimation ne doit être calculée qu'à des fins d'illustration, afin de montrer que le programme fonctionne correctement. Aucune ne devrait être utilisée à des fins de totalisation ou d'analyse.

2. Méthode de création des fichiers fictifs

2.1 Création de catégories d'enregistrements

Pour commencer, on a classé les enregistrements de chaque fichier principal selon l'âge, le sexe et les caractéristiques géographiques. Dans le cas des fichiers longitudinaux, la classification est en outre fonction du profil de réponse longitudinale. Pour le fichier Santé transversal, la classification fait aussi la distinction entre l'échantillon sélectionné par CA (achat d'unités supplémentaires par les provinces) et l'échantillon principal pour l'Alberta, ainsi que pour les enfants au Manitoba, puisque nombre de questions n'ont pas été posées aux membres de l'échantillon sélectionnés par CA en Alberta et que certaines questions n'ont été posées qu'aux enfants sélectionnés par CA au Manitoba et en Alberta. L'un des objectifs consistait à former des catégories regroupant des enregistrements pour lesquels le cheminement à travers le questionnaire était le même, de sorte que, après permutation aléatoire des données à l'intérieur des catégories, chaque enregistrement artificiel résultant contienne des éléments de données cohérents.

Dans le cas des fichiers longitudinaux fictifs, les catégories sont légèrement différentes. Puisque, pour ces fichiers, les catégories se fondent sur l'âge au moment de l'interview de 1994-1995 ainsi que de 1996-1997, le nombre de catégories est devenu très élevé. Afin de ne garder qu'un nombre aussi faible que possible d'enregistrements du fichier principal dans chaque catégorie, on ne s'est pas fondé sur les caractéristiques géographiques pour effectuer la classification des enregistrements.

Des questions spéciales supplémentaires ont été posées en Alberta et au Manitoba en 1994-1995, et en Alberta en 1996-1997. Pour contourner ce problème, on a commencé par créer un fichier fictif ne renfermant pas ce contenu spécial, puis on a ajouté le contenu spécial aux enregistrements du fichier fictif provenant de l'Alberta et du Manitoba à une étape ultérieure. À ce stade-là, le problème des catégories trop petites ne se posait plus, puisque le nombre de catégories selon l'âge établi pour le contenu spécial était assez faible.

2.2 Création de blocs de variables

La deuxième étape a consisté à créer des blocs de variables. On a produit les enregistrements artificiels devant figurer dans le fichier fictif en procédant à la permutation aléatoire de blocs de variables entre les enregistrements d'une même catégorie. On a d'abord regroupé les variables en blocs élémentaires correspondant aux diverses sections du questionnaire, puis on a regroupés ces blocs élémentaires pour former les blocs de variables. Dans le cas des sections dont les questions ont été posées ou non selon les réponses aux questions d'une section antérieure, on a intégré dans un même bloc les blocs élémentaires de variables correspondant à cette section, afin d'assurer la cohérence interne des enregistrements. Seules font exception à cette règle les variables utilisées pour former les catégories, puisque leurs effets sur le cheminement au long du questionnaire avaient déjà été pris en compte.

Le principe directeur était 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, permettaient d'identifier des personnes ont donc été placées dans des blocs différents.

Pour créer les blocs, on a réparti les variables entre les blocs élémentaires suivants (outre un certain nombre de variables géographiques et d'établissement du plan de sondage, de caractéristiques des personnes et des ménages, ainsi que de variables relatives à l'administration de l'enquête) :

- | | |
|--|---|
| 1. Revenu du ménage | 12. Accès aux services — Tension artérielle (1996 uniquement) |
| 2. Incapacité au cours des deux dernières semaines | 13. Accès à — Test Pap (1996 uniquement) |
| 3. Utilisation des services de santé | 14. Accès à — Mammographie (1996 uniquement) |
| 4. Limitation des activités | 15. Accès à — Examen des seins (1996 uniquement) |
| 5. Problèmes de santé chroniques | 16. Accès à — Examen général (1996 uniquement) |
| 6. Caractéristiques socio-démographiques | 17. Accès à — Vaccination antigrippale (1996 uniquement) |
| 7. Niveau de scolarité | 18. Accès à — Visites chez le dentiste (1996 uniquement) |
| 8. Population active | |
| 9. État de santé général | |
| 10. Taille/poids | |
| 11. Mesures préventives en santé (1994 uniquement) | |

19. Accès à — Examen des yeux (1996 uniquement)
20. Accès à — Services d'urgence (1996 uniquement)
21. VIH (1996 uniquement)
22. Renseignements sur la santé (Alberta, 1996 uniquement)
23. État de santé
24. Activité physique
25. Bronzage (Alberta, 1996 uniquement)
26. Mouvements répétitifs (1996 uniquement)
27. Blessures
28. Consommation de médicaments
29. Usage du tabac
30. Consommation d'alcool
- 30a. Consommation d'alcool (Alberta, 1996 uniquement)
31. Stress (1994 uniquement)
32. Événements récents de la vie (1994 uniquement)
33. Traumatismes (1994 uniquement)
34. Stress au travail (1994 uniquement)
35. Estime de soi et contrôle (1994 uniquement)
36. Sentiment de cohérence (1994 uniquement)
37. Santé mentale
38. Soutien social
- 38a. Soutien social (Alberta, 1996 uniquement)
39. Opinions au sujet des parents (Alberta, 1996 uniquement)
40. Services de santé (Alberta, 1996 uniquement)
41. Santé sexuelle (1996 uniquement)
- 41a. Santé sexuelle (Alberta, 1996 uniquement)
42. Sécurité routière (1996 uniquement)
43. Violence/sécurité personnelle (Alberta, 1996 uniquement)
- 44a. Sens de la débrouillardise (Alberta, 1994)
- 44b. Sens de la débrouillardise (Alberta, 1996)
45. Rapport avec les gens (Manitoba, 1994 uniquement)

On notera que les blocs élémentaires n^{os} 22, 25, 39, 40, 43, 44 et 45 renferment le contenu spécial pour le Manitoba et l'Alberta. Ce contenu a été ajouté à une troisième étape et, en général, les catégories étaient fondées sur l'âge, le sexe et le profil de réponse longitudinale.

2.3 Variables supprimées

Les variables énumérées ci-après ont été supprimées durant la création des fichiers fictifs.

Fichier principal Général fictif :

AM56_RNO, AM36_PAR, AM36_LP, AM36_LNG, AM56_LNG, AM56_SHA, AM66_SHA, AM36_SRC, AM56_SRC, AM56_CAS, AM56_BMM, AM56_BDD, AM56_BYY, AM56_EMM, AM56_EDD, AM56_EYY, SP36_CPS, SP36_TYP, SP36DPC, CYCLE, SUBCYCLE, MEMCYCLE, WT56_C, WT56_R.

Fichier principal Santé fictif :

Toutes les variables énumérées sous le titre « Fichier Général », sauf AM56_RNO, WT56_C et WT56_R;

AM66_RNO, AM66_LNG, AM66_CAS, AM66_RC, AM66_SRC, AM66_BMM, AM66_BDD, AM66_BYY, AM66_EMM, AM66_EDD, AM66_EYY, AM66_LNK, AM66_HNC, AM66_5, WT66_C, WT66_R, WT66_NR.

Fichiers principaux longitudinaux fictifs :

Toutes les variables énumérées sous le titre « Fichier Santé », sauf AM66_RNO, WT66_C, WT66_R, WT66_NR;

AM34_LP, AM34_SRC, AM34_LNG, SP34_CPS, SP34_TYP, SP34_MET, AM64_LNK, AM54_BMM, AM54_BDD, AM54_BYY, AM54_SRC, AM54_EMM, AM54_EDD, AM54_EYY, AM64_BMM, AM64_BDD, AM64_BYY, AM54_LNG, AM64_STA, AM64_SRC, AM64_RC, AM64_LNG, AM64_EMM, AM64_EDD, AM64_EYY, DHC6ADOD, SP34_DPC, SP36FOUT, SP36_STA, AM56_STA, AM66_STA, AM66LDUR, AM66_HN.

2.4 Fichiers figurant sur le CD_R

Fichiers ASCII :

DUMYH35.TXT	Fichier Général (H35) fictif
DUMYH356.TXT	Fichier Santé (H356) fictif
DUMYLNFG.TXT	Réponses longitudinales complètes, fichier fictif
DUMYLNFGP.TXT	Réponses longitudinales partielles, fichier fictif

Fichier des dictionnaires de données :

Général — H35 :

DICGEN_F.PDF	Liste des dictionnaires de données en français
DICGEN_E.PDF	Liste des dictionnaires de données en anglais
INDGEN_F.PDF	Index alphabétique des variables en français
INDGEN_E.PDF	Index alphabétique des variables en anglais
LAYGEN_F.PDF	Cliché d'enregistrement en français
LAYGEN_E.PDF	Cliché d'enregistrement en anglais
TOPGEN_F.PDF	Index thématique des variables en français
TOPGEN_E.PDF	Index thématique des variables en anglais

Santé — H356 :

DICHLT_F.PDF	Liste des dictionnaires de données en français
DICHLT_E.PDF	Liste des dictionnaires de données en anglais
INDHLT_F.PDF	Index alphabétique des variables en français
INDHLT_E.PDF	Index alphabétique des variables en anglais
LAYHLT_F.PDF	Cliché d'enregistrement en français
LAYHLT_E.PDF	Cliché d'enregistrement en anglais
TOPHLT_F.PDF	Index thématique des variables en français
TOPHLT_E.PDF	Index thématique des variables en anglais

Réponses longitudinales complètes - LNFG

DICLNF_F.PDF	Liste des dictionnaires de données en français
DICLNF_E.PDF	Liste des dictionnaires de données en anglais
INDLNF_F.PDF	Index alphabétique des variables en français
INDLNF_E.PDF	Index alphabétique des variables en anglais
LAYLNF_F.PDF	Cliché d'enregistrement en français
LAYLNF_E.PDF	Cliché d'enregistrement en anglais
TOPLNF_F.PDF	Index thématique des variables en français
TOPLNF_E.PDF	Index thématique des variables en anglais

Réponses longitudinales partielles - LNGP

DICLNP_F.PDF	Liste des dictionnaires de données en français
DICLNP_E.PDF	Liste des dictionnaires de données en anglais
INDLNP_F.PDF	Index alphabétique des variables en français
INDLNP_E.PDF	Index alphabétique des variables en anglais
LAYLNP_F.PDF	Cliché d'enregistrement en français
LAYLNP_E.PDF	Cliché d'enregistrement en anglais
TOPLNP_F.PDF	Index thématique des variables en français
TOPLNP_E.PDF	Index thématique des variables en anglais

Fichiers périphériques SAS :

H35_IN.SAS	H35 cliché d'enregistrement d'entrée
H35_FMT.SAS	H35 énoncé format
H35_LBF.SAS	H35 étiquettes de variable en français (nom + concept en français (49 caractères))
H35_LBE.SAS	H35 étiquettes de variable en anglais (nom + concept en anglais (49 caractères))
H35_OUT.SAS	H35 cliché d'enregistrement de sortie
H35_PFF.SAS	H35 format proc en français
H35_PFE.SAS	H35 format proc en anglais
H356_IN.SAS	H356 cliché d'enregistrement d'entrée
H356_FMT.SAS	H356 énoncé format
H356_LBF.SAS	H356 étiquettes de variable en français (nom + concept en français (49 caractères))
H356_LBE.SAS	H356 étiquettes de variable en anglais (nom + concept en anglais (49 caractères))
H356_OUT.SAS	H356 cliché d'enregistrement de sortie
H356_PFF.SAS	H356 format proc en français
H356_PFE.SAS	H356 format proc en anglais
LNGF_IN.SAS	LNGF cliché d'enregistrement d'entrée
LNGF_FMT.SAS	LNGF énoncé format
LNGF_LBF.SAS	LNGF étiquettes de variable en français (nom + concept en français (49 caractères))
LNGF_LBE.SAS	LNGF étiquettes de variable en anglais (nom + concept en anglais (49 caractères))
LNGF_OUT.SAS	LNGF cliché d'enregistrement de sortie
LNGF_PFF.SAS	LNGF format proc en français
LNGF_PFE.SAS	LNGF format proc en anglais

LNGP_IN.SAS	LNGP cliché d'enregistrement d'entrée
LNGP_FMT.SAS	LNGP énoncé format
LNGP_LBF.SAS	LNGP étiquettes de variable en français (nom + concept en français (49 caractères))
LNGP_LBE.SAS	LNGP étiquettes de variable en anglais (nom + concept en anglais (49 caractères))
LNGP_OUT.SAS	LNGP cliché d'enregistrement de sortie
LNGP_PFF.SAS	LNGP format proc en français
LNGP_PFE.SAS	LNGP format proc en anglais

Fichiers périphériques SPSS :

H35_IN.SPS	H35 cliché d'enregistrement d'entrée
H35VALF.SPS	H35 codes de valeur en français
H35VALE.SPS	H35 codes de valeur en anglais
H35VARF.SPS	H35 étiquettes de variable en français (nom + concept en français (49 caractères))
H35VARE.SPS	H35 étiquettes de variable en anglais (nom + concept en anglais (49 caractères.))
H356_IN.SPS	H356 cliché d'enregistrement d'entrée
H356VALF.SPS	H356 codes de valeur en français
H356VALE.SPS	H356 codes de valeur en anglais
H356VARF.SPS	H356 étiquettes de variable en français (nom + concept en français (49 caractères))
H356VARE.SPS	H356 étiquettes de variable en anglais (nom + concept en anglais (49 caractères))
LNGF_IN.SPS	LNGF cliché d'enregistrement d'entrée
LNGFVALF.SPS	LNGF codes de valeur en français
LNGFVALE.SPS	LNGF codes de valeur en anglais
LNGFVARF.SPS	LNGF étiquettes de variable en français (nom + concept en français (49 caractères))
LNGFVARE.SPS	LNGF étiquettes de variable en anglais (nom + concept en anglais (49 caractères))
LNGP_IN.SPS	LNGP cliché d'enregistrement d'entrée
LNGPVALF.SPS	LNGP codes de valeur en français
LNGPVALE.SPS	LNGP codes de valeur en anglais
LNGPVARF.SPS	LNGP étiquettes de variable en français (nom + concept en français (49 caractères))
LNGPVARE.SPS	LNGP étiquettes de variable en anglais (nom + concept en anglais (49 caractères))

Programme de variance utilisant le bootstrap

Poids bootstrap utilisés par le programme BOOTVAR.SAS:

B1H35.TXT	Poids bootstrap pour le fichier dummyH35.TXT
B5H356.TXT	Poids bootstrap pour le fichier dummyH356.TXT
B5H356A.TXT	Poids bootstrap pour le fichier dummyH356.TXT, Questions de l'Enquête de la promotion sur la santé
B5H356C.TXT	Poids bootstrap pour le fichier dummyH356.TXT, Questions sur les Services de santé pour enfants
B5LONGF.TXT	Poids bootstrap pour le fichier dummyLNGF.TXT
B5LONGP.TXT	Poids bootstrap pour le fichier dummyLNGP.TXT

Cliché d'article:

BOOT_IN.SAS

Documentation :

BTDOCFR.PDF	Documentation en français du programme BOOTVAR.SAS
BTDOCENG.PDF	Documentation en anglais du programme BOOTVAR.SAS
BOOTVARF.SAS	Programme de variance utilisant le Bootstrap avec commentaires en français
BOOTVAR.SAS	Programme de variance utilisant le Bootstrap avec commentaires en anglais
BOOTSASF.SAS	Documentation en français du programme BOOTSASF.SAS
BOOTSAS.SAS	Documentation en anglais du programme BOOTSAS.SAS
EX1A_FR.SAS	Exemple 1A (Commentaires en français)
EX1A_ENG.SAS	Exemple 1A (Commentaires en anglais)
EX1B_FR.SAS	Exemple 1B (Commentaires en français)
EX1B_ENG.SAS	Exemple 1B (Commentaires en anglais)
EX1C_FR.SAS	Exemple 1C (Commentaires en français)
EX1C_ENG.SAS	Exemple 1C (Commentaires en anglais)
EX2A_FR.SAS	Exemple 2A (Commentaires en français)
EX2A_ENG.SAS	Exemple 2A (Commentaires en anglais)
EX2B_FR.SAS	Exemple 2B (Commentaires en français)
EX2B_ENG.SAS	Exemple 2B (Commentaires en anglais)
EX2C_FR.SAS	Exemple 2C (Commentaires en français)
EX2C_ENG.SAS	Exemple 2C (Commentaires en anglais)
EX2BTSTF.SAS	Exemple 4 (Commentaires en français)
EX2BTSTE.SAS	Exemple 4 (Commentaires en anglais)