

Enquête sur la santé dans les collectivités canadiennes (ESCC)

Documentation sur les poids-ménages



Statistique
Canada

Statistics
Canada

Canada

TABLE DES MATIÈRES

1.	INTRODUCTION	1
2.	VUE D'ENSEMBLE DE LA PONDÉRATION	1
3.	PONDÉRATION DE L'ÉCHANTILLON.....	2
3.1	NON-RÉPONSE PERSONNE (I2).....	3
3.2	WINSORIZATION (I3)	3
3.3	CALAGE AUX MARGES (I4)	4
3.4	PARTICULARITÉS DE LA PONDÉRATION POUR LES TROIS TERRITOIRES	4
3.5	LA CRÉATION D'UN POIDS POUR LE FICHIER PARTAGÉ (WTS_SHH)	5
3.6	MODULE DE SÉCURITÉ ALIMENTAIRE EN 2010.....	5

1. Introduction

Ce document décrit la stratégie de pondération pour la production du poids des ménages de l'Enquête sur la santé dans les collectivités canadiennes (ESCC). En utilisant le poids des ménages, les utilisateurs doivent être conscients que l'enquête est conçue afin de représenter les individus et non pas les ménages. Le plan de sondage est établi pour s'assurer que l'échantillon est représentatif de différents groupes démographiques au niveau des personnes. Il est possible que la qualité de la représentativité de l'échantillon par rapport à certains types de ménages soit moindre. De plus, puisque le calage aux marges pour le poids des ménages est effectué au niveau des provinces, il est possible de produire des estimations fiables aux échelles nationale et provinciale seulement. D'autre part, il peut être tentant d'utiliser les réponses des individus à certaines questions comme si elles représentaient le ménage alors que ce n'est pas le cas. Les utilisateurs doivent noter que le poids des ménages devrait être utilisé seulement pour les variables où on peut clairement affirmer que les réponses de l'individu représentent le ménage et que la réponse resterait la même si n'importe quel autre individu du ménage répondait. Ceci est appuyé par le fait que, tout au long du document, on fait référence aux personnes répondantes. Les utilisateurs doivent donc se souvenir qu'il est sous-entendu que les réponses de la personne sélectionnée représentent le ménage lorsqu'on utilise le poids des ménages.

Les utilisateurs familiers avec l'ESCC remarqueront que les ajustements de poids reliés aux poids des ménages sont similaires aux ajustements des autres poids au niveau de la personne, qui étaient créés dans le passé. Une interview de l'ESCC peut être vue comme un processus en deux étapes. D'abord, l'intervieweur effectue le listage des membres du ménage (réponse de ménage). Ensuite, il/elle interviewe la personne sélectionnée dans le ménage (réponse de personne). Dans le calcul du poids des ménages, on utilise les réponses individuelles pour représenter les ménages. Cependant, les ajustements de non-réponse pour les deux étapes sont effectués lors de la production des poids des ménages, comme lors de la production des poids au niveau de la personne, parce que la non-réponse peut survenir à n'importe laquelle des étapes du processus de collecte. Les ajustements de non-réponse du ménage sont donc basés sur les caractéristiques de l'individu répondant. Étant donné que l'enquête est conçue pour recueillir les renseignements auprès de l'individu, les caractéristiques de ce dernier peuvent avoir un effet sur la non-réponse du ménage.

2. Vue d'ensemble de la pondération

Pour que les estimations produites à partir de données d'enquête soient représentatives de la population couverte, et non pas seulement représentatives de l'échantillon comme tel, l'utilisateur doit incorporer les facteurs de pondération, appelés ici les poids d'enquête, dans ses calculs. Un poids d'enquête est attribué à chaque personne incluse dans l'échantillon final, c'est-à-dire dans l'échantillon de personnes ayant répondu à l'enquête. Ce poids correspond au nombre de ménages représentés par le répondant dans l'ensemble de la population.

L'ESCC a recours à trois bases de sondage pour la sélection de son échantillon : une base aréolaire de logements agissant comme base principale, puis deux bases formées de numéros de téléphone utilisées pour compléter la base aréolaire. Puisque seulement quelques différences mineures

distinguent les deux bases de numéros de téléphone pour la pondération, elles sont traitées ensemble. On réfère à celles-ci comme faisant partie de la base téléphonique.

La stratégie de pondération traite séparément la base aréolaire et la base téléphonique. Les poids résultant de ces deux bases sont ensuite combinés en un seul ensemble de poids lors d'une étape appelée « intégration ». Suite à quelques ajustements, ce poids intégré devient le poids final. Noter que dépendamment du besoin, une seule ou deux bases peuvent être utilisées pour la sélection de l'échantillon dans une région sociosanitaire donnée. La stratégie de pondération s'occupe de cette particularité lors de l'étape d'intégration.

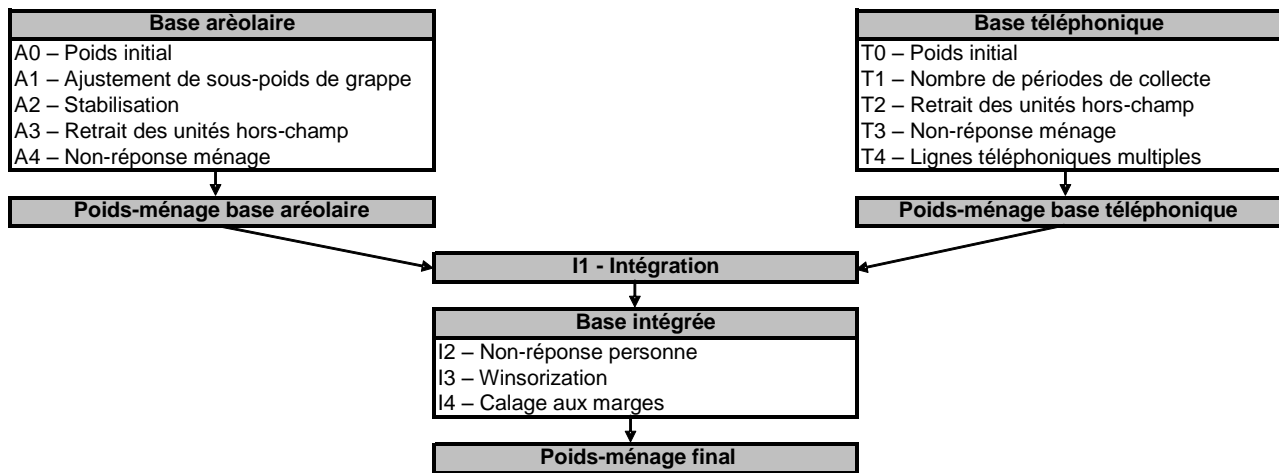
3. Pondération de l'échantillon

Tel que mentionné auparavant, les unités des bases aréolaire et téléphonique sont traitées séparément jusqu'à l'étape d'intégration. Ces étapes de pondération des ménages sont les mêmes que celles du poids principal jusqu'à l'intégration des bases (inclusivement). Veuillez faire référence au Guide de l'utilisateur de l'ESCC pour des renseignements plus complets concernant ces étapes. Les trois étapes finales de pondération, c'est-à-dire l'ajustement pour la non-réponse personnelle, la winsorization et le calage aux marges sont expliquées dans les sous-sections 3.1 à 3.3.

Malgré que les deux bases soient utilisées pour couvrir les trois territoires, certaines modifications doivent être faites relativement à leur utilisation. Ces modifications affectent substantiellement la pondération de ces trois régions et celles-ci sont rapportées dans la sous-section 3.4.

Le diagramme A présente un sommaire des différents ajustements qui font partie de la stratégie de production des poids des ménages dans l'ordre qu'ils sont appliqués. Un système de numérotation est utilisé pour identifier chaque ajustement apporté au poids et sera utilisé tout au long de la section. Les lettres A et T sont utilisées comme préfixes pour référer aux ajustements appliqués aux unités des bases aréolaire et téléphonique respectivement. Le préfixe I est, quant à lui, utilisé pour identifier les ajustements d'intégration et ceux qui suivent.

Diagramme A : Sommaire de la stratégie de pondération (Poids-ménages)



3.1 Non-réponse personne (I2)

Dans le cadre de l'ESCC, une interview peut être vue comme un processus en deux étapes. Dans un premier temps, l'intervieweur obtient la liste complète des personnes vivant dans le ménage, puis par la suite, interviewe la personne sélectionnée dans le ménage. Dans certains cas, les intervieweurs ne réussissent qu'à compléter la première étape, soit parce qu'ils ne peuvent entrer en contact avec la personne sélectionnée, ou encore parce que la personne sélectionnée refuse d'être interviewée. De tels cas sont définis comme étant des non-réponses à l'échelle de la personne et un facteur d'ajustement doit être appliqué aux poids des personnes répondantes pour compenser pour cette non-réponse. Tout comme pour la non-réponse à l'échelle du ménage (voir le Guide de l'utilisateur, section 8.2 – A4) l'ajustement est appliqué à l'intérieur des classes définies à partir des caractéristiques disponibles pour les répondants et les non-répondants. Toutes les caractéristiques recueillies lors du listage des membres du ménage, en plus des données géographiques, sont en fait disponibles pour créer ces classes. La méthode du score est utilisée afin de définir les classes. Enfin, un facteur d'ajustement est calculé à l'intérieur de chaque classe de la façon suivante :

$$\frac{\text{Somme des poids } I1 \text{ pour toutes les personnes sélectionnées}}{\text{Somme des poids } I1 \text{ pour toutes les personnes sélectionnées répondantes}}$$

Le poids I1 est donc multiplié par ce facteur d'ajustement pour produire le poids I2. Les personnes non-répondantes sont éliminées de la pondération à partir de ce point.

3.2 Winsorization (I3)

Suite à la série d'ajustements appliqués sur les poids, il est possible que certaines unités se soient retrouvées avec des poids se démarquant des autres unités dans le même domaine d'intérêt, au point même de devenir aberrantes. Pour les poids des ménages, ces domaines incluent la province par la taille du ménage où la taille du ménage est définie par : ménage d'une personne, ménage de 2 personnes et ménage d'au moins 3 personnes. Certains répondants peuvent effectivement

représenter une partie anormalement élevée de leur croisement Province-taille du ménage et ainsi influencer fortement la variance. Afin d'éviter cette situation, le poids des unités aberrantes, qui représente une grande partie de leur domaine, est ajusté à la baisse selon une méthode appelée «winsorization».

3.3 Calage aux marges (I4)

La dernière étape nécessaire afin d'obtenir le poids final des ménages de l'ESCC est le calage aux marges. Le calage aux marges est appliqué en utilisant CALMAR afin d'assurer que la somme des poids finaux corresponde aux estimations du compte de ménages définies à l'échelle des provinces par taille du ménage. Ces groupes d'intérêt sont définis par les tailles : ménage d'une personne, ménage de 2 personnes et ménage d'au moins 3 personnes. En même temps, les poids sont ajustés de façon saisonnière afin d'assurer que chaque période de collecte (période de deux mois) soit représentée de façon égale dans l'échantillon. En ce qui concerne la géographie, tout le calage aux marges est effectué au niveau provincial seulement.

Les estimations du compte de ménages sont basées sur le plus récent recensement. Les moyennes de ces estimations mensuelles sont utilisées pour le calage aux marges dans les poststrates, définies par le croisement Province-Taille du ménage, à l'intérieur d'une période de collecte. Le poids I3 est ainsi ajusté pour obtenir le poids final I4, à l'aide du facteur d'ajustement I4 défini comme suit :

$$\frac{\text{Estimation de population pour le croisement Province-Taille du ménage}}{\text{Somme des poids I3 pour le croisement Province-Taille du ménage}}$$

Par conséquent, le poids I4 correspond au **poids-ménage final de l'ESCC** que l'on retrouve dans le fichier de poids-ménage portant les noms de variable WTS_MHH pour le poids maître et WTS_SHH pour le poids partagé.

À partir de 2010, on peut noter des différences dans les estimations du nombre de ménages formés de 3 personnes ou plus, pour certaines régions. Ces différences sont dues à un changement dans les méthodes utilisées pour estimer le nombre de ménages ainsi qu'à l'utilisation des données provenant du recensement de 2006. Ces nouvelles estimations du nombre de ménages représentent une amélioration par rapport à la version précédente.

3.4 Particularités de la pondération pour les trois territoires

Le plan d'échantillonnage utilisé pour les trois territoires est quelque peu différent de celui utilisé dans les 10 provinces. La stratégie de pondération doit donc être adaptée pour répondre à ces différences. Cette section résume les changements apportés aux étapes de pondération pour les territoires.

D'abord pour la base aréolaire, une étape additionnelle de sélection est ajoutée pour les territoires. Chaque territoire est initialement stratifié selon des regroupements de communautés, à l'intérieur desquels on sélectionne aléatoirement une communauté à l'intérieur de chaque groupe. Les capitales de chaque territoire forment chacune une strate et sont donc toutes trois sélectionnées

automatiquement lors de ce premier stage. Cette particularité n'a d'effet que dans le calcul de la probabilité de sélection, et donc dans la valeur du poids initial (A0). Une fois ce poids initial calculé, la même série d'ajustements (A1 à A4) est appliquée aux unités de la base aréolaire. Les classes d'ajustement pour les ménages et les personnes sont construites de la même manière que pour les provinces, à l'aide du même ensemble de variables disponibles.

Pour ce qui est de la pondération des unités de la base téléphonique, mentionnons tout d'abord que seule la base de sondage à composition aléatoire est utilisée pour les territoires, et ce, uniquement à l'intérieur des capitales du Yukon et des Territoires du Nord-Ouest. Tous les ajustements téléphoniques standard sont appliqués.

Les deux ensembles de poids (aréolaire et téléphonique) sont ensuite intégrés, puis ajustés pour la non-réponse personnelle et la winsorization, et enfin poststratifiés de façon semblable à ce qui est fait pour les provinces, à l'exception de quatre détails. D'abord, l'intégration est appliquée uniquement pour les unités situées dans les capitales du Yukon et des Territoires du Nord-Ouest; les autres communautés sont couvertes uniquement par la base aréolaire. De plus, pour le Nunavut, les comptes de la population des ménages utilisés pour le calage aux marges représentent seulement les dix plus grandes communautés (73% des ménages) étant donné le sous-dénombrement de la base aréolaire (voir la section 5.4.1 du Guide de l'utilisateur pour plus de détails). Troisièmement, pour le Yukon et les Territoires du Nord-Ouest, le calage aux marges est utilisé, à partir des périodes de référence 2008 et 2007-2008, pour contrôler la proportion des ménages situés dans les capitales territoriales par rapport à la proportion des ménages situés à l'extérieur de ces capitales. La même approche a été adaptée pour le Nunavut à partir de 2009. Enfin, en raison des différences dans les stratégies de collecte, le nombre de périodes de collecte utilisées dans le calage aux marges pour l'effet saisonnier dans les territoires est différent de celui des provinces. Depuis 2009, deux périodes de six mois sont utilisées dans les trois territoires.

3.5 La création d'un poids pour le fichier partagé (WTS_SHH)

En plus du fichier maître et du FMGD qui contiennent tous les répondants de l'ESCC, un fichier partagé est créé qui contient seulement une portion (> 90%) des répondants originaux de l'ESCC. Les individus sur le fichier partagé ont consenti à partager leurs données avec certains partenaires. Pour compenser la perte des répondants sur le fichier, les poids des « partageurs » doivent être ajustés par le facteur :

$$\frac{\text{Somme des poids I4 pour tous les répondants}}{\text{Somme des poids I4 pour les partageurs}}$$

Pareillement à ce qui est fait pour les ajustements pour la non-réponse, ce facteur est créé à l'intérieur de groupes homogènes, formés dans le cas présent des individus ayant des propensions à partager semblables. Le poids final après cet ajustement est WTS_SHH.

3.6 Module de sécurité alimentaire en 2010

Le Module de sécurité alimentaire (FSC) est un des rares modules de l'ESCC pour lequel l'utilisation des poids-ménages peut être appropriée. En 2007-2008, le module FSC faisait partie

du contenu commun, et était donc demandé aux répondants de toutes les provinces et de tous les territoires. Pour la période de collecte 2009-2010, le module FSC fait partie du contenu optionnel et n'a pas été sélectionné pour certaines provinces. Compte tenu de ce changement, il n'est plus approprié de calculer des estimations au niveau national avec les variables de ce module, puisque les résultats ne représenteraient que les provinces et territoires pour lesquels le module a été sélectionné. Pour plus d'informations sur la sélection du contenu optionnel, veuillez référer aux annexes du Guide de l'utilisateur.