

## MICRODATA FILE - 1994 INCOME

### HOUSEHOLD INCOME (1994), FACILITIES AND EQUIPMENT (1995)

#### SAMPLE DESIGN AND IMPLICATIONS

This HIFE data set has been created by linking data collected from the same households for several surveys. These surveys are:

Household Facilities and Equipment Survey, (HFE)  
May 1995  
Labour Force Survey, (LFS) April 1995  
Survey of Consumer Finances, (SCF) April 1995

Rent Survey, April 1995

The HFE and HES Surveys were supplements to the May 1995 LFS and the other three surveys were conducted simultaneously in April 1995. The sample employed for these surveys is the Labour Force Survey sample, which is a multi-stage stratified clustered probability sample.

Following each decennial census, the sample design of the Labour Force Survey is revised to incorporate recent developments in sampling methodology and to reflect population changes indicated by the most recent census data. The latest redesigned sample was implemented in January 1985, to replace the design in use since 1975.

For information on the redesign of the Labour Force Survey sample, see the feature articles published in **The Labour Force, October 1994, December 1994 and January 1995**, Catalogue No. 71-001 and the publication **Methodology of the Canadian Labour Force Survey, 1984-1990**, Catalogue No. 71-526.

Briefly, this sample is designed to represent approximately 98% of the population. Excluded population groups are:

1. Residents of the Yukon and Northwest Territories;
2. Residents of Indian reserves;
3. Residents of military barracks; and
4. Inmates of institutions such as prisons, penitentiaries, jails, reformatories, mental hospitals, tuberculosis hospitals, sanatoria, orphanages and homes for the aged.

## FICHIER DE MICRODONNÉES - REVENU DE 1994

### ÉQUIPEMENT MÉNAGER (1995) SELON LE REVENU DE 1994

#### PLAN DE SONDAGE ET IMPLICATIONS

Ce fichier RMEM de microdonnées découle du raccordement pour pluri-s enquêtes suivantes sur les mêmes ménages:

Enquête sur l'équipement ménager, (EÉM) mai 1995

Enquête sur la population active, (EPA) avril 1995  
Enquête sur les finances des consommateurs, (EFC) avril 1995  
Enquête sur les loyers, avril 1995

L'EÉM et l'EME ont été effectuées en même temps que l'EPA de mai 1995 tandis que les trois autres enquêtes ont été effectuées simultanément en avril 1995. L'échantillon utilisé pour ces enquêtes est celui de l'Enquête sur la population active qui consiste en un échantillon aléatoire à plusieurs degrés stratifié par grappes.

Après chaque recensement décennal, le plan de sondage de l'Enquête sur la population active est révisé pour y incorporer les dernières modifications aux méthodes d'échantillonnage et pour tenir compte des changements démographiques mis à jour par les données du recensement le plus récent. Le nouveau plan de sondage a été introduit en janvier 1985 pour remplacer celui utilisé depuis 1975.

Pour obtenir des renseignements sur le remaniement du plan de sondage de l'enquête sur la population active, voir l'article du mois paru dans **La population active, octobre 1994, décembre 1994 et janvier 1995**, n° 71-001 au catalogue et la publication **Méthodologie de l'enquête sur la population active au Canada, 1984-1990**, n° 71-526 au catalogue.

En somme, cet échantillon est conçu de manière à représenter environ 98% de la population. Sont exclus les groupes suivants:

1. Les résidents du Yukon et des Territoires du Nord-Ouest;
2. Les résidents des réserves indiennes;
3. Les résidents de baraquements militaires; et
4. Les pensionnaires d'institutions comme les prisons, les pénitenciers, les maisons de correction, les hôpitaux pour malades mentaux, les hôpitaux pour tuberculeux, les sanatoriums, les orphelinats et les foyers pour personnes âgées.

## **MICRODATA FILE - 1994 INCOME**

### **HOUSEHOLD INCOME (1994), FACILITIES AND EQUIPMENT (1995)**

#### **SAMPLE DESIGN AND IMPLICATIONS - continued**

In addition, households living in collective-type dwellings such as hotels, large lodging houses, clubs, logging and construction camps are excluded from this file.

Because of varying sampling and response ratios each record on the file has a weight attached to it. This weighting factor reflects the sample design and incorporates the inverse of the sampling ratio (which varies significantly by geographic area) and differential response rates for households, among other things. For further information on the weighting procedure see Catalogue No. 13-207, and "Final Universal Weight" of this documentation as well as Catalogues Nos. 71-526, 71-528 and the issues of Catalogue No. 71-001 mentioned above.

The sample design and weighting procedure have important implications for data use. Each variable on the file must be multiplied by the weight in order to produce valid statistical estimates of population parameters such as means, medians, modes, estimated numbers and aggregates. However, the appropriateness of using these sample survey weights in regression analysis is another issue. A useful summary of this issue is presented in DuMouchel and Duncan (1983) "Using Sample Survey Weights in Multiple Regression Analysis of Stratified Samples". Journal of American Statistical Association Volume 78, December 1983, pp 535-543.

## **FICHIER DE MICRODONNÉES - REVENU DE 1994**

### **ÉQUIPEMENT MÉNAGER (1995) SELON LE REVENU DE 1994**

#### **PLAN DE SONDAGE ET IMPLICATIONS - suite**

De plus, les ménages demeurant dans des logements collectifs tels que des hôtels, de grandes pensions, des clubs, des camps de bûcherons et des baraquements de travailleurs ne sont pas compris dans ce fichier.

En raison de la variation des fractions de sondage et des taux de réponse, nous avons attribué un poids à chaque dossier du fichier. Ce coefficient de pondération tient compte du plan de sondage et comporte l'inverse de la fraction de sondage (qui varie beaucoup d'une région à l'autre) et, entre autres, différents taux de réponse pour les ménages. Pour obtenir de plus amples renseignements sur la méthode de pondération, voir la publication n° 13-207 au catalogue ainsi que les publications n°s 71-526, 71-528 au catalogue et les parutions de n° 71-001 au catalogue cité plus haut. Et la rubrique "Pondération universelle finale" de la présente documentation.

Le plan de sondage et la méthode de pondération comportent d'importantes implications à l'égard de l'utilisation des données. Il faut multiplier chaque variable du fichier par le poids pour obtenir des estimations statistiques valides des paramètres démographiques, comme des moyennes, des médianes, des modes et des estimations d'agrégats. Quoi qu'il en soit, l'intérêt que présente l'utilisation de ces poids de sondage dans l'analyse de régression est un autre sujet. L'article intitulé "Using Sample Survey Weights in Multiple Regression Analysis of Stratified Samples", de Dumouchel et Duncan (1983), Journal of the American Statistical Association, Volume 78, December 1983, pp. 535-543, présente un sommaire utile sur la question.

## MICRODATA FILE - 1994 INCOME

### HOUSEHOLD INCOME (1994), FACILITIES AND EQUIPMENT (1995)

#### SAMPLE DESIGN AND IMPLICATIONS - concluded

The information necessary to estimate the standard errors is not given on this file. Most regression programs are not capable of considering design effects, even if the relevant information were available. However, generalized programs such as SAS, BMDP and SPSS produce coefficients using weighted least squares. The standard errors calculated by SPSS are much too low because the sum of the weights is used for the degrees of freedom. This can be corrected by dividing the weight on each record by the sum of the weights and multiplying by the sample size. This transformation will produce standard errors equivalent to those produced by SAS and BMDP. Specific requests to produce standard errors of estimates can also be handled on a cost-recoverable basis.

Secondly, standard errors calculated for these data using the formula for data from simple random samples may be very misleading. Sampling errors calculated from this sample design are generally higher than those using the sample random formula. As an illustration of their magnitude, see the standard error symbols published in Catalogue No. 64-202, as well as the standard error of average incomes (which estimate the sampling error) published in Catalogue No. 13-218. For information on standard errors of more complex analytic statistics (i.e. regression estimates) see the uncatalogued paper "**An Investigation of Variance Estimation in Family Expenditure Surveys**" and associated bibliography.

For further information on sampling and non-sampling errors associated with the survey, see Catalogue Nos. 13-218, 64-202 and 13-207.

## FICHIER DE MICRODONNÉES - REVENU DE 1994

### ÉQUIPEMENT MÉNAGER (1995) SELON LE REVENU DE 1994

#### PLAN DE SONDAGE ET IMPLICATIONS - fin

Les renseignements dont nous avons besoin pour estimer les erreurs types ne figurent pas dans ce fichier. La plupart des programmes de régression ne peuvent pas prendre en considération les effets du plan de sondage, même si les renseignements pertinents sont connus. Toutefois, les programmes généraux comme le SAS, le BMDP et le PSSS produisent des coefficients au moyen des moindres carrés pondérés. Les erreurs types calculées d'après le PSSS sont beaucoup trop faibles parce qu'on y utilise la somme des poids pour établir les degrés de liberté. On peut remédier à la situation en divisant le poids de chaque dossier par la somme des poids et en multipliant par la taille de l'échantillon. On obtiendra ainsi des erreurs types équivalentes à celles que produisent le SAS et le BMDP. Nous pourrions répondre à ceux qui nous demanderont de produire des erreurs types particulières et ce, moyennant le recouvrement des frais.

Deuxièmement, les erreurs types calculées pour ces données au moyen de la formule s'appliquant aux données d'échantillons aléatoires simples peuvent être très trompeuses. Les erreurs d'échantillonnage calculées à partir de ce plan de sondage-ci sont généralement plus élevées que celles qui sont calculées à l'aide de la formule de l'échantillon aléatoire simple. Pour en avoir une idée, voir les indicateurs littéraux de l'erreur type publiés dans le catalogue n° 64-202, de même que les erreurs types des revenus moyens (qui estiment l'erreur d'échantillonnage) publiées dans la publication n° 13-218 au catalogue. Pour obtenir des renseignements sur les erreurs types de statistique analytiques plus complexes (par exemple, des estimations de régression), voir le document hors catalogue intitulé "**L'Estimation de la variance dans les enquêtes sur les dépenses des familles**" et la bibliographie qui s'y rattache.

Pour de plus amples renseignements sur les erreurs d'échantillonnage et les erreurs non dues à l'échantillonnage dans l'enquête, voir les publications n°s 13-218, 64-202, 13-207 au catalogue.

## **MICRODATA FILE - 1994 INCOME**

### **HOUSEHOLD INCOME (1994), FACILITIES AND EQUIPMENT (1995)**

#### **PROCEDURES FOR MANUSCRIPT REVIEW**

Because of the complexities inherent in data from household surveys, users of these data are advised to consult with the Household Surveys Division prior to publication of the data. A manuscript review service is provided by that Division to assist users. Users of microdata tapes are also advised prior to publishing their research to consult the relevant pages of the tape documentation provided regarding data reliability and the use of small samples.

The subscriber should file with Statistics Canada a copy of published research or other results based on the use of these microdata. The subscriber shall use the following form of accreditation in reference to the microdata:

"This analysis is based on the Statistics Canada microdata tape (Household Income (1994), Facilities and Equipment, 1995) which contains data collected in the (Survey of Consumer Finances, 1995 and the Survey of Household Facilities and Equipment). All computations on these microdata were done by (name of user organization) and the responsibility for the use and interpretation of these data is entirely that of the author(s)."

The user should distinguish clearly in the manuscript between data derived directly from the tape on the one hand and all added assumptions, adjustments and modifications introduced by him/her on the other. For example, true wage rates cannot be calculated from the data on the tape and the researcher must specify his assumptions (for which the survey does not provide direct evidence) that would make the averages calculated from the tape appropriate proxies for such a variable.

Manuscripts can be forwarded to:

Director  
Household Surveys Division  
Statistics Canada  
Jean Talon Building, 5th floor  
Tunney's Pasture  
Ottawa, Ontario  
K1A 0T6

## **FICHIER DE MICRODONNÉES - REVENU DE 1994**

### **ÉQUIPEMENT MÉNAGER (1995) SELON LE REVENU DE 1994**

#### **PROCÉDURES POUR FAIRE RÉVISER UN MANUSCRIT**

En raison des complexités inhérentes aux données provenant des enquêtes-ménages, il est recommandé de consulter la Division des enquêtes-ménages avant la publication de ces données. Un service de révision de manuscrits est offert par la Division pour assister les utilisateurs. Ceux qui utilisent les bandes de microdonnées pour fins de publication, devraient consulter les pages de la documentation qui traitent de la fiabilité des données et de l'usage de petits échantillons.

L'abonné devrait remettre à Statistique Canada un exemplaire de toute publication portant sur les recherches ou d'autres résultats basés sur l'utilisation des bandes de microdonnées. Dans le texte du manuscrit, il faudra indiquer la source des données de la façon suivante:

"L'analyse est basée sur la bande de microdonnées de Statistique Canada (Revenu (1994) des ménages et l'équipement ménager, 1995) qui renferme des données recueillies dans le cadre de l'enquête sur les finances des consommateurs de 1995 et de l'enquête sur l'équipement ménager de 1995. Tous les calculs ont été effectués par (nom de l'organisme), et l'auteur est seul responsable (les auteurs sont seuls responsables) de l'usage et de l'interprétation qui sont faits des données."

Dans son manuscrit, l'utilisateur devra faire clairement la distinction entre les données tirées directement de la bande et les hypothèses, rajustements et modifications qu'il a apportés. Par exemple, les données tirées de la bande ne permettent pas de calculer les taux réels de salaire; le chercheur doit préciser les hypothèses (pour lesquelles l'enquête ne renferme pas de preuves directes) qui feront des moyennes calculées à partir des données de la bande des valeurs de substitution appropriées pour cette variable.

Les manuscrits peuvent être envoyés au:

Directeur  
Division des enquêtes-ménages  
Statistique Canada  
Immeuble Jean Talon, 5<sup>e</sup> étage  
Parc Tunney  
Ottawa, Ontario  
K1A 0T6

## **MICRODATA FILE - 1994 INCOME**

### **HOUSEHOLD INCOME (1994), FACILITIES AND EQUIPMENT (1995)**

#### **ESTIMATES OBTAINED FROM MICRODATA FILE**

The following pages contain counts and aggregates obtained from this tape and are given for your use in checking computer output from the tape. On microdata files prior to the 1984 household income file (1985 facilities and equipment), households with exceptionally high incomes, large income losses from self-employment or unusual characteristics were excluded from the file. As of 1984, records with these conditions have had selected characteristics zeroed out rather than removing them from the file. Counts on the following pages include these records.

We suggest that users not release or analyse data based on a sample of less than 100 records. In fact, due to our procedure of income imputation we feel that a slightly larger minimal sample would be more appropriate. Users employing small samples may wish to consult the...

**Income and Housing Surveys Section  
Household Surveys Division  
Statistics Canada  
Phone (613) 951-4633 or (613) 951-4643**

## **FICHIER DE MICRODONNÉES - REVENU DE 1994**

### **ÉQUIPEMENT MÉNAGER (1995) SELON LE REVENU DE 1994**

#### **ESTIMATIONS PROVENANT DU FICHIER DE MICRODONNÉES**

Les comptes et agrégats provenant de cette bande vous sont fournis pour vous permettre de vérifier les sorties d'ordinateur produites à partir de cette bande. Les fichiers de microdonnées antérieurs au fichier sur le revenu des ménages de 1984 (équipement ménager de 1985) n'englobaient pas les ménages dont le revenu était exceptionnellement élevé, ceux qui ont déclaré des pertes considérables découlant d'un travail autonome, ceux présentant des caractéristiques inhabituelles. à partir de 1984, on a annulé certaines caractéristiques des enregistrements portant sur ces ménages au lieu de supprimer ces enregistrements du fichier. Les montants de contrôle aux pages suivantes comprennent ces enregistrements.

Nous recommandons aux utilisateurs de ne pas publier ni analyser des données fondées sur un échantillon de moins de 100 enregistrements. De fait, en raison de la méthode d'imputation utilisée pour le revenu, nous croyons que le nombre minimum d'enregistrements devrait être plus élevé. Les utilisateurs qui se servent de petits échantillons désireront peut-être consulter la...

**Section des enquêtes sur le revenu et le logement  
Division des enquêtes-ménages  
Statistique Canada  
Téléphone (613) 951-4633 ou (613) 951-4643**