

N° 62M0004XCB au catalogue

Guide de l'utilisateur : fichier de microdonnées à grande diffusion

Enquête sur les dépenses des ménages, 2003

Mai 2005

**Division de la statistique du revenu
Statistique Canada, Ottawa, K1A 0T6
Téléphone : 613 951-7355**

This document is available in English

La «Division de la statistique du revenu, Statistique Canada» devrait être mentionnée lorsqu'une partie quelconque du présent document est reproduite ou citée.

Table des matières

1. Introduction	3
1.1 Renseignements généraux.....	3
1.1.1 Contexte	3
1.1.2 Nouveau pour 2002	3
1.1.3 Autres documents.....	3
1.1.4 Pour de plus amples renseignements	4
1.2 Caractéristiques techniques du fichier.....	5
2. Information technique.....	6
2.1 Méthodologie de l'enquête	6
2.1.1 L'univers de l'enquête.....	6
2.1.2 Contenu de l'enquête et période de référence	7
2.1.3 L'échantillon.....	7
2.1.4 Collecte des données	8
2.1.5 Traitement des données et contrôle de la qualité.....	8
2.1.6 Pondération	8
2.2 La qualité des données	10
2.2.1 Erreur d'échantillonnage.....	10
2.2.2 Erreur non due à l'échantillonnage.....	11
2.2.3 Effet des valeurs extrêmes	13
2.2.4 Comparabilité dans le temps	14
2.3 Lignes directrices pour la totalisation, l'analyse et la diffusion	15
2.3.1 Note importante aux utilisateurs concernant les ménages constitués pour toute l'année ou pour une partie de l'année seulement	15
2.3.2 Lignes directrices pour l'arrondissement	15
2.3.3 Lignes directrices pour la pondération de l'échantillon en vue de la totalisation.....	16
2.3.4 Genres d'estimations: type nominal par opposition à quantitatives	17
2.3.5 Lignes directrices pour l'analyse statistique	19
2.3.6 Lignes directrices pour la diffusion	19
2.4 Confidentialité des microdonnées à grande diffusion.....	32

1. Introduction

1.1 Renseignements généraux

1.1.1 Contexte

Ce fichier de microdonnées à grande diffusion présente les données de l'Enquête sur les dépenses des ménages de 2003 menée de janvier à mars 2004. Les renseignements sur les habitudes de dépenses, les caractéristiques du logement et l'équipement ménager des ménages canadiens pendant l'année 2003 ont été obtenus en demandant aux gens dans les dix provinces et les trois territoires de se rappeler les dépenses engagées au cours de l'année civile précédente (pour les habitudes de dépenses) ou l'information au 31 décembre (pour les caractéristiques du logement et l'équipement ménager).

Menée depuis 1997, l'Enquête sur les dépenses des ménages tente de reprendre une bonne part du contenu de l'Enquête sur les dépenses des familles ainsi que du contenu de l'Enquête sur l'équipement ménager. Plusieurs données provenant de ces enquêtes sont comparables aux données de l'Enquête sur les dépenses des ménages. Cependant, certaines différences liées à la méthodologie, à la qualité des données et aux définitions doivent être considérées lorsqu'on veut comparer ces données. Voir la section 1.1.4 «Pour de plus amples renseignements.»

1.1.2 Nouveau pour 2002

Les variables suivantes ont été discontinuées dans le fichier de microdonnées à grande diffusion de 2003 :

- O403 – Paiements en argent et pension alimentaire à des personnes habitant au Canada
- O404 – Paiements en argent et pension alimentaire à des personnes habitant à l'étranger
- O406 – Contributions aux organismes religieux
- O407 – Contributions aux organismes de charité non religieux

Les variables suivantes ont été introduites dans le fichier de microdonnées à grande diffusion de 2003 :

- O402 – Paiements en argent et pension alimentaire
- O404 – Dons de bienfaisance

1.1.3 Autres documents

- Dictionnaire de données (spécifications des variables, ensembles de codes et autres renseignements) est disponible en format pdf—EDM2003-PUMDF-Dictionnaire de données.

- Cliché d'enregistrement est disponible en format Excel—EDM2003-PUMDF-Cliché d'enregistrements.
- Les annexes sont disponibles en format Excel—EDM2003-PUMDF-Annexes.
 - L'annexe A présente les comptes de fréquence des variables non monétaires du fichier de microdonnées à grande diffusion. Ils sont inclus pour vous aider à vérifier vos totalisations.
 - L'annexe B présente les données sur les dépenses en utilisant le fichier de microdonnées à grande diffusion et également en utilisant la base de données interne de l'enquête. Ils sont inclus pour vous aider à vérifier vos totalisations.
 - L'annexe C contient un tableau indiquant les variables de dépense présentées dans les fichiers de microdonnées à grande diffusion précédents de l'Enquête sur les dépenses des ménages et l'Enquête sur les dépenses des familles.
 - L'annexe D présente tout changement dans les variables comparativement à l'année précédente.
 - L'annexe E présente les coefficients de variation pour les données publiées provenant de l'EDM de 2003.

1.1.4 Pour de plus amples renseignements

De l'information additionnelle au sujet de l'EDM peut maintenant être obtenue gratuitement sur le site Web de Statistique Canada (www.statcan.ca). Voir particulièrement:

- Note aux anciens utilisateurs de données provenant de l'Enquête sur les dépenses des familles (62F0026MIF2000002)
- Note aux anciens utilisateurs de données provenant de l'Enquête sur l'équipement ménager (62F0026MIF2000003)
- Guide de l'utilisateur de l'Enquête sur les dépenses des ménages de 2003 (62F0026MIF2004003)
- La méthodologie de l'Enquête sur les dépenses des ménages (62F0026MIF2001003)
- Enquête sur les dépenses des ménages de 2002 — Indicateurs de qualité des données (62F0026MIF2004001)

Pour obtenir plus de renseignements au sujet des résultats actuels de l'enquête et des produits et services connexes, ou pour s'informer sur les concepts, les méthodes ou la qualité des données de l'Enquête sur les dépenses des ménages, communiquez avec les Services aux clients (613-951-7355; 1-888-

297-7355; télécopieur : 613-951-3012; revenu@statcan.ca), Division de la statistique du revenu.

1.2 Caractéristiques techniques du fichier

Contenu : Les habitudes de dépenses des ménages, les caractéristiques des logements et l'équipement ménager, 2003

Source : Enquête sur les dépenses des ménages de 2003
Division de la statistique du revenu
Statistique Canada

Ensemble de données :

Nom de l'ensemble de données	SHS2003.TXT
Nombre d'enregistrements	17 265
Structure	
Taille de l'enregistrement	2 064

2. Information technique

2.1 Méthodologie de l'enquête

(Pour de plus amples renseignements, veuillez consulter "La méthodologie de l'Enquête sur les dépenses des ménages" disponible gratuitement sur le site Internet de Statistique Canada au www.statcan.ca.)

2.1.1 L'univers de l'enquête

L'Enquête sur les dépenses des ménages de 2003 a été effectuée dans les ménages privés des dix provinces et les trois territoires¹ du Canada.

Les groupes suivants ont été exclus de l'enquête :

- les personnes qui vivent dans les réserves indiennes et les terres de la couronne (à l'exception des Territoires);
- les représentants officiels de pays étrangers qui vivent au Canada et leur famille;
- les membres d'ordres religieux et d'autres groupes vivant en communauté;
- les membres des Forces armées canadiennes vivant dans des camps militaires;
- les personnes vivant dans des résidences pour personnes âgées; et
- les personnes qui vivent à plein temps dans les institutions : par exemple, les détenus des pénitenciers et les malades chroniques qui vivent dans des hôpitaux et dans des établissements de soins de longue durée.

L'enquête couvre environ 98 % de la population dans les dix provinces. Dans les territoires, la couverture se limite à 88 % au Yukon, 92 % dans les Territoires du Nord-Ouest et 89 % au Nunavut.

Aucuns renseignements n'ont été recueillis auprès des personnes qui vivent temporairement loin de leur famille (p. ex. les étudiants universitaires) parce que l'information est obtenue auprès de leurs familles, si les familles sont choisies dans l'échantillon. On peut ainsi éviter de compter deux fois ces personnes.

Les données qui proviennent des ménages constitués pendant une partie de l'année seulement sont exclues des estimations des dépenses des ménages. Cependant, ces données sont incluses dans les estimations des caractéristiques des logements et dans le calcul du taux de réponse de l'Enquête sur les dépenses des ménages. Les données qui proviennent des ménages constitués pendant une partie de l'année seulement servent également à la production des comptes nationaux de Statistique Canada. Les ménages constitués pendant une partie de l'année seulement sont des ménages qui comprennent uniquement des personnes qui étaient membres d'autres ménages pour une partie de l'année de

1. Dans le but de réduire le fardeau de réponse pour les ménages du nord, l'EDM est menée dans le nord seulement tous les deux ans, débutant en 2001

l'enquête. Dans l'échantillon de 2003, il y avait 570 ménages constitués pendant une partie de l'année seulement.

2.1.2 Contenu de l'enquête et période de référence

Des renseignements détaillés au sujet des dépenses pour les produits et services de consommation, les variations de l'actif, les hypothèques et d'autres prêts, et le revenu annuel ont été recueillis. Ces renseignements furent recueillis pour l'année civile 2003 (l'année de référence de l'enquête). Nous avons aussi recueilli des renseignements sur les caractéristiques des logements (p. ex. le type et l'âge de l'appareil de chauffage) et de l'équipement ménager (p. ex. les appareils ménagers, le matériel de communications et les véhicules). Ces derniers renseignements furent recueillis en date du 31 décembre de l'année de référence.

Comme l'Enquête sur les dépenses des ménages est conçue principalement pour donner des renseignements détaillés sur les dépenses autres que les dépenses d'alimentation, seule une estimation globale de ces dépenses est indiquée. On peut trouver des renseignements détaillés sur ces dépenses dans l'Enquête sur les dépenses alimentaires, qui est menée à tous les quatre à six ans. Elle a été faite la dernière fois en 2001. En février 2003, les résultats ont été publiés dans *Dépenses alimentaires au Canada, 2001*, n° 62-554-XIF au catalogue.

2.1.3 L'échantillon

La taille de l'échantillon pour l'Enquête sur les dépenses des ménages de 2003 a été de 23 869 ménages admissibles.

Cette année, un échantillon supplémentaire d'environ 2 200 logements au Québec a été financé par le Ministère des finances du Québec afin de mieux répondre aux besoins analytiques de celui-ci. Cet échantillon a été sélectionné d'une liste de logements selon un plan stratifié à deux degrés d'échantillonnage. La base liste, qui est construite à partir du Registre des adresses de Statistique Canada, est constitué des logements d'un certain nombre d'aires de diffusion du Recensement de 2001.

L'échantillon régulier de l'EDM a été stratifié à plusieurs degrés choisi à partir de la base de sondage de l'Enquête sur la population active (EPA). La sélection de l'échantillon s'est faite en deux étapes : la sélection de grappes (petites aires géographiques) de la base de sondage de l'EPA et la sélection de logements dans les grappes choisies. La base de sondage de l'EPA utilise principalement la géographie du recensement de 1991 et les dénombremens de la population de 1991.²

2. Consulter la publication *Méthodologie de l'Enquête sur la population active du Canada* Statistique Canada, no 71-526-XPB au catalogue, pour obtenir une description détaillée de la base de sondage de l'Enquête sur la population active.

2.1.4 Collecte des données

L'Enquête sur les dépenses des ménages de 2003 s'est déroulée au cours des mois de janvier à mars 2004. La collecte des renseignements fut effectuée par un intervieweur au moyen d'une entrevue en personne. Un questionnaire papier fut utilisé pour recueillir l'information. Une copie de ce questionnaire est disponible sur demande.

2.1.5 Traitement des données et contrôle de la qualité

L'entrée des données et la vérification automatisée pour l'Enquête sur les dépenses des ménages de 2003 se sont faites dans les bureaux régionaux de Statistique Canada. On pouvait ainsi communiquer avec les répondants lorsqu'on avait besoin de plus de renseignements pour résoudre des anomalies dans leurs questionnaires.

Après l'entrée des données, on a vérifié au moyen d'un système automatisé de vérification s'il y avait eu des erreurs dans l'entrée. Les données devaient ensuite être vérifiées par un système à deux niveaux comportant une vérification dont les résultats devaient être probants et où l'on déterminait si les réponses aux questionnaires étaient logiques et cohérentes, ainsi que des «avertissements» qui indiquaient qu'une situation particulière était anormale et devrait peut-être être corrigée. L'utilisation de l'un ou l'autre type de vérification nécessitait l'intervention d'un membre de l'une des équipes de contrôle spécialement formées à cette fin. L'ensemble des données a fait l'objet d'une autre vérification au bureau central, lorsque des réponses erronées étaient corrigées.

Les réponses manquantes ont été imputées par la méthode du plus proche voisin. Le Système canadien de contrôle et d'imputation du recensement (SCANCIR) de Statistique Canada a été utilisé pour insérer des valeurs à partir d'enregistrements donneurs ayant des caractéristiques similaires. Le choix des caractéristiques est effectué en fonction de chaque variable. Par exemple, le revenu total a été utilisé pour la plupart des variables. Le type de logement, la taille du ménage et la province ont également été fréquemment utilisés.

Pour l'Enquête sur les dépenses des ménages de 2003, la totalisation a été effectuée par un système PC/client-serveur. Ce système fournit des outils (capacités d'interrogation, de recherche et de visualisation) pour déceler les erreurs systématiques.

2.1.6 Pondération

L'estimation des caractéristiques de la population selon une enquête par échantillonnage est fondée sur le postulat que chaque unité échantillonnée représente un certain nombre d'unités de la population. Un poids de base a été associé à chaque enregistrement dans l'échantillon pour tenir compte de cette représentation. Pour l'échantillon régulier, ces poids de base ont été rajustés pour tenir compte de la non-réponse pour des régions métropolitaines choisies, pour des régions géographiques supplémentaires ainsi que pour des strates de

haut revenu. Les régions géographiques supplémentaires comprennent le reste des régions métropolitaines, ainsi que les régions urbaines et rurales selon les définitions du recensement, mais ne correspondent pas nécessairement exactement. Pour les définitions de ces termes, veuillez consulter le *Dictionnaire du recensement, 1996*, n° 92-351-XPB au catalogue.

Pour l'échantillon supplémentaire au Québec, la correction pour la non-réponse des ménages est faite à l'intérieur de groupes d'aires de diffusion du Recensement de 2001. Ces groupes ont été construits à l'aide du logiciel « Knowledge Seeker » de façon à regrouper les ménages possédant des propensions similaires de répondre à l'enquête. Ils tiennent compte de la région métropolitaine de recensement, de l'indicateur d'urbanisation et de caractéristiques de la population et des logements des aires de diffusion en date du Recensement de 2001, tels l'âge, le revenu et le type de logement. Une fois la correction pour la non-réponse terminée, un ajustement additionnel est fait aux poids des ménages de l'échantillon supplémentaire et de l'échantillon régulier du Québec. Cet ajustement est effectué afin de tenir compte du fait que l'échantillon supplémentaire du Québec et qu'une partie de l'échantillon régulier de cette province couvrent la même sous-population.

Pour accroître la fiabilité des estimations, les poids ont été rajustés de sorte que les estimations sur les caractéristiques pertinentes pour la population correspondent aux totaux de la population en provenance de sources autres que l'enquête. Pour les dix provinces, deux ensembles de totaux sont utilisés.

Le premier ensemble, pour les groupes d'âge/sexe, la taille des ménages et le type de ménage au niveau provincial, sont basés sur les projections démographiques à la mi-janvier 2004 utilisant le recensement de la population de 1996 (ajusté pour tenir compte de l'effet du sous-dénombrement net). Des totaux pour 18 groupes d'âge/sexe sont utilisés. Ils sont combinés avec des totaux du nombre de ménages d'une seule personne, de deux personnes et de plus de deux personnes. Il y a de plus des totaux pour le nombre de ménages monoparentaux et de couples avec enfants célibataires. Enfin, pour les 14 régions métropolitaines sélectionnées, seulement deux groupes d'âges sont utilisés : le nombre de personnes de moins de 18 ans et le nombre de personnes de 18 ans et plus.

À cause de leur plus petite population, on a utilisé seulement deux groupes d'âge/sexe pour les trois territoires (Yukon, Territoires du Nord-Ouest et Nunavut) : le nombre de personnes de moins de 18 ans et le nombre de personnes de 18 ans et plus, combinés avec des totaux du nombre de ménages d'une seule personne, de deux personnes et de plus de deux personnes pour le Yukon.

Le deuxième ensemble de totaux provient de l'information des relevés T4 de l'Agence des douanes et du revenu du Canada (ADRC, auparavant Revenu Canada); il permet de s'assurer que la distribution pondérée du revenu (sur la base des salaires et traitements) dans l'ensemble de données correspond à celle de la population canadienne.

L'utilisation des totaux de la population basé sur la population du Recensement de 1996 au lieu de ceux de 1991, utilisés antérieurement, et l'utilisation de l'information des relevés T4 de l'ADRC ont été introduites avec l'EDM de 1999. Les estimations révisées de l'EDM pour les enquêtes antérieures sont disponibles et devraient être utilisées pour les comparaisons d'une année à l'autre.

2.2 La qualité des données

(Pour de plus amples renseignements, veuillez consulter *Les indicateurs de la qualité des données de l'Enquête sur les dépenses des ménages*, qui sera bientôt disponible gratuitement sur le site Internet de Statistique Canada au www.statcan.ca.)

2.2.1 Erreur d'échantillonnage

Les erreurs d'échantillonnage résultent du fait que l'on tire certaines conclusions au sujet de l'ensemble de la population à partir des renseignements obtenus auprès de seulement un échantillon de cette population. L'importance de l'erreur d'échantillonnage est fonction du plan de sondage, de la variabilité des données et de la taille de l'échantillon. De plus, pour un plan d'échantillonnage donné, on obtiendra des erreurs d'échantillonnage différentes pour différentes méthodes d'estimation.

Le plan retenu pour l'Enquête sur les dépenses des ménages de 2003 était un échantillonnage stratifié à plusieurs degrés. Dans un échantillonnage à plusieurs degrés, les erreurs d'échantillonnage sont normalement plus importantes que dans un échantillonnage aléatoire simple de même taille. Toutefois, les avantages sur le plan opérationnel l'emportent sur cet inconvénient, et le fait que l'échantillon soit aussi stratifié améliore la précision des estimations.

La variabilité des données est la différence entre les unités de la population par rapport aux dépenses pour un article ou la présence d'une caractéristique spécifique du logement ou d'une pièce d'équipement ménager. En général, plus ces différences sont grandes, plus l'erreur d'échantillonnage de cet article est élevée. Également, plus la taille de l'échantillon est grande, plus l'erreur d'échantillonnage est faible.

2.2.1.1 Erreur type et coefficient de variation

Une mesure courante de l'erreur d'échantillonnage est l'erreur type (ET). L'erreur type est le degré de variation que l'on observe dans les estimations par suite du choix d'un échantillon particulier plutôt qu'un autre. Il a été démontré que la « vraie » valeur de la caractéristique d'intérêt se situe à l'intérieur de l'intervalle de +/- 1 erreur type de l'estimation pour 68 % de tous les échantillons, et de +/- 2 erreurs types de l'estimation pour 95 % de tous les échantillons.

Le coefficient de variation (CV) est l'erreur type exprimée en tant que pourcentage de l'estimation. Il est utilisé pour indiquer le degré d'incertitude

associé à une estimation. Par exemple, si l'estimation du nombre de ménages pour une caractéristique du logement donnée est 10 000 ménages, et que le CV correspondant est de 5 %, la « vraie » valeur se situe entre 9 500 et 10 500 ménages, 68 % du temps, et entre 9 000 et 11 000 ménages, 95 % du temps.

Les erreurs types pour l'Enquête sur les dépenses des ménages de 2003 ont été estimées selon la méthode du « bootstrap ». Cette méthode permet l'estimation de la variance de statistiques non lisses comme les quantiles. Pour plus de renseignements sur les erreurs types et les coefficients de variation, voir la publication de Statistique Canada, intitulée *Méthodologie de l'enquête sur la population active du Canada*, n° 71-526XPB au catalogue.

Les coefficients de variation pour les données publiées sont présentés aux tableaux techniques 1 et 2 de l'annexe E.

2.2.1.2 Suppression de données

Pour des raisons de fiabilité, les données pour lesquelles les CV ont été estimés à plus de 33 % devraient être supprimées. Puisque les CV ne sont pas calculés pour toutes les estimations, la suppression de données pour l'Enquête sur les dépenses des ménages a été fondée sur la relation entre le CV et le nombre de ménages qui déclaraient des dépenses pour un article. L'analyse d'enquêtes antérieures indique que les CV atteignent habituellement ce niveau lorsque le nombre de ménages qui déclarent un article chute à environ 30. Par conséquent, on a supprimé les dépenses moyennes des ménages et le pourcentage de ménages déclarants pour les articles déclarés par moins de 30 ménages.

Toutefois, les données relatives aux articles supprimés sont incluses dans les variables sommaires. Par exemple, les dépenses pour une catégorie particulière d'habillement pourraient être supprimées, mais ce montant fait partie de l'estimation des dépenses totales au titre de l'habillement.

2.2.2 Erreur non due à l'échantillonnage

Les erreurs non dues à l'échantillonnage se produisent parce qu'il est difficile, en raison de certains facteurs, d'obtenir des réponses exactes et de faire en sorte que ces réponses restent exactes à toutes les étapes du traitement. Contrairement à l'erreur d'échantillonnage, l'erreur non due à l'échantillonnage n'est pas facilement quantifiable. On peut dégager quatre sources d'erreurs non dues à l'échantillonnage : les erreurs de couverture, les erreurs de réponse, les erreurs de non-réponse et les erreurs de traitement.

2.2.2.1 Erreur de couverture

Les erreurs de couverture découlent d'une représentation inadéquate de la population observée. Ces erreurs peuvent survenir au cours de l'établissement du plan d'échantillonnage ou du tirage de l'échantillon, ou encore pendant la collecte ou le traitement des données.

2.2.2.2 Erreur de réponse

Les erreurs de réponse peuvent être attribuables à divers facteurs comme l'ambiguïté du questionnaire, une mauvaise interprétation des questions par les intervieweurs ou les répondants ou la déclaration de données inexactes par les répondants. Dans l'Enquête sur les dépenses des ménages, la différence entre les rentrées d'argent et les sorties d'argent est calculée pour vérifier la capacité de rappel des répondants. Cet important outil de contrôle de la qualité consiste en la mise en équilibre des rentrées d'argent (revenu et autres montants reçus par le ménage) et des sorties d'argent (dépenses totales et la variable *Flux monétaires — biens, prêts et autres dettes*) pour chaque questionnaire. Si la différence est supérieure à 10 % des rentrées d'argent ou des sorties d'argent, le montant le plus élevé étant retenu, on communique de nouveau avec les répondants pour obtenir des renseignements supplémentaires. De cette façon, il est garanti, au moins au niveau agrégé, que les dépenses sont proportionnées aux revenus et aux autres sources de fonds du ménage.

Plusieurs caractéristiques de l'enquête aident les répondants à se rappeler leurs dépenses de la façon la plus exacte possible. Premièrement, la période de l'enquête est l'année civile, car dans l'esprit des gens elle est probablement définie plus clairement que toute autre période de longueur semblable. Deuxièmement, les dépenses alimentaires (environ 11 % du budget moyen en 2003) peuvent être estimées comme étant des dépenses hebdomadaires ou mensuelles, selon les habitudes du répondant. Troisièmement, les dépenses pour de petits articles achetés à intervalles réguliers sont normalement estimées en fonction du montant et de la fréquence de l'achat. Les achats de gros articles (p. ex. une automobile) sont assez faciles à se rappeler, tout comme les dépenses au titre du loyer, des taxes foncières, et des versements mensuels pour emprunts hypothécaires. Toutefois, même dans le cas de ces articles, l'exactitude des données dépend de la capacité du répondant de se rappeler et de sa volonté de consulter des dossiers.

2.2.2.3 Erreur de non-réponse

Les erreurs de non-réponse sont présentes dans toutes les enquêtes-échantillons, car les répondants éventuels n'acceptent pas tous de collaborer pleinement. L'importance de cette erreur varie d'une non-réponse partielle à une non-réponse totale.

Une non-réponse totale est enregistrée dans le cas où l'intervieweur se trouve devant l'impossibilité de contacter le répondant, lorsqu'aucun membre du ménage n'est capable de fournir l'information recherchée, ou que le répondant refuse de participer à l'enquête. S'il s'agit d'une non-réponse totale, on ajuste le poids de base des ménages répondants afin de compenser pour les ménages qui n'ont pas répondu. Pour l'Enquête des dépenses des ménages de 2003, le taux global de réponse est 72,3 %. Consulter le Tableau 1 pour le taux de réponse par province et territoire.

Dans la plupart des cas, on se trouve devant une non-réponse partielle à l'enquête lorsque le répondant ne comprend pas la question ou l'interprète mal, refuse d'y répondre, ou encore n'arrive pas à se souvenir des renseignements

demandés. Ce type de non-réponse est compensé en imputant les valeurs manquantes.

L'importance de cette erreur est inconnue mais de façon générale cette erreur n'est pas négligeable lorsqu'un groupe de personnes possèdent certaines caractéristiques communes refusent de participer à l'enquête et que ces caractéristiques exercent un effet déterminant sur les résultats de l'enquête.

Tableau 1
Taux de réponse, Canada, provinces et territoires, 2003

	Ménages admissibles (1)	Non contactés	Ayant refusé de participer	Inutilisables (2)	Utilisables	Taux de réponse (3)
Terre-Neuve-et-Labrador	1 772	134	198	69	1 371	77,4 %
Île-du-Prince-Édouard	781	52	121	27	581	74,4 %
Nouvelle-Écosse	2 048	150	347	132	1 419	69,3 %
Nouveau-Brunswick	1 845	98	303	118	1 326	71,9 %
Québec	4 817	360	927	68	3 462	71,9 %
Ontario	3 149	282	787	103	1 977	62,8 %
Manitoba	1 887	104	312	49	1 422	75,4 %
Saskatchewan	1 837	72	280	77	1 408	76,6 %
Alberta	2 087	128	336	18	1 605	76,9 %
Colombie-Britannique	2 580	205	500	45	1 830	70,9 %
Yukon	418	28	74	9	307	73,4 %
Territoires du Nord-Ouest	410	21	42	8	339	82,7 %
Nunavut	238	13	7	-	218	91,6 %
Canada	23 869	1 647	4 234	723	17 265	72,3 %

(1) Les ménages constitués pour une partie de l'année sont inclus dans le calcul des taux de réponse. Il y avait 570 ménages de ce genre en 2003.

(2) Rejetées à l'étape de la vérification.

(3) Utilisables/ménages admissibles x 100

2.2.2.4 Erreur de traitement

Les erreurs de traitement surviennent lors du traitement des données, c'est-à-dire durant la saisie, la vérification, la pondération et la totalisation. Consulter la section Traitement des données et contrôle de la qualité (plus haut) pour une description des moyens utilisés pour réduire l'erreur de traitement.

2.2.3 Effet des valeurs extrêmes

Dans tout échantillon, la présence ou l'absence de valeurs extrêmes au sein de l'échantillon peut avoir une influence sur les estimations. Ces valeurs extrêmes sont plus susceptibles d'être rencontrées au sein des populations affichant une répartition asymétrique positive. Or, de par sa nature même, l'Enquête sur les dépenses des ménages se prête à l'apparition de telles valeurs extrêmes. La présence de ces valeurs influe fortement sur les estimations de totaux, de moyennes et d'erreurs types.

2.2.4 Comparabilité dans le temps

Menée depuis 1997, l'Enquête sur les dépenses des ménages tente de reprendre une bonne part du contenu de l'Enquête sur les dépenses des familles ainsi que du contenu de l'Enquête sur l'équipement ménager. De nombreuses variables provenant de ces enquêtes sont comparables à celles de l'Enquête des dépenses des ménages. Avant d'établir des comparaisons, on doit cependant tenir compte de certaines différences reliées à la méthodologie, à la qualité des données et aux définitions.

Pour plus d'information, veuillez vous référer à la *Note aux anciens utilisateurs de données provenant de l'Enquête sur les dépenses des familles*, n° 62F0026MIF2000002 au catalogue et la *Note aux anciens utilisateurs de données provenant de l'Enquête sur l'équipement ménager* n° 62F0026MIF2000003 au catalogue. Ces deux documents sont disponibles gratuitement, sur le site Internet de Statistique Canada (www.statcan.ca).

Les données historiques des enquêtes sur les dépenses des ménages de 1997 et de 1998, de l'Enquête sur les dépenses des familles de 1996 et de l'Enquête sur l'équipement ménager de 1996 ont été pondérées de nouveau au moyen de la méthode de pondération décrite à la section « Pondération ». Les comparaisons historiques entre les données de ces enquêtes et les données des dernières années de l'Enquête sur les dépenses des ménages devraient, en général, se faire à partir de données repondérées, même si les différences entre les estimations de l'enquête établies au moyen de l'ancienne méthode et de la nouvelle méthode semblent être minimales au niveau sommaire. Cependant, certaines populations ou variables sont susceptibles d'être plus fortement touchées que d'autres.

Débutant avec l'Enquête sur les dépenses des ménages de 1997, on a retranché des dépenses d'entretien, réparations et modifications payés par le locataire et des primes d'assurances des locataires le montant correspondant au pourcentage du loyer attribué à des fins commerciales. Ceci peut avoir un effet sur toutes comparaisons aux données antérieures.

Pour l'année de référence de 2001, on a ajouté des questions supplémentaires pour que les données de l'enquête puissent servir à la pondération de l'Indice des prix à la consommation. Ce changement peut nuire aux comparaisons historiques à l'égard de certaines variables. Par exemple, on a ajouté des questions à la rubrique « Produits de soins personnels » afin de recueillir des renseignements additionnels sur les produits de soins, le maquillage, les parfums, les désodorisants et les produits pour l'hygiène buccale. Ainsi, il se peut que les répondants aient donné des renseignements plus précis et la hausse relative à l'estimation pour les dépenses en « Produits de soins personnels » en 2001 peut avoir été attribué à la capacité accrue qu'ont les répondants de se souvenir de certains détails. L'effet des questions supplémentaires sur les estimations est difficile à quantifier. Cependant, en 2002 lorsque les questions additionnelles ont été enlevées, l'estimation pour les dépenses en soins personnels a diminué de nouveau.

Le calcul des versements hypothécaires a été révisé pour les années 1997 à 2001. Cela a eu pour effet d'amener des révisions à la baisse pour les catégories

de dépenses «Versements hypothécaires périodiques», «Logements appartenant à l'occupant», «Résidence principale», «Logement», «Consommation courante totale», et «Dépenses totales». La révision a produit une diminution moyenne d'environ 0,5 % dans la catégorie des «Dépenses totales» pour ces années.

2.3 Lignes directrices pour la totalisation, l'analyse et la diffusion

Cette section décrit les lignes directrices que doivent suivre les utilisateurs qui totalisent, analysent, publient ou diffusent des données tirées du fichier de microdonnées à grande diffusion.

2.3.1 Note importante aux utilisateurs concernant les ménages constitués pour toute l'année ou pour une partie de l'année seulement

En 1997, l'Enquête sur les dépenses des familles (EDFAM) et l'Enquête sur l'équipement ménager (EÉM) ont été remplacées par l'Enquête sur les dépenses des ménages (EDM). Les fichiers de microdonnées de l'EDFAM n'incluaient que les ménages présents l'année entière,³ car seuls ces ménages peuvent fournir une bonne indication du revenu et des dépenses sur toute l'année. Les microdonnées de l'EÉM incluaient quant à elles tous les ménages, car les données recueillies l'étaient en date du 31 décembre. Pour satisfaire les besoins des utilisateurs, tous les ménages sont présents sur le fichier de l'EDM, avec une variable indiquant le statut de chaque ménage (année entière / une partie de l'année).

Afin de créer des statistiques sur les dépenses annuelles moyennes, les utilisateurs devraient employer les enregistrements relatifs aux ménages pour l'année entière seulement. Pour mettre en tableaux des caractéristiques du logement ou de l'équipement ménager, ou pour créer d'autres types de statistiques sur les dépenses telles que les totaux (agrégats) ou la part de marché, les utilisateurs devraient utiliser les enregistrements relatifs aux ménages pour l'année entière ainsi que ceux pour une partie de l'année.

2.3.2 Lignes directrices pour l'arrondissement

Afin que les estimations calculées à partir de ce fichier de microdonnées et destinées à la publication ou à toute autre forme de diffusion correspondent à celles qui seraient obtenues par Statistique Canada, nous conseillons vivement

3. Un *ménage présent « une partie de l'année »* se compose entièrement de membres qui ne sont présents qu'une partie de l'année seulement; un *ménage pour l'année entière* est celui qui compte au moins un membre présent toute l'année. Un membre présent une partie de l'année est un membre du ménage qui a été présent pendant moins de 52 semaines. Les données sur les dépenses et le revenu des membres présents une partie de l'année ne sont recueillies que pour la portion de l'année durant laquelle ils ont fait partie du ménage.

aux utilisateurs de se conformer aux lignes directrices suivantes pour l'arrondissement de ces estimations:

- a) Les estimations dans le corps d'un tableau statistique doivent être arrondies à la centaine près au moyen de la technique d'arrondissement classique. Selon cette technique, si le premier ou le seul chiffre à supprimer se situe entre 0 et 4, le dernier chiffre à conserver ne change pas. Si le premier ou le seul chiffre à supprimer se situe entre 5 et 9, on augmente de 1 la valeur du dernier chiffre à conserver. Par exemple, lorsqu'on veut arrondir à la centaine près de la façon classique, si les deux derniers chiffres se situent entre 00 et 49, il faut les remplacer par 00 et laisser le chiffre précédent (le chiffre des centaines) tel quel. Si les deux derniers chiffres se situent entre 50 et 99, on les remplace par 00 et on augmente de 1 le chiffre précédent.
- b) Les totaux partiels marginaux et les totaux marginaux des tableaux statistiques doivent être calculés à partir de leurs éléments correspondants non arrondis, puis arrondis à leur tour à la centaine près à l'aide de la technique d'arrondissement classique.
- c) Les moyennes, les proportions, les taux et les pourcentages doivent être calculés à partir d'éléments non arrondis (c'est-à-dire les numérateurs et/ou les dénominateurs), puis arrondis à une décimale au moyen de la technique d'arrondissement classique.
- d) Les sommes et les différences d'agrégats (ou de rapports) doivent être calculées à partir de leurs éléments correspondants non arrondis, puis arrondis à leur tour à la centaine près (ou à la décimale près) en employant la technique d'arrondissement classique.
- e) Si, en raison de limitations d'ordre technique ou de toute autre nature, on utilise une autre technique que l'arrondissement classique, de sorte que les estimations à publier ou à diffuser sous une forme quelconque diffèrent des estimations correspondantes qui seraient obtenues par Statistique Canada en utilisant ce fichier de microdonnées, nous conseillons vivement aux utilisateurs d'indiquer la raison de ces divergences dans le ou les documents à publier ou à diffuser.
- f) Des estimations non arrondies ne doivent en aucun cas être publiées ou diffusées de quelque façon que ce soit par les utilisateurs. Des estimations qui ne sont pas arrondies laissent faussement supposer qu'elles sont beaucoup plus précises qu'elles ne le sont en réalité.

2.3.3 Lignes directrices pour la pondération de l'échantillon en vue de la totalisation

Le plan d'échantillonnage utilisé pour l'EDM n'est pas autopondéré, c'est-à-dire que les ménages faisant partie de l'échantillon n'ont pas tous le même poids d'échantillonnage. Pour produire des estimations simples, y compris des tableaux statistiques ordinaires, les utilisateurs doivent employer le poids d'échantillonnage approprié. Sinon, les estimations calculées à partir des

fichiers de microdonnées ne pourront être considérées comme représentatives de la population observée et ne correspondront pas à celles qui seraient obtenues par Statistique Canada en utilisant ce fichier de microdonnées. Voir la section 2.1.6 «Pondération.»

Les utilisateurs doivent également noter qu'en raison de la manière dont ils traitent le champ du poids, certains progiciels ne permettent peut-être pas d'obtenir des estimations qui correspondent exactement à celles de Statistique Canada en utilisant ce fichier de microdonnées.

2.3.4 Genres d'estimations: type nominal par opposition à quantitatives

Avant de discuter de la façon dont on peut totaliser et analyser les données de l'EDM, il est utile de décrire les deux principaux genres d'estimations qui peuvent être produites à partir du fichier de microdonnées de l'Enquête sur les dépenses des ménages.

2.3.4.1 Estimations de type nominal

Les estimations de type nominal sont des estimations du nombre ou du pourcentage de ménages dans la population visée par l'enquête qui possèdent certaines caractéristiques ou qui font partie d'une catégorie définie. Le nombre de ménages ayant déclaré une dépense particulière est un exemple d'estimation de ce genre.

Exemples de questions de type nominal :

Aviez-vous un téléphone cellulaire pour usage personnel? oui non

Quand votre logement a-t-il été originellement construit?

- en 1920 ou avant
- 1921-1945
- 1946-1960
- 1961-1970
- 1971-1980
- 1981-1990
- 1991-2000
- en 2001-2002
- en 2003

Le 31 décembre 2003, votre logement était-il:

- Possédé sans hypothèque par votre ménage?
- Possédé avec une (des) hypothèque (s) par votre ménage?
- Loué par votre ménage?
- Occupé gratuitement par votre ménage?

Totalisation d'estimations de type nominal

On peut obtenir des estimations du nombre de personnes qui possèdent une caractéristique donnée à partir du fichier de microdonnées en additionnant les

poids finals de tous les enregistrements contenant la ou les caractéristiques voulues. On obtient les proportions et les rapports de la forme X/Y comme suit:

- a) en additionnant les poids finals des enregistrements contenant la caractéristique voulue pour le numérateur X;
- b) en additionnant les poids finals des enregistrements contenant la caractéristique voulue pour le dénominateur Y;
- c) en divisant l'estimation du numérateur par celle du dénominateur.

2.3.4.2 Estimations quantitatives

Les estimations quantitatives sont des estimations de totaux ou de moyennes, de médianes ou d'autres mesures de tendance centrale de quantités basées sur tous les membres de la population observée ou sur certains d'entre eux. Elles comprennent aussi explicitement des estimations de la forme X/Y où X est une estimation de la quantité totale pour la population observée et Y est une estimation du nombre de personnes dans la population observée qui contribuent à cette quantité totale.

Un exemple d'estimation quantitative est la dépense moyenne annuelle pour les soins personnels et de santé par ménage de la population cible. Le numérateur correspond à une estimation de la dépense totale annuelle pour les soins personnels et de santé, et le dénominateur à une estimation du nombre de ménages dans la population.

Exemple de question quantitative :

En 2003, combien votre ménage a-t-il dépensé pour un service téléphonique?

Totalisation d'estimations quantitatives

On peut obtenir des estimations de quantités à partir du fichier de microdonnées en multipliant la valeur de la variable voulue par le poids final de chaque enregistrement, puis en additionnant cette quantité pour tous les enregistrements qui nous intéressent. Par exemple, pour obtenir une estimation des dépenses totales en électricité pour les ménages qui étaient propriétaires au 31 décembre, on multiplie la valeur déclarée à la question « En 2003, combien votre ménage a-t-il dépensé pour l'électricité? » par le poids final de l'enregistrement, puis on additionne ce résultat pour tous les enregistrements ayant une réponse positive à la question « Le 31 décembre 2003, votre logement était-il : 'Possédé sans hypothèque par votre ménage' ou 'Possédé avec une (des) hypothèque(s) par votre ménage' ».

Afin d'obtenir une moyenne pondérée exprimée sous la forme X/Y, le numérateur X est calculé comme une estimation quantitative et le dénominateur Y comme une estimation de type nominal. Par exemple, pour estimer les dépenses moyennes des ménages en électricité par les propriétaires, il faut:

- a) estimer la dépense totale en électricité des ménages propriétaires de leur logement, selon la méthode décrite ci-haut;

- b) estimer le nombre de ménages propriétaires en additionnant les poids finals de tous les enregistrements ayant une réponse positive à la question « Le 31 décembre 2003, votre logement était-il : 'Possédé sans hypothèque par votre ménage' ou 'Possédé avec une (des) hypothèque(s) par votre ménage'»; puis,
- c) diviser l'estimation obtenue en a) par celle calculée en b).

Nota : Puisqu'il s'agit ici d'une estimation de dépenses moyennes, les ménages « pour une partie de l'année » devraient préalablement être exclus des calculs (Pour plus de détails, voir la section 2.3.1 «Note importante aux utilisateurs concernant les ménages constitués pour toute l'année ou pour une partie de l'année seulement.»)

2.3.5 Lignes directrices pour l'analyse statistique

L'Enquête sur les dépenses des ménages est basée sur un plan complexe qui comporte une stratification et de multiples degrés de sélection, ainsi que des probabilités inégales de sélection des répondants. L'utilisation de données provenant d'enquêtes aussi complexes pose des difficultés aux analystes, car le plan d'enquête et les probabilités de sélection influent sur les méthodes d'estimation et de calcul de la variance à utiliser.

Bien que de nombreuses méthodes analytiques intégrées aux progiciels statistiques permettent d'utiliser des poids, le sens ou la définition du poids diffère de celui qui convient à une enquête par sondage. Par conséquent, si les estimations faites au moyen de ces progiciels sont exactes dans bien des cas, **les variances calculées n'ont pratiquement aucune signification.**

Dans le cas de nombreuses techniques analytiques (p. ex., la régression linéaire, la régression logistique, l'analyse de variance), il existe un moyen de rendre l'application des progiciels courants plus significative. Si l'on transforme les poids des enregistrements contenus dans le fichier de manière à ce que le poids moyen soit de (1), les résultats produits par les progiciels courants seront plus raisonnables et tiendront compte des probabilités inégales de sélection, bien qu'ils continueront à ne faire aucun cas de la stratification et de la répartition en grappes du plan d'échantillonnage. On peut effectuer cette transformation en employant dans l'analyse un poids égal au poids original divisé par la moyenne des poids originaux des unités d'échantillonnage (ménages) qui contribuent à l'estimateur en question. Cependant, parce que cette méthode ne tient pas toujours compte de la stratification et des grappes du plan d'échantillonnage, les estimations de la variance calculées de cette façon sont fort probablement des sous-estimations des vraies valeurs.

2.3.6 Lignes directrices pour la diffusion

Avant de diffuser et/ou de publier des estimations tirées du fichier de microdonnées, les utilisateurs doivent d'abord déterminer le niveau de fiabilité des estimations. La qualité des données est affectée par l'erreur d'échantillonnage et l'erreur non due à l'échantillonnage comme décrit précédemment. Cependant, le niveau de fiabilité des estimations est déterminé seulement sur la base des erreurs échantillonnales telles qu'évaluées par le

coefficient de variation (CV) comme le montre le tableau ci-dessous. En plus de calculer les CVs, les utilisateurs devraient également lire la section de ce document concernant les caractéristiques de la qualité des données.

Quel que soit le CV obtenu pour une estimation tirée de ce fichier de microdonnées, les utilisateurs devraient déterminer le nombre de répondants échantillonnés qui contribuent au calcul de l'estimation. Si ce nombre est inférieur à 30, l'estimation pondérée ne devrait pas être diffusée quelle que soit la valeur du CV pour cette estimation. En ce qui concerne les estimations pondérées fondées sur des tailles d'échantillon de 30 ou plus, les utilisateurs devraient déterminer le CV de l'estimation arrondie en suivant les lignes directrices ci-dessous.

Tableau 2
Lignes directrices relatives à la variabilité de l'échantillonnage

Type d'estimation	CV (en %)	Lignes directrices
1. Acceptable	0,0 – 16,5	Les estimations peuvent être envisagées pour la diffusion générale sans restriction. N'exige aucune notation particulière.
2. Marginal	16,6 – 33,3	Les estimations peuvent être envisagées pour la diffusion générale sans restriction mais devraient être accompagnées d'un avertissement mettant les utilisateurs subséquents en garde contre la haute variabilité d'échantillonnage associée aux estimations. De telles estimations devraient être identifiées par la lettre M (ou d'une autre façon semblable).
3. Inacceptable	Supérieur à 33,3	Statistique Canada ne recommande pas de diffuser des estimations de qualité inacceptable. Cependant, si l'utilisateur choisit de le faire, les estimations devraient être marquées par la lettre I (ou d'une autre manière semblable) et l'avertissement suivant devrait accompagner les estimations : « On informe l'utilisateur que... (précisez les données)... ne satisfont pas aux normes de qualité de Statistique Canada pour ce programme statistique. Les conclusions fondées sur ces données ne seront pas fiables et seront fort probablement non valables.»

2.3.6.1 Calcul de CV approximatifs

Afin de fournir une façon d'évaluer la qualité des estimations, Statistique Canada a produit une table des coefficients de variation (table des CV) qui s'applique aux estimations de moyennes, de proportions et de totaux obtenus de ce fichier de microdonnées à grande diffusion pour les principales variables de l'EDM par province et à l'échelle du Canada (voir l'annexe E). Le CV d'une estimation est défini comme la racine carrée de la variance de l'estimation divisée par l'estimation elle-même; il est exprimé sous forme de pourcentage. Le numérateur du CV est une mesure de l'erreur d'échantillonnage de l'estimation, appelée l'erreur-type, et il est calculé à Statistique Canada au moyen de la méthode du « bootstrap ». Cette méthode exige, entre autres, certains renseignements au sujet des strates et des grappes qui ne peuvent pas être diffusés dans les fichiers de microdonnées à grande diffusion pour des raisons de confidentialité. Afin que les utilisateurs puissent estimer des CV pour des variables qui ne sont pas incluses dans les tableaux de CV, Statistique Canada a établi un ensemble de règles pour obtenir les CV approximatifs d'une diversité d'estimations. Il faudrait noter que ces règles fournissent des CV approximatifs, donc non officiels. Toutefois, la qualité de l'approximation est assez satisfaisante, surtout dans le cas des estimations les plus fiables. Notons que la précision est moindre lorsque les domaines deviennent plus petits. Ainsi, l'utilisation de l'approximation du CV doit être utilisée avec prudence lorsque les domaines sont plus petits. Le document sur la qualité des données de l'enquête de 1997 renferme des résultats de l'évaluation de la performance de cette méthode d'approximation du CV.

Comment obtenir des CV approximatifs

Les règles suivantes devraient permettre à l'utilisateur de déterminer les coefficients de variation approximatifs pour les estimations des totaux, des moyennes ou des proportions et des différences entre ces estimations pour les sous-populations (domaines) pour lesquels le CV « Jackknife » n'est pas fourni dans les tableaux de CV.

Important : Si le nombre d'observations sur lequel se fonde une estimation est inférieur à 30, l'estimation pondérée ne devrait pas être diffusée quelle que soit la valeur du CV pour cette estimation.

Règle 1 : Approximation de CV pour des estimations de totaux (agrégats)

Toutes les étapes décrites ci-dessous doivent être suivies afin d'obtenir un CV approximatif (CVA) pour l'estimation d'un total (soit un nombre de ménages possédant certaines caractéristiques (estimation catégorique) ou un total de certaines dépenses pour tous les ménages (estimation quantitative)) pour une sous-population (domaine) d'intérêt:

- 1) Créer une variable binaire pour chaque ménage, p. ex., I, qui est égale à 1 si le ménage fait partie du domaine d'intérêt, c.-à-d. possède la caractéristique désirée et 0 sinon;
- 2) Pour l'estimation d'une variable quantitative, créer une variable Y, représentant le produit de la variable binaire I et de la variable d'intérêt. Pour l'estimation d'une variable catégorique, créer une variable Z, qui est

- égale à 1 si la variable catégorique a la valeur d'intérêt et à 0 sinon, et définir la variable Y comme le produit de I et Z;
- 3) Effectuer les étapes (4) à (9) pour chaque province séparément;
 - 4) Calculer la somme pour tous les ménages du produit du poids final (section Pondération), et Y; (notons que cette somme représente l'estimation du total pour le domaine d'intérêt dans la province traitée)
 - 5) Calculer la somme pour tous les ménages du produit du poids final et la taille du ménage;
 - 6) Diviser le résultat obtenu à l'étape (4) par le résultat obtenu à l'étape (5);
 - 7) Pour chaque ménage, multiplier le résultat obtenu à l'étape (6) par la taille du ménage;
 - 8) Pour chaque ménage, définir une variable, p. ex., E, en soustrayant le résultat obtenu à l'étape (7) de Y;
 - 9) Calculer la somme pour tous les ménages du produit du poids final moins 1, du poids final et de E au carré; (notons que cette somme représente l'estimation de la variance du total estimé à l'étape 4)
 - 10) Additionner le résultat obtenu à l'étape (9) pour chaque province;
 - 11) Le CVA est défini comme étant 100 fois la racine carrée du résultat obtenu à l'étape (10), divisée par l'estimation. L'estimation est la somme pour toutes les provinces du résultat obtenu à l'étape (4).

Plus formellement, les étapes 1 à 10 ci-dessus peuvent être obtenues au moyen de la formule suivante:

$$\sum_{p=1}^{12} \sum_{k \in S_p} (w_k - 1) w_k \left(Y_k - m_k \frac{\sum_{k \in S_p} w_k Y_k}{\sum_{k \in S_p} w_k m_k} \right)^2$$

où l'indice p correspond aux provinces, S_p est l'échantillon de répondants pour la province p , l'indice k correspond aux ménages, w_k est le poids final pour le k^e ménage, m_k est la taille du k^e ménage et Y_k est la valeur de la variable Y, définie à l'étape (2) ci-dessus, pour le k^e ménage. On notera que l'indice p , indicateur de la province, prend des valeurs allant de 1 à 12. On retrouve 12 codes de province distincts sur le fichier de microdonnées: un pour chacune des dix provinces, un pour l'ensemble des territoires et un code de province « 00 » assigné à une série d'enregistrements pour des raisons de confidentialité (Voir la section 2.4 «Confidentialité des microdonnées à grande diffusion.»)

Nota : Deux variables concernant la taille de ménage sont disponibles sur le fichier de microdonnées. Pour le calcul des CV approximatifs, la variable à utiliser pour définir la taille du ménage est « Taille du ménage au 31 décembre », et non « Taille du ménage (Nombre de personnes membres à un moment quelconque pendant l'année de référence) ».

Important : Lors de l'estimation de la variance pour un domaine donné, on ne doit pas se limiter à utiliser seulement les unités faisant partie du domaine. L'estimation de la variance doit toujours se faire en utilisant la totalité de l'échantillon. Les unités ne faisant pas partie du domaine d'intérêt n'interviennent pas lors de l'estimation ponctuelle du total, mais contribuent lors de l'estimation de la variance.

Règle 2 : Approximation de CV pour des estimations des moyennes ou des proportions

On entend par moyenne ou proportion, le rapport entre deux estimations de totaux. Pour une proportion, le numérateur est l'estimation d'un sous-ensemble du dénominateur, p. ex., la proportion des dépenses des ménages au Manitoba par rapport aux ménages canadiens. Le CV d'une estimation d'une moyenne ou d'une proportion tend à être légèrement inférieur au CV correspondant du numérateur. Le CV d'une estimation d'une moyenne ou d'une proportion peut ainsi être approximé par le CV du numérateur et la technique décrite à la règle (1) peut être utilisée.

Règle 3 : Approximation de CV pour des estimations de ratios

On entend par ratio, le rapport entre deux estimations pour lesquelles la règle (2) ne s'applique pas. Alors le CV approximatif pour tout autre type de ratio, peut être calculé avec la formule suivante :

$$CVA_R = \sqrt{CVA_N^2 + CVA_D^2}$$

où CVA_R est le CV approximatif du ratio, CVA_N est le CV approximatif du numérateur du ratio et $CVAD$ est le CV approximatif du dénominateur du ratio. La formule aura tendance à surestimer le CV si les deux estimations qui forment le ratio sont en corrélation positive, et à le sous-estimer si les deux estimations sont en corrélation négative.

Règle 4 : Approximation de CV pour des estimations de différences

Le CV approximatif d'une différence entre n'importe quelles estimations ($EST_{DIFF} = EST_1 - EST_2$) est déterminé par la formule suivante:

$$CVA_{DIFF} = \frac{\sqrt{(EST_1 CVA_1)^2 + (EST_2 CVA_2)^2}}{|EST_{DIFF}|}$$

où CVA_1 est le CV approximatif associé à EST_1 et CVA_2 est le CV approximatif associé à EST_2 . La formule aura tendance à surestimer le CV si les deux estimations impliquées dans la différence sont en corrélation positive, et à le sous-estimer si les deux estimations sont en corrélation négative.

Exemples

Dans un premier temps, les détails du calcul des CV approximatifs pour l'estimation de totaux seront présentés à l'aide d'un cas fictif. Ensuite, des cas réels d'estimation de totaux, de moyennes (ou de proportions), de ratios et de différences, basés sur les données du fichier de microdonnées, seront présentés pour permettre à l'utilisateur de vérifier ses résultats et d'ainsi s'assurer de la validité de sa démarche.

Partie 1 : Cas fictif: Détails des calculs d'une approximation de CV pour l'estimation d'un total

A) Variable quantitative

Supposons que l'on veuille estimer le total pour une variable de dépense « X » (variable quantitative), pour les ménages comprenant au moins une personne de moins de 18 ans. Pour illustrer la procédure, on utilisera un échantillon fictif (voir tableau 3) et on présentera les détails des calculs (voir tableau 4) pour chacune des onze étapes décrites ci-haut. La procédure étant appliquée indépendamment à l'intérieur de chaque province, on se limitera à décrire les calculs pour une seule d'entre elles.

Supposons l'échantillon suivant en Ontario:

Tableau 3
Exemple fictif

Données de départ						
Identificateur	Province	Année entière	Poids	Taille du ménage	Nombre d'enfants 0-17 ans	Variable d'intérêt "X"
00001	Ontario	Oui	5	3	2	30
00002	Ontario	Oui	20	5	3	0
00003	Ontario	Oui	25	2	1	20
00004	Ontario	Non	5	4	2	50
00005	Ontario	Oui	15	3	0	20
00006	Ontario	Oui	10	1	0	10
00007	Ontario	Oui	15	4	0	15

À l'étape 1, on définit le domaine d'intérêt en créant une variable binaire, qui est égale à 1 pour toutes les unités faisant partie du domaine, les ménages comprenant au moins un enfant âgé de 0 à 17 ans dans le cas présent. On procède ensuite aux étapes 2 à 9 pour obtenir l'estimation de la variance, qui nous conduira au calcul du CV. On obtient ainsi les résultats suivants :

Tableau 4
Détails des calculs pour l'approximation de CV d'un total (étapes 1 à 9)

	Étape 1	Étape 2	Étape 4	Étape 5	Étape 6	Étape 7	Étape 8	Étape 9
Ident.	Variable binaire "I"	Variable quantitative "Y"	Y pondéré	Variable "K"		Étape 6 * taille	(Y - étape 7)	(Poids - 1) * Poids * (Étape 8) ²
		(X * I)	(poids * Y)	(poids * taille)				
00001	1	30 * 1 = 30	5 * 30 = 150	5 * 3 = 15		3 * 3 = 9	30 - 9 = 21	(4 * 5 * 21 * 21) = 8 820
00002	1	0 * 1 = 0	20 * 0 = 0	20 * 5 = 100		3 * 5 = 15	0 - 15 = -15	(19 * 20 * (-15) * (-15)) = 85 500
00003	1	20 * 1 = 20	25 * 20 = 500	25 * 2 = 50		3 * 2 = 6	20 - 6 = 14	(24 * 25 * 14 * 14) = 117 600
00004	1	50 * 1 = 50	5 * 50 = 250	5 * 4 = 20		3 * 4 = 12	50 - 12 = 38	(4 * 5 * 38 * 38) = 28 880
00005	0	20 * 0 = 0	15 * 0 = 0	15 * 3 = 45		3 * 3 = 9	0 - 9 = -9	(14 * 15 * (-9) * (-9)) = 17 010
00006	0	10 * 0 = 0	10 * 0 = 0	10 * 1 = 10		3 * 1 = 3	0 - 3 = -3	(9 * 10 * (-3) * (-3)) = 810
00007	0	15 * 0 = 0	15 * 0 = 0	15 * 4 = 60		3 * 4 = 12	0 - 12 = -12	(14 * 15 * (-12) * (-12)) = 30 240
			Total: 900	Total: 300	900 / 300 = 3			Total = 288 860

Si on est intéressé au CV pour l'Ontario, on procède au calcul suivant:

$$CV_{ONT} = 100 * \frac{\sqrt{Variance_{ONT}}}{Estimation_{ONT}} = 100 * \frac{\sqrt{Étape 9_{ONT}}}{Étape 4_{ONT}} = 100 * \frac{\sqrt{288860}}{900} = 59,7$$

Si on est intéressé au CV pour le Canada, on procède de façon similaire, en totalisant les résultats pour chaque province. Autrement dit,

$$CV_{CAN} = 100 * \frac{\sqrt{Variance_{CAN}}}{Estimation_{CAN}}$$

$$= 100 * \frac{\sqrt{Variance_{T-N} + \dots + Variance_{C-B} + Variance_{Ferr} + Variance_{pPROV00}}}{Estimation_{T-N} + \dots + Estimation_{C-B} + Estimation_{Ferr} + Estimation_{pPROV00}}$$

Remarque : Dans cet exemple, on était intéressé à l'estimation du total pour la variable de dépenses « X ». Si, p. ex., on avait voulu utiliser ce total pour calculer la moyenne de dépenses « X » par ménage, l'approche aurait été légèrement différente. L'unité 00004 étant un ménage « pour une partie de l'année », elle n'aurait alors pas été considérée comme faisant partie du domaine d'intérêt, et la variable binaire « I » aurait alors eu la valeur 0. Le résultat final aurait ainsi été différent. (Pour plus de détails, voir la section 2.3.1 «Note importante aux utilisateurs concernant les ménages constitués pour toute l'année ou pour une partie de l'année seulement.»)

B) Variable qualitative (catégorique)

Il est à noter que dans le cas de l'estimation d'une variable catégorique, les étapes pour le calcul du CV approximatif seraient les mêmes que dans l'exemple

présenté pour le cas d'une variable quantitative. Au lieu d'une valeur quantitative pour la variable d'intérêt X , on créerait une variable dichotomique, qui serait égale à 1 si le ménage possède la caractéristique que l'on désire estimer, et serait égale à 0 sinon.

Pour l'estimation de variables catégoriques, différentes approches sont possibles lors de la définition du domaine et de la variable d'intérêt, procurant toutes le même résultat.

Supposons que l'on veuille estimer le nombre de ménages composés de plus d'une personne et habitant une maison individuelle. On pourrait procéder de différentes façons :

- 1) La variable binaire « I » est égale à 1 pour tous les ménages, et la variable X est égale à 1 pour les ménages composés de plus d'une personne et habitant une maison individuelle.
- 2) La variable binaire « I » est égale à 1 pour tous les ménages composé de plus d'une personne, et la variable X est égale à 1 pour tous les ménages habitant une maison individuelle.
- 3) La variable binaire « I » est égale à 1 pour tous les ménages habitant une maison individuelle, et la variable X est égale à 1 pour tous les ménages composés de plus d'une personne.
- 4) La variable binaire « I » est égale à 1 pour tous les ménages composés de plus d'une personne habitant une maison individuelle, et X est égale à 1 pour tous les ménages.

Peu importe l'approche utilisée, la variable « Y » qui en résultera (étape 2) sera égale à 1 si le ménage possède toutes les caractéristiques (plus d'une personne et habitant une maison individuelle), et sera à égale à zéro sinon. Les résultats en terme d'estimation ponctuelle et d'estimation de la variance (CV) seront donc les mêmes.

Partie 2 : Cas réels basés sur le fichier de microdonnées

Exemple 1a : Approximation de CV pour l'estimation de totaux (variable quantitative)

Supposons qu'on ait estimé que les dépenses en ameublement et équipement ménager pour les ménages composés d'une seule personne au Manitoba totalisent 85 777 298 \$. On doit estimer le CV approximatif pour cette estimation. L'utilisateur doit donc suivre les étapes (1) à (11) de la règle 1.

- 1) Il faut d'abord créer une variable binaire I dont la valeur est 1 si le ménage n'est composé que d'une seule personne et réside au Manitoba, sinon I vaut 0.
- 2) Y est défini pour chaque ménage comme étant le produit de la variable binaire I et de la variable « total des dépenses en articles et accessoires d'ameublement ».

Notons que l'estimation des dépenses en ameublement et équipement ménager est obtenue en sommant le produit de la variable Y définie en 2) et du poids final du ménage.

Le tableau 5 donne les résultats de quelques étapes du calcul du CV approximatif.

Tableau 5
Calcul du CVA

Étape	Total des dépenses en ameublement et équipement ménager pour les ménages composés d'une personne au Manitoba
4	85 777 298
5	1 066 289
6	80,44
9	8,4180 x10 ¹³
10	8,4180 x10 ¹³
11	10,70

Exemple 1b : Approximation de CV pour l'estimation de totaux (variable qualitative)

Supposons maintenant que l'on veuille estimer le nombre total de ménages canadiens composés d'une personne, ainsi que le nombre total de ménages canadiens composés d'une personne habitant différents types de logement.

Dans ce cas, la variable I est définie comme ayant la valeur 1 si le ménage est composé d'une personne et 0 sinon. Il faut créer 5 variables Z, Z1 étant une variable de compte valant 1 pour tous les enregistrements, Z2 valant 1 si le type de logement occupé est «maison individuelle» et 0 sinon, Z3 vaut 1 si le type de logement occupé est «maison jumelée» et 0 sinon, Z4 vaut 1 si le type de logement occupé est «maison en rangée» et 0 sinon et finalement Z5 vaut 1 si le type de logement occupé est «autre» et 0 sinon. Y1 est définie comme étant le produit de I et de Z1, Y2 comme le produit de I et Z2 etc.

Les estimations obtenues sont de 3 164 930 pour l'ensemble des ménages composés d'une personne, de 1 028 603 pour les maisons individuelles⁴, 114 802 pour les maisons jumelées⁵, 169 425 pour les maisons en rangée⁶ et 1 852 100 pour les autres.⁷ On veut calculer les CV approximatifs de ces estimations.

Le tableau 6 donne les résultats de quelques étapes du calcul du CV approximatif. Les résultats présentés pour les étapes 4 à 9 sont les résultats pour le Manitoba (présentés à titre d'exemple, pour une province, ils serviront à titre de comparaison dans l'exemple suivant), tandis que ceux présentés pour les étapes 10 et 11 sont les résultats au niveau du Canada.

4. Maison individuelle = individuelle non attenante

5. Maison jumelée = double

6. Maison en rangée = en rangée

7. Autre = duplex, appartement, hôtel, maison mobile, autre

Tableau 6
Calcul du CVA

Étape	Nombre de ménages composés d'une personne	Nombre de ménages composés d'une personne habitant dans une maison individuelle	Nombre de ménages composés d'une personne habitant dans une maison jumelée	Nombre de ménages composés d'une personne habitant dans une maison en rangée	Nombre de ménages composés d'une personne habitant dans d'autres logements
4	117 285	58 050	747	3194	55 294
5	1 066 289	1 066 289	1 066 289	1 066 289	1 066 289
6	0,11	0,05	0,00	0,00	0,05
9	48 310 931	22 087 229	342 917	938 852	20 415 020
10	5 675 505 959	1 694 346 875	208 459 405	280 243 854	3 070 767 506
11	2,38	4,00	12,58	9,88	2,99

Exemple 1c : Approximation de CV pour l'estimation de totaux servant au calcul d'une dépense moyenne

Supposons que l'on veuille estimer les dépenses moyennes en ameublement et équipement ménager pour les ménages composés d'une personne au Manitoba. Pour ce faire, on devra estimer le nombre de ménages composés d'une personne au Manitoba, ainsi que le total de leurs dépenses en ameublement et équipement ménager.

Puisqu'on est ici intéressé au calcul de dépenses moyennes, les « ménages pour une partie de l'année » sont exclus du domaine d'intérêt. (Pour plus de détails, voir la section 2.3.1 «Note importante aux utilisateurs concernant les ménages constitués pour toute l'année ou pour une partie de l'année seulement.») C'est pourquoi les estimations de totaux diffèrent légèrement de celles obtenues dans les deux exemples précédents.

Tableau 7
Calcul du CVA

Étape	Nombre de ménages composés d'une personne au Manitoba	Total des dépenses en ameublement et équipement ménager pour les ménages composés d'une personne au Manitoba
4	110 577	79 951 359
5	1 066 289	1 066 289
6	0,10	74,98
9	44 588 794	$7,6904 \times 10^{13}$
10	44 588 794	$7,6904 \times 10^{13}$
11	6,04	10,97

L'estimation de la moyenne serait égale à $79\,951\,359\ \$ / 110\,577 = 723\ \$$.
Comment déterminer le CV de cette estimation?

La règle (2) devrait être appliquée dans ce cas. Ainsi, le CV de cette moyenne peut donc être approximé par le CV du numérateur, soit le total des dépenses en ameublement et équipement ménager pour les ménages composés d'une personne au Manitoba. Celui-ci est 10,97%.

Exemple 2 : Approximation de CV pour l'estimation de ratios

Supposons que l'on veuille estimer le ratio entre les dépenses totales en ameublement et équipement ménager pour les ménages urbains composés d'une personne au Manitoba par rapport aux ménages ruraux.

Tableau 8
Calcul du CVA

Étape	Total des dépenses en ameublement et équipement ménager pour les ménages composés d'une personne au Manitoba (urbain)	Total des dépenses en ameublement et équipement ménager pour les ménages composés d'une personne au Manitoba (rural)
4	71 607 232	14 170 066
5	1 066 289	1 066 289
6	67,16	13,29
9	$7,4117 \times 10^{13}$	$8,7789 \times 10^{12}$
10	$7,4117 \times 10^{13}$	$8,7789 \times 10^{12}$
11	12,02	20,91

L'estimation du ratio serait égale à $71\,607\,232\ \$ / 14\,170\,066\ \$ = 5,1$ (les ménages d'une personne du milieu urbain dépensent 5 fois plus en ameublement que ceux en milieu rural). Comment déterminer le CV de cette estimation?

On a déjà calculé les CV de chacune des deux estimations impliquées dans l'estimation du ratio. On doit appliquer la règle (3) afin d'obtenir le CV voulu :

$$CVA_R = \sqrt{CVA_N^2 + CVA_D^2} = \sqrt{12,02^2 + 20,91^2} = 24,12$$

Il faut noter qu'un tel CV devrait porter la mention « Marginal » (voir la section 2.3.6 «Lignes directrices pour la diffusion») puisqu'il est assez élevé, se situant entre 16,6% et 33,3%.

Exemple 3 : Approximation de CV pour l'estimation de différences

Supposons qu'un utilisateur veuille estimer la différence entre le total des dépenses en ameublement et équipement ménager en Alberta et au Manitoba, ainsi que le CV de cette différence.

On estime le total des dépenses en ameublement et équipement ménager, ainsi que leur CV respectif, pour le Manitoba (total = 639 539 796, CV = 4,04) et pour l'Alberta (total = 2 130 282 394; CV = 3,84).

L'estimation de la différence est donc de 2 130 282 394 - 639 539 796 = 1 490 742 598. La règle (4) peut être appliquée afin d'obtenir le CV voulu :

$$CVA_{DIFF} = \frac{\sqrt{(EST_1 CVA_1)^2 + (EST_2 CVA_2)^2}}{|EST_{DIFF}|}$$

$$= \frac{\sqrt{(2\,130\,282\,394 * 3,84)^2 + (639\,539\,796 * 4,04)^2}}{|1\,490\,742\,598|} = 5,75$$

2.3.6.2 Comment obtenir les limites de confiance

Bien que les coefficients de variation soient utilisés sur une grande échelle, une mesure intuitivement plus utile de l'erreur d'échantillonnage est l'intervalle de confiance d'une estimation. Un intervalle de confiance constitue une indication du niveau de confiance selon lequel la vraie valeur de la population se situe à l'intérieur d'un intervalle déterminé de valeurs. Par exemple, un intervalle de confiance de 95% peut être décrit de la façon suivante.

Si l'échantillonnage d'une population est répété un grand nombre de fois, et que chaque échantillon mène à un nouvel intervalle de confiance pour une estimation, pour 95% des échantillons, l'intervalle couvre la vraie valeur de la population.

Au moyen du CV d'une estimation, on peut obtenir ces intervalles de confiance en supposant que dans le cas d'échantillonnage répété, les différentes estimations sont normalement réparties autour de la vraie valeur. En utilisant cette hypothèse, les probabilités sont d'environ 68 sur 100 que la différence entre une estimation d'échantillon et la vraie valeur de la population soit inférieure à un écart-type, d'environ 95 sur 100 que la différence soit inférieure à deux écarts-types, et d'environ 99 sur 100 que la différence soit inférieure à trois écarts-types. Ces différents degrés de confiance sont désignés par niveaux de confiance.

Les intervalles de confiance d'une estimation, EST, sont en général exprimés comme deux chiffres, un inférieur à l'estimation et l'autre supérieur à l'estimation, comme (EST - k, EST + k) où k est déterminé selon le niveau de confiance voulu et l'erreur d'échantillonnage de l'estimation.

Les intervalles de confiance pour une estimation peuvent être calculés d'abord en déterminant le CVA de l'estimation puis en utilisant la formule suivante afin de passer à un intervalle de confiance (IC):

$$(EST - z \times EST \times CVA / 100, EST + z \times EST \times CVA / 100)$$

où

$z = 1$ si l'on veut obtenir un intervalle de confiance de 68%,

$z = 1,6$ si l'on veut obtenir un intervalle de confiance de 90%,

$z = 2$ si l'on veut obtenir un intervalle de confiance de 95%,

$z = 3$ si l'on veut obtenir un intervalle de confiance de 99%.

Nota : Les lignes directrices relatives à la diffusion qui s'appliquent à l'estimation s'appliquent aussi à l'intervalle de confiance. Par exemple, si l'estimation ne peut pas être diffusée, l'intervalle de confiance ne peut pas être diffusé non plus.

Exemple 4

Un intervalle de confiance de 95% de la moyenne estimée des dépenses en ameublement et équipement ménager pour les ménages composés d'une personne au Manitoba pourrait être calculé de la manière suivante:

$$EST = 723$$

$$z = 2$$

$$CVA = 10,97$$

$$IC = (723 - 2 \times 723 \times 10,97/100 ; 723 + 2 \times 723 \times 10,97/100) = (564,4; 881,6)$$

2.3.6.3 Comment effectuer un test Z

Les coefficients de variation peuvent aussi être utilisés afin d'effectuer une vérification d'hypothèse, une procédure servant à distinguer entre les paramètres de la population au moyen d'estimations de l'échantillon. Les estimations des échantillons peuvent être des totaux, des moyennes, des ratios, etc. Les tests peuvent être effectués à différents niveaux de confiance lorsqu'un niveau de confiance est la probabilité de conclure que les caractéristiques sont différentes quand, en fait, elles sont identiques.

Supposons que EST_1 et EST_2 sont deux estimations d'échantillon pour deux caractéristiques d'intérêt. Supposons que le CV approximatif de la différence $EST_1 - EST_2$ soit CVA_{DIFF} .

Si $z = 100 / CVA_{DIFF}$ est plus petit que 2, aucune conclusion au sujet de la différence entre les caractéristiques n'est justifiée au niveau de confiance de 5%. Si, toutefois, ce ratio est supérieur à 2, la différence observée est significative au niveau de 5%.

Exemple 5

Supposons que l'on veuille tester, au niveau de confiance de 5%, l'hypothèse selon laquelle il n'y a pas de différence entre le total des dépenses en ameublement et équipement ménager en Alberta et le même total au Manitoba. De l'exemple 3, on a trouvé que le CV approximatif de la différence entre ces deux estimations est de 5,75, et donc $z = 17,4$. Comme cette valeur est supérieure à 2, on doit conclure qu'il y a une différence significative entre les deux estimations au niveau de confiance de 0,05.

2.4 Confidentialité des microdonnées à grande diffusion

Le fichier de microdonnées « à grande diffusion » diffère à bien des égards du fichier « principal » de l'enquête conservé par Statistique Canada. Ces écarts sont dus aux mesures prises pour protéger l'anonymat des répondants à l'enquête.

La confidentialité du fichier à grande diffusion est assurée principalement par la réduction de l'information, c'est-à-dire la suppression de variables ou de certains détails qu'elles comprennent, ou encore le regroupement de ces détails.

Pour protéger la confidentialité

- Toutes les données d'identification explicites, comme les numéros d'identification, ont été supprimées du fichier (les noms et adresses ne sont pas des données saisies).
- Le *code de la province* de 112 enregistrements a été fixé à 0 en raison de caractéristiques particulières telles des valeurs excessivement élevées. Ces enregistrements ont été pondérés de nouveau.
- Les autres enregistrements ont été pondérés de nouveau pour des fins de confidentialité.
- Les trois territoires sont regroupés sous un même code.
- Il y a eu *plafonnement* et *regroupement* des jeux de codes pour les variables non liées aux dépenses.
- Les valeurs du revenu du ménage, de la personne de référence et du (de la) conjoint(e) de la personne de référence ont été *arrondies* de la façon suivante :
 - Pour des valeurs de revenu entre 1 \$ et 9 999 \$: arrondissement aux 100 \$ les plus près
 - Pour des valeurs de revenu entre 10 000 \$ et 99 999 \$: arrondissement aux 1 000 \$ les plus près
 - Pour des valeurs de revenu entre 100 000 \$ et 999 999 \$: arrondissement aux 10 000 \$ les plus près
 - Pour des valeurs de revenu entre 1 000 000 \$ et 9 999 999 \$: arrondissement aux 100 000 \$ les plus près
 - Pour des valeurs de revenu entre 10 000 000 \$ et 99 999 999 \$: arrondissement aux 1 000 000 \$ les plus près (on ne retrouve pas de telles valeurs sur le fichier de 2003).

Les variables «Prix d'achat du logement» et «Prix de vente du logement» ont également été arrondies.