



## **Guide de l'utilisateur des microdonnées**

# **Enquête canadienne sur le don, le bénévolat et la participation**

**2004**



Statistique  
Canada

Statistics  
Canada

**Canada**



## Table des matières

<b>1.0</b>	<b>Introduction</b>	<b>5</b>
<b>2.0</b>	<b>Contexte</b>	<b>7</b>
<b>3.0</b>	<b>Objectifs</b>	<b>9</b>
<b>4.0</b>	<b>Concepts et définitions</b>	<b>11</b>
<b>5.0</b>	<b>Méthodologie de l'enquête de la composante provinciale</b>	<b>17</b>
5.1	Population visée	17
5.2	Plan de sondage	17
5.2.1	Stratification	17
5.2.2	Répartition de l'échantillon	17
5.3	Tirage de l'échantillon	18
5.4	Taille de l'échantillon selon la province	19
<b>6.0</b>	<b>Méthodologie de l'enquête de la composante territoriale (la composante du Nord)</b>	<b>21</b>
6.1	Population visée	21
6.2	Plan de sondage	21
6.2.1	Renouvellement de l'échantillon	21
6.2.2	Modifications apportées au plan de l'Enquête sur la population active pour l'Enquête canadienne sur le don, le bénévolat et la participation	22
6.3	Taille de l'échantillon	22
<b>7.0</b>	<b>Collecte des données</b>	<b>23</b>
7.1	Conception du questionnaire	23
7.2	Supervision et contrôle de qualité	23
7.3	Méthodologie de la collecte de données	23
7.3.1	Composante provinciale	23
7.3.2	Composante territoriale	24
7.4	Non-réponse	24
<b>8.0</b>	<b>Traitement des données</b>	<b>25</b>
8.1	Saisie des données	25
8.2	Vérification	25
8.3	Codage des questions ouvertes	25
8.4	Imputation	26
8.5	Création de variables dérivées	26
8.6	Pondération	27
8.7	Suppression de renseignements confidentiels	27
<b>9.0</b>	<b>Qualité des données</b>	<b>29</b>
9.1	Taux de réponse	29
9.1.1	Réponse à la composante provinciale	29
9.1.2	Réponse à la composante territoriale	30
9.2	Erreurs relatives à l'enquête	31
9.2.1	Collecte des données	31
9.2.2	Traitement des données	31
9.2.3	Non-réponse et imputation	31
9.2.4	Mesure de l'erreur d'échantillonnage	33

<b>10.0</b>	<b>Lignes directrices pour la totalisation, l'analyse et la diffusion de données</b>	<b>35</b>
10.1	Lignes directrices pour l'arrondissement d'estimations	35
10.2	Lignes directrices pour la pondération de l'échantillon en vue de la totalisation	36
10.3	Définitions de types d'estimations : catégoriques et quantitatives	36
10.3.1	Estimations catégoriques	36
10.3.2	Estimations quantitatives	36
10.3.3	Totalisation d'estimations catégoriques	37
10.3.4	Totalisation d'estimations quantitatives	37
10.4	Lignes directrices pour l'analyse statistique	38
10.5	Lignes directrices pour la diffusion de coefficients de variation	38
10.6	Seuils pour la diffusion des estimations pour l'Enquête canadienne sur le don, le bénévolat et la participation	40
<b>11.0</b>	<b>Tables de variabilité d'échantillonnage approximative</b>	<b>41</b>
11.1	Comment utiliser les tables de coefficients de variation pour les estimations catégoriques	42
11.1.1	Exemples d'utilisation des tables de coefficients de variation pour des estimations catégoriques	43
11.2	Comment utiliser les tables de coefficients de variation pour obtenir des limites de confiance	47
11.2.1	Exemple d'utilisation des tables de coefficients de variation pour obtenir des limites de confiance	48
11.3	Comment utiliser les tables de coefficients de variation pour effectuer un test t	48
11.3.1	Exemple d'utilisation des tables de coefficients de variation pour effectuer un test t	49
11.4	Coefficients de variation pour des estimations quantitatives	49
11.5	Tables des coefficients de variation	49
<b>12.0</b>	<b>Pondération</b>	<b>51</b>
12.1	Pondération de la composante provinciale	51
12.2	Pondération de la composante territoriale	55
<b>13.0</b>	<b>Questionnaires</b>	<b>57</b>
<b>14.0</b>	<b>Structure des fichiers</b>	<b>59</b>
<b>15.0</b>	<b>Désignations des variables</b>	<b>61</b>
<b>16.0</b>	<b>Cliché d'enregistrement à valeurs univariées</b>	<b>63</b>

## **1.0 Introduction**

L'Enquête canadienne sur le don, le bénévolat et la participation (ECDBP) est une composante de l'Initiative sur le secteur bénévole et communautaire, un programme de collaboration entre le gouvernement fédéral et le secteur bénévole. Statistique Canada a réalisé l'ECDBP dans les dix provinces de la mi-septembre à décembre 2004 et de la fin d'août à la mi-novembre 2004 dans les trois territoires.

Ce manuel a été produit pour faciliter la manipulation du fichier de microdonnées portant sur les résultats de l'enquête.

Toutes les questions concernant l'ensemble de données ou son utilisation devraient être adressées à :

Statistique Canada

Services à la clientèle

Division des enquêtes spéciales

Téléphone : 613-951-3321 ou appelez sans frais : 1-800-461-9050

Télécopieur : 613-951-4527

Courriel : [des@statcan.ca](mailto:des@statcan.ca)



## 2.0 Contexte

Malgré leurs vies occupées et leurs nombreuses obligations, des millions de Canadiens font un effort conscient pour aider leur prochain et leur collectivité en versant des dons de bienfaisance, en faisant du bénévolat au sein d'organismes sans but lucratif et de bienfaisance et en aidant leurs concitoyens directement, de leur propre chef.

En 1997, l'Enquête nationale sur le don, le bénévolat et la participation (ENDBP) a brossé le premier tableau détaillé des contributions que se sont faites les Canadiens en donnant de leur temps et de leur argent. L'ENDBP est issue d'un partenariat tout particulier entre divers ministères fédéraux et des organismes sans but lucratif et bénévoles, parmi lesquels figurent Bénévoles Canada, le Centre canadien de philanthropie (qui a maintenant adopté le nom Imagine Canada), Développement des ressources humaines Canada, Patrimoine canadien, Santé Canada, et Statistique Canada. En utilisant un cadre semblable, on a de nouveau réalisé cette enquête en 2000 dans le cadre de l'Initiative sur le secteur bénévole et communautaire du gouvernement fédéral. En 2001, l'administration fédérale a financé la mise en œuvre d'un programme d'enquête permanent sur le don de bienfaisance, le bénévolat et la participation au sein de Statistique Canada. On a rebaptisé l'enquête « Enquête canadienne sur le don, le bénévolat et la participation » (ECDBP) pour la distinguer d'enquêtes semblables menées dans d'autres pays.

L'établissement d'un programme permanent d'enquête a permis de réexaminer la conception du questionnaire pour veiller à ce qu'il produise des renseignements de grande qualité sur une base régulière. Des consultations ont été menées auprès de divers intervenants du secteur bénévole et sans but lucratif, des administrations publiques et du milieu universitaire en vue d'améliorer l'enquête. En 2004, on a recueilli, pour la première fois, des données auprès de la population du Nord (Yukon, Territoires du Nord-Ouest et Nunavut), où un échantillon représentatif de 1 332 répondants âgés de 15 ans et plus ont pris part à l'enquête. On a augmenté l'effectif de l'échantillon dans les dix provinces, celui-ci passant de 14 724 en 2000 à 20 832 en 2004, afin d'accroître les possibilités d'estimations à l'échelle provinciale et à l'échelle des grandes régions urbaines. Le questionnaire a été modifié de diverses façons à la lumière de l'expérience acquise lors des enquêtes antérieures. On a reformulé certaines questions pour que les répondants les comprennent mieux. On a ajouté des questions pour recueillir de nouveaux renseignements présentant de l'intérêt. On a également éliminé certaines questions de l'enquête. Dans la mesure où l'enquête s'inscrit maintenant dans un programme permanent, il est possible de moduler des ensembles de questions d'un cycle à l'autre.

On a également modifié la base de l'enquête. L'ENDBP a été réalisée auprès d'un sous-échantillon de répondants à l'Enquête sur la population active (EPA) de Statistique Canada. Pour éviter d'alourdir indûment le fardeau de réponse dans l'Enquête sur la population active, on a mené la composante provinciale de l'ECDBP à titre d'enquête téléphonique à composition aléatoire (CA) dans le cadre de laquelle les répondants ont été choisis explicitement pour prendre part à l'ECDBP.

L'ECDBP de 2004 propose une nouvelle méthode de mesurer le don, le bénévolat et la participation, qui remplace la méthode employée dans le cadre de l'ENDBP de 1997 et de 2000. Les modifications apportées sont telles qu'on ne devrait pas comparer les résultats de l'ECDBP de 2004 à ceux des ENDBP antérieures.





### **3.0 Objectifs**

L'Enquête canadienne sur le don, le bénévolat et la participation poursuit trois objectifs :

- 1) recueillir des données à l'échelle nationale pour combler une lacune en matière de renseignements sur les comportements contributifs des particuliers, y compris le bénévolat, les dons à des organismes de bienfaisance et la participation civique;
- 2) fournir des données fiables et ponctuelles au Système de comptabilité nationale;
- 3) communiquer à la fois au public et aux secteurs bénévoles les décisions en matière de politiques et de programmes qui ont trait aux secteurs de bienfaisance et de bénévolat.



## 4.0 Concepts et définitions

Ce chapitre donne un aperçu des concepts et des définitions d'intérêt pour les utilisateurs.

### Bénévole

C'est une personne qui a fait du bénévolat, c'est-à-dire qui a fourni un service sans rémunération par l'entremise d'un groupe ou d'un organisme sans but lucratif ou de bienfaisance, au moins une fois au cours des 12 mois ayant précédé l'enquête. Cette définition comprend l'aide non rémunérée fournie à une école, un organisme religieux, une association communautaire ou une association de sports.

### Bénévole non encadré

L'Enquête canadienne sur le don, le bénévolat et la participation (ECDBP) définit ainsi un bénévole non encadré (ou un fournisseur d'aide directe) : personne qui a aidé quelqu'un de son propre chef, c'est-à-dire non par l'entremise d'un groupe ou d'une organisation, durant la période de référence de 12 mois ayant précédé l'enquête. Cette définition comprend l'aide directe donnée aux amis, aux voisins et aux membres de la famille, mais ne tient pas compte de l'aide offerte aux membres du ménage. Comme ces activités ne sont pas encadrées par une organisation, elles ne sont pas incluses dans la définition du bénévolat.

### Classification des organismes

On a demandé aux répondants de fournir certains renseignements sur les organismes pour lesquels ils ont fait du bénévolat ou auxquels ils ont fait des dons. On leur a d'abord demandé de fournir le nom de l'organisme. On s'est servi d'une liste préétablie des organismes les plus souvent mentionnés par les répondants des enquêtes de 1997 et 2000. Si un répondant signalait un organisme qui ne figurait pas sur cette liste, il était prié de fournir des renseignements sur la raison d'être de celui-ci pour qu'on puisse le placer dans l'une des catégories générales.

Les organismes ont été classés à l'aide de la Classification internationale des organismes à but non lucratif (CIOBNL)<sup>1</sup>. Bien qu'ils soient classifiés selon leur activité principale, certains organismes oeuvrent au sein de plusieurs domaines. La CIOBNL a l'avantage d'être largement utilisée par de nombreux autres pays, ce qui permet de faire des comparaisons internationales. Ce système de classification a également été développé pour le secteur des organismes à but non lucratif et bénévoles afin de refléter l'éventail et la nature des activités usuelles. La CIOBNL, établie dans le cadre du Johns Hopkins Comparative Nonprofit Sector Project et modifiée pour fin d'usage au Canada, se divise en 15 grandes catégories d'activités, y compris une catégorie fourre-tout « groupes non classés ailleurs ». Les 15 catégories sont à nouveau regroupées en 12 catégories.

Les 15 catégories sont :

- 1) *Arts et culture* : Cette catégorie regroupe des organismes et des activités des domaines généraux et spécialisés des arts et de la culture. Elle comprend : médias et communications; les arts visuels, l'architecture, poterie; arts d'interprétation; sociétés historiques, littéraires et humanistes; musées; zoos et aquariums.
- 2) *Sports et loisirs* : Cette catégorie regroupe les organismes et les activités des domaines généraux et spécialisés des sports et des loisirs. Elle comprend deux sous-groupes : (1) les sports amateurs (comprend les centres de conditionnement physique et de mieux-être); (2) les clubs de loisirs et amicales (comprend les clubs sociaux).
- 3) *Éducation et recherche* : Cette catégorie regroupe des organismes et des activités d'éducation et de recherche, qu'il s'agisse d'administration, de prestation, de promotion, de mise en oeuvre, de soutien ou de services. Elle comprend trois sous-groupes : (1) les organismes se consacrant à l'enseignement primaire ou secondaire; (2) les organismes se consacrant à d'autres formes d'enseignement (c'est-à-dire éducation des adultes et éducation permanente, écoles de

---

<sup>1</sup> Cette classification est basée sur L.M. Salamon and H.K. Anheier, 1997. *Defining the Nonprofit Sector: A Cross-national Analysis*. Manchester University Press.

formation professionnelle et technique); (3) les organismes se consacrant à la recherche (c'est-à-dire recherche médicale, sciences et technologie, sciences sociales). Veuillez noter que les organismes qui se consacrent principalement à l'éducation et à la recherche dans les domaines des maladies spécifiques (par exemple, la Société canadienne du cancer, la Fondation des maladies du cœur du Canada) sont inclus sous la catégorie 5, *Santé*.

- 4) *Universités et collèges* : Cette catégorie regroupe les organismes et les activités liés à l'enseignement supérieur. Elle comprend les universités, les écoles de gestion des affaires, de droit et de médecine.
- 5) *Santé* : Cette catégorie regroupe les organismes dont les activités sont liées à la santé et qui consistent principalement à fournir des services aux malades externes. Elle comprend deux sous-groupes : (1) le traitement externe des maladies mentales, les services d'intervention d'urgence; (2) autres services (c'est-à-dire éducation en santé et mieux-être public, soins ambulatoires, services de consultation externe, services médicaux de réadaptation externes, et services médicaux d'urgence). Elle comprend aussi les organismes qui se consacrent principalement à l'éducation, la recherche ou le soutien dans les domaines des maladies spécifiques (par exemple, la Société canadienne du cancer, la Fondation des maladies du cœur du Canada) ainsi que les organismes fournissant du soutien aux personnes victimes de maladies mortelles (par exemple, les carrefours et d'autres types de soins palliatifs).
- 6) *Hôpitaux* : Cette catégorie regroupe les organismes dont les activités sont liées à la santé et qui consistent principalement à fournir des services aux malades internes. Elle comprend deux sous-groupes : (1) les hôpitaux et des activités liées à la réadaptation; (2) les maisons de repos.
- 7) *Services sociaux* : Cette catégorie regroupe des organismes et des établissements fournissant des services sociaux à une collectivité ou à un public cible. Elle comprend trois sous-groupes : (1) services sociaux (dont les organismes fournissant des services aux enfants, aux jeunes, aux familles, aux personnes handicapées et âgées, ou encore des services sociaux personnels ou d'entraide); (2) services d'urgence et de secours; (3) services de soutien et maintien du revenu.
- 8) *Environnement* : Cette catégorie regroupe des organismes voués à la protection de l'environnement qui offrent des services axés sur la sauvegarde de l'environnement, la lutte antipollution et la prévention de la pollution, l'éducation relative à l'environnement et à la salubrité de l'environnement et la défense des animaux. Elle comprend deux sous-groupes : (1) l'environnement; (2) la protection des animaux.
- 9) *Développement et logement* : Cette catégorie regroupe des organismes offrant des programmes et des services visant à favoriser le développement des collectivités et l'amélioration du bien-être économique et social de la société. Elle comprend trois sous-groupes : (1) développement économique, social et communautaire (dont les organismes communautaires et les organisations de quartier); (2) logement; (3) emploi et formation.
- 10) *Droit, défense des intérêts et politique* : Cette catégorie regroupe des organismes et des groupes qui oeuvrent pour la protection et la promotion des droits de la personne et des autres droits, qui défendent les intérêts sociaux et politiques de la population en général ou de groupes particuliers, qui offrent des services juridiques et qui servent à promouvoir la sécurité du public. Elle comprend trois sous-groupes : (1) associations civiques et organismes de défense; (2) services juridiques; (3) organismes politiques.
- 11) *Octroi de subventions, collecte de fonds et promotion du bénévolat* : Cette catégorie regroupe des organismes philanthropiques et les organismes dont le but est de promouvoir les activités non lucratives comme les fondations accordant des bourses et des subventions, les organismes faisant la promotion du bénévolat et les organismes de collecte de fonds.
- 12) *Organismes internationaux* : Cette catégorie regroupe des organismes qui favorisent la bonne entente entre les gens de nationalités et de cultures diverses et qui, de plus, fournissent des secours d'urgence et travaillent au développement et au mieux-être à l'étranger.

- 13) *Religion* : Cette catégorie regroupe des organismes qui mettent en valeur les croyances religieuses, célèbrent des services et des rites religieux (par exemple, les églises, les mosquées, les synagogues, les temples, les sanctuaires, les séminaires, les monastères et autres institutions religieuses du genre), ainsi que leurs organismes auxiliaires.
- 14) *Associations d'affaires et professionnelles, syndicats* : Cette catégorie regroupe des organismes qui soutiennent, régissent et protègent les intérêts du milieu professionnel, des affaires et du travail.
- 15) *Groupes non classés ailleurs*.

La concordance entre la classification de 12 catégories et la classification de 15 catégories est comme suit :

CIOBNL 12 catégories	CIOBNL 15 catégories
1) Culture et divertissements	1) Arts et culture 2) Sports et loisirs
2) Éducation et recherche	3) Éducation et recherche 4) Universités et collèges
3) Santé	5) Santé 6) Hôpitaux
4) Services sociaux	7) Services sociaux
5) Environnement	8) Environnement
6) Développement et logement	9) Développement et logement
7) Droit, défense des intérêts et politique	10) Droit, défense des intérêts et politique
8) Intermédiaires de bienfaisance et de bénévolat	11) Octroi de subventions, collecte de fonds et promotion du bénévolat
9) Organismes internationaux	12) Organismes internationaux
10) Religion	13) Religion
11) Associations d'affaires et professionnelles, syndicats	14) Associations d'affaires et professionnelles, syndicats
12) Groupes non classés ailleurs	15) Groupes non classés ailleurs

### **Don de biens ou don en nature**

C'est un don non monétaire fait à un organisme sans but lucratif ou de bienfaisance, par exemple, un don de vêtements, d'articles ménagers ou d'aliments.

### **Don en argent**

Un don en argent est le montant monétaire versé à un organisme à but non lucratif ou de bienfaisance pendant la période de référence de 12 mois ayant précédé l'enquête. L'argent versé au même organisme plusieurs fois en réponse à la même méthode de sollicitation est considéré comme étant un seul don. Par exemple, tout l'argent donné à un organisme religieux au cours des 12 mois ayant précédé l'enquête par une collecte au lieu de culte est considéré comme étant un seul don.

### **Donateur**

Un donateur est une personne qui a effectué au moins un don en argent à un organisme à but non lucratif ou de bienfaisance pendant la période de référence de 12 mois ayant précédé l'enquête.

### **Industrie et profession**

L'ECDBP de 2004 procure des données par industrie et par profession sur les personnes occupées seulement (c.-à-d. au sujet du travail que la personne faisait dans la semaine ayant précédé l'interview). Pour la composante industrie, les statistiques ont été fournies selon le Système de classification des industries de l'Amérique du Nord (SCIAN) de 1997 et de 2002. Pour la composante profession, à la fois

la Classification type des professions (CTP) de 1991 et la Classification nationale des professions – statistiques (CNP-S) de 2001 ont été utilisées.

### **Participant**

L'ECDBP définit un participant comme étant une personne qui était membre d'au moins un organisme ou groupe communautaire à n'importe quel moment au cours de l'année ayant précédé l'enquête. Ceci comprend associations professionnelles ou syndicats; clubs sociaux ou sociétés fraternelles; organismes politiques; organismes à vocation culturelle ou pédagogique ou de loisir; clubs sportifs ou récréatifs; groupes d'appartenance religieuse; groupes pour les personnes âgées ou les jeunes; programmes de soutien ou d'auto-assistance; groupes de conservation ou écologique; organismes communautaires ou scolaires.

### **Période de référence**

Pour la plupart des questions dans le questionnaire de l'ECDBP, la période de référence était les 12 mois ayant précédé l'interview. Pour la composante provinciale, les interviews ont été menées du 13 septembre au 19 décembre 2004. Pour la composante territoriale (la composante du Nord), les interviews ont été menées du 30 août au 15 novembre 2004.

### **Situation vis-à-vis de l'activité**

C'est la situation du répondant sur le marché du travail. Pour l'ECDBP de 2004, les estimations de la situation vis-à-vis de l'activité se rapportent à la population sondée âgée de 15 à 75 ans, car on n'a pas posé aux répondants de 76 ans et plus la série de questions connexe.

Les trois catégories de situation vis-à-vis de l'activité sont les suivantes : « personnes occupées, « chômeurs » et « inactifs ». Aux fins de l'ECDBP, les trois catégories de situation vis-à-vis de l'activité sont définies comme suit :

#### **Personnes occupées**

Les personnes occupées sont celles qui, au cours de la semaine ayant précédé l'interview :

- a) ont fait un travail<sup>2</sup> quelconque dans le cadre d'un emploi ou dans une entreprise; ou
- b) avaient un emploi mais n'étaient pas au travail à cause d'une maladie ou d'une incapacité, pour obligations personnelles ou familiales, pour des vacances, par suite d'un conflit de travail ou du fait de tout autre facteur (cela n'inclut pas les personnes mises à pied ou celles qui étaient inactives entre deux emplois occasionnels).

#### **Chômeurs**

Les chômeurs sont les personnes qui, au cours de la semaine ayant précédé l'interview :

- a) avaient été mises à pied temporairement (à l'exclusion des étudiants à plein temps); ou
- b) étaient sans emploi, avaient activement cherché un emploi au cours des quatre dernières semaines (à l'exclusion des étudiants à plein temps et les retraités).

#### **Inactifs**

Les inactifs sont ceux qui n'avaient pas travaillé pendant la semaine ayant précédé l'interview et :

- a) soit étaient incapables de travailler en permanence;
- b) soit étaient étudiants à plein temps et avaient un emploi, mais étaient absents du travail parce qu'ils avaient été mis à pied ou qu'ils étaient entre deux emplois occasionnels;
- c) soit étaient étudiants à plein temps ou retraités qui n'avaient pas d'emploi et étaient à la recherche de travail;

---

<sup>2</sup> On entend par « travail » tout travail fait contre rémunération ou en vue d'un bénéfice, c'est-à-dire le travail fait contre rémunération pour un employeur ou son propre compte. Il peut également s'agir d'un travail familial non rémunéré, c'est-à-dire d'un travail non rémunéré qui contribue directement à l'exploitation d'une ferme, d'une entreprise ou d'un bureau professionnel, qui est possédé ou exploité par un membre apparenté du ménage. Parmi les tâches en question, on retrouve la comptabilité, la vente de produits et le service aux tables. Les tâches ménagères ou l'entretien de la maison ne sont pas considérés comme faisant partie du travail familial non rémunéré.

- d) soit n'avaient pas d'emploi et n'étaient pas à la recherche de travail.

**Travail communautaire obligatoire**

Il s'agit de l'aide non rémunérée pour le compte d'un groupe ou organisme qui était obligé ou requis par une école, un employeur, un organisme sans but lucratif ou de bienfaisance ou autres autorités. Dans l'ECDBP de 2004, on tient compte du travail communautaire obligatoire sous la définition du bénévolat.





## **5.0 Méthodologie de l'enquête de la composante provinciale**

L'Enquête canadienne sur le don, le bénévolat et la participation (ECDBP) de 2004 a été menée dans les dix provinces du 13 septembre au 19 décembre 2004 par composition aléatoire (CA), une technique permettant de générer des numéros de téléphone au hasard par ordinateur. Les interviews ont été effectuées par téléphone.

### **5.1 Population visée**

La population cible est formée des personnes âgées de 15 ans et plus ne vivant pas dans un établissement et résidant dans les dix provinces du Canada.

La population enquêtée ne comprend pas les personnes vivant dans un ménage qui ne dispose pas d'une ligne téléphonique terrestre, c'est-à-dire que les personnes vivant au sein de ménages sans téléphone ou qui ne disposent que de téléphones cellulaires sont exclues. On estime que, en 2004, environ 4,2 %<sup>3</sup> des ménages des dix provinces du Canada n'étaient pas branchés à une ligne téléphonique terrestre – 1,5 % des ménages n'avaient pas le téléphone et 2,7 % n'avaient que des téléphones cellulaires. Il importe de noter que, bien que ces personnes aient été exclues de la population visée par l'enquête, les estimations ont été pondérées pour en tenir compte. On présume donc que les personnes vivant dans ces ménages ont les mêmes caractéristiques et les mêmes comportements que les personnes visées par l'enquête.

### **5.2 Plan de sondage**

#### **5.2.1 Stratification**

L'échantillon de la composante provinciale de l'ECDBP se fonde sur un plan stratifié faisant appel à un échantillonnage probabiliste. La stratification a été effectuée à l'échelle de la province/région métropolitaine de recensement (RMR). Vingt-sept strates ont été constituées. Chaque province a été subdivisée en un certain nombre de strates définies selon la RMR (allant de zéro à l'Île-du-Prince-Édouard à quatre en Ontario) et une strate résiduelle supplémentaire « non-RMR » couvrant le reste de la province a également été créée.

#### **5.2.2 Répartition de l'échantillon**

La taille de l'échantillon a été déterminée de manière à permettre la production des données suivantes :

- 1) estimations transversales des bénévoles à l'échelle provinciale et pour les trois principales RMR;
- 2) estimations transversales des non-bénévoles à l'échelle provinciale et pour les trois principales RMR;
- 3) estimations transversales nationales des immigrants;
- 4) estimations longitudinales nationales des personnes dont la situation au regard du bénévolat change.

On a déterminé qu'il fallait obtenir 40 000 réponses pour atteindre ces objectifs. Une allocation de puissance (puissance = 0,20) a été utilisée pour répartir toutes les réponses escomptées entre les trois principales RMR, le reste de la province de ces RMR, et toutes les autres provinces. On a ensuite procédé à la répartition proportionnelle au sein de la province pour les autres strates. Le taux de réponse prévu ayant été établi à 80 %,

---

<sup>3</sup> Selon les données de l'Enquête sur le service téléphonique résidentiel de décembre 2004 menée par Statistique Canada.

on a déterminé qu'il fallait un échantillon de 50 000 unités pour obtenir les 40 000 réponses requises. En raison du recours à la CA, on doit tenir compte du fait que les numéros de téléphone ne seront pas tous des numéros résidentiels valides. Un échantillon sélectionné par CA comprendra un nombre important de numéros d'affaires et hors service. De plus, pendant la collecte des données, on trouvera inévitablement des numéros dont le statut (affaires ou résidentiel) restera indéterminé. On a donc augmenté la taille de l'échantillon pour tenir compte de tous ces facteurs, à la lumière de l'expérience acquise par Statistique Canada dans le cadre de l'Enquête sociale générale, et on a ainsi obtenu une taille d'échantillon de 120 650 unités.

### **5.3 Tirage de l'échantillon**

L'échantillon de la composante provinciale de l'ECDBP a été sélectionné au moyen d'une technique perfectionnée d'échantillonnage par CA appelée « élimination des banques non valides » (ÉBNV). Dans chaque strate, une liste des banques valides (indicatif régional suivi des cinq prochains chiffres) a été dressée à partir des fichiers administratifs des compagnies de téléphone. Pour les fins des enquêtes sociales, une banque valide se définit comme une banque renfermant au moins un numéro de téléphone résidentiel valide. Par conséquent, toutes les banques qui renferment uniquement des numéros de téléphone non assignés, d'appareils cellulaires, non valides ou d'affaires sont exclues de la base de sondage.

Un échantillon systématique des banques (avec remplacement) a été sélectionné dans chacune des strates. Pour chaque banque sélectionnée, un numéro à deux chiffres (de 00 à 99) a été généré aléatoirement. Ce numéro aléatoire a été ajouté à la banque pour former un numéro de téléphone complet. Cette méthode permet de sélectionner des numéros de téléphone résidentiels inscrits et non inscrits à l'annuaire, des numéros d'affaires et des numéros non valides (qui ne sont pas ou qui n'ont jamais été en service). Une procédure automatisée de filtrage pré composition, visant à éliminer les numéros hors service et d'entreprise, a été effectuée avant l'envoi de l'échantillon à l'équipe chargée de l'interview téléphonique assistée par ordinateur (ITAO). L'échantillon définitif, transmis à l'équipe de l'ITAO, comportait 90 721 numéros de téléphone.

On a composé tous les numéros de téléphone figurant dans l'échantillon de l'ITAO pour déterminer s'ils étaient rattachés à un ménage. Dans l'affirmative, on a demandé à la personne ayant pris l'appel de fournir des renseignements sur les différents membres du ménage. Un membre du ménage âgé de 15 ans ou plus a ensuite été choisi au hasard pour prendre part à l'enquête. Les interviews par procuration n'ont pas été permises.

On a posé aux répondants sélectionnés une série de 15 questions pour déterminer leur situation au regard du bénévolat. Les répondants ayant fait du bénévolat ont répondu aux autres questions de l'interview. Par contre, les non-bénévoles ont fait l'objet d'un sous-échantillonnage à un taux de 50 %, et seuls les répondants sous échantillonnés ont dû répondre aux sections pertinentes du reste du questionnaire. Lors de la création du fichier de l'échantillon, on a inclus un indicateur dont la valeur était aléatoirement établie à une dans 50 % des cas et à zéro dans 50 % des cas. Les répondants non bénévoles associés à un indicateur dont la valeur a été établie à une ont dû répondre aux autres questions; pour ceux associés à un indicateur dont la valeur était de zéro, l'interview a pris fin au terme de la première série de 15 questions.

## 5.4 Taille de l'échantillon selon la province

Le tableau suivant présente le nombre de numéros de téléphone générés pour la composante provinciale de l'ECDBP de 2004, ainsi que le nombre de répondants avant et après le sous-échantillonnage des non-bénévoles :

Provinces	Nombre de numéros de téléphone générés	Nombre de réponses avant le sous-échantillonnage des non-bénévoles	Nombre de réponses après le sous-échantillonnage des non-bénévoles
Terre-Neuve-et-Labrador	9 003	1 990	1 407
Île-du-Prince-Édouard	5 803	1 314	936
Nouvelle-Écosse	8 619	2 182	1 612
Nouveau-Brunswick	9 623	2 113	1 510
Québec	17 914	4 510	2 948
Ontario	25 217	5 421	4 071
Manitoba	8 445	2 554	1 834
Saskatchewan	9 034	2 272	1 688
Alberta	9 172	2 487	1 807
Colombie-Britannique	17 820	4 188	3 019
<b>Toutes les provinces</b>	<b>120 650</b>	<b>29 031</b>	<b>20 832</b>



## **6.0 Méthodologie de l'enquête de la composante territoriale (la composante du Nord)**

L'Enquête canadienne sur le don, le bénévolat et la participation (ECDBP) a été réalisée dans les trois territoires du 30 août au 15 novembre 2004 auprès d'un sous-échantillon de ménages ayant fait partie de l'Enquête sur la population active (EPA) pendant une période de trois mois. Par conséquent, le plan d'échantillonnage de l'ECDBP est étroitement lié à celui de l'EPA. L'ECDBP n'a pas été menée à titre de supplément de l'EPA à proprement parler puisque la collecte des données n'a pas été effectuée en même temps que celle de l'EPA. Il a donc fallu, dans le cadre de l'ECDBP, répéter la collecte des renseignements de base de même que des données relatives aux variables de l'EPA présentant de l'intérêt.

### **6.1 Population visée**

La population cible est formée des personnes âgées de 15 ans et plus résidant dans les trois territoires du Canada. Sont exclus de la population cible :

- les personnes vivant en établissement;
- les résidents des réserves indiennes (à l'exception des résidents de la réserve Hay River dans les Territoires du Nord-Ouest qui font partie de la population cible);
- les membres à temps plein des Forces armées canadiennes.

Au Yukon et dans les Territoires du Nord-Ouest, l'EPA ne vise que la population de certaines collectivités. Pour des raisons d'ordre opérationnel et de coûts, les très petites collectivités sont exclues. On estime que les collectivités visées couvrent plus de 90 % de la population âgée de 15 ans et plus au Yukon et dans les Territoires du Nord-Ouest. Au Nunavut, les collectivités retenues aux fins de l'échantillonnage couvrent moins de 70 % de la population âgée de 15 ans et plus. Les estimations sont toutefois pondérées en fonction de l'ensemble de la population âgée de 15 ans et plus.

### **6.2 Plan de sondage**

Dans le Nord, l'EPA s'appuie sur un plan d'échantillonnage à plusieurs degrés. L'unité primaire d'échantillonnage (UPÉ) est la collectivité. L'échantillonnage des UPÉ est suivi de l'échantillonnage des ménages.

#### **6.2.1 Renouvellement de l'échantillon**

L'EPA dans le Nord emploie un plan de sondage avec renouvellement de panel, l'échantillon étant constitué de huit panels ou groupes de renouvellement. On communique avec les ménages du panel tous les trois mois et ces ménages restent dans l'échantillon pendant huit trimestres, de sorte que chaque ménage fait partie de l'échantillon pendant près de deux ans. L'enquête est réalisée tous les mois. Un tiers de l'échantillon trimestriel est joint chaque mois, si bien que 1/24<sup>e</sup> de l'échantillon est renouvelé mensuellement.

### **6.2.2 Modifications apportées au plan de l'Enquête sur la population active pour l'Enquête canadienne sur le don, le bénévolat et la participation**

L'échantillon de l'ECDBP comprend tous les ménages faisant partie de l'échantillon de l'EPA des mois d'octobre, novembre et décembre 2004, exception faite des ménages apparaissant dans l'échantillon de l'EPA pour la première fois.

L'ECDBP a fait appel à sept des huit groupes de renouvellement de l'échantillon de l'EPA d'octobre, novembre et décembre 2004. Le premier groupe de renouvellement a été exclu. Les renseignements de base ont été recueillis pour tous les membres du ménage, puis un membre du ménage âgé de 15 ans ou plus a été choisi au hasard pour répondre aux autres questions du questionnaire de l'ECDBP. Les réponses par procuration n'ont pas été permises. Contrairement à la composante provinciale, la composante territoriale n'a pas donné lieu à un sous-échantillonnage des non-bénévoles. Tous les non-bénévoles ont été invités à répondre aux questions de toutes les sections pertinentes du questionnaire.

### **6.3 Taille de l'échantillon**

L'échantillon est formé des groupes de renouvellement (exception faite du premier) de l'échantillon trimestriel des mois d'octobre, novembre et décembre 2004 de l'EPA. L'échantillon initial comportait 1 831 unités. Le tableau suivant présente la répartition de l'échantillon selon le territoire.

<b>Territoire</b>	<b>Taille de l'échantillon initial</b>	<b>Nombre de répondants</b>
Nunavut	438	335
Territoires du Nord-Ouest	680	489
Yukon	713	508
<b>Territoires total</b>	<b>1 831</b>	<b>1 332</b>

## **7.0 Collecte des données**

### **7.1 Conception du questionnaire**

L'ECDBP de 2004 propose une nouvelle méthode de mesurer le don, le bénévolat et la participation, qui remplace la méthode employée dans le cadre des enquêtes nationales sur le don, le bénévolat et la participation (ENDBP) de 1997 et de 2000. L'expérience acquise lors de l'ENDBP de 2000 a suggéré la nécessité d'apporter certaines modifications reliées au contenu du questionnaire.

Pour se préparer pour l'ECDBP de 2004, les représentants des gouvernements fédéral, provinciaux et territoriaux ainsi que les représentants du secteur bénévole et de l'enseignement ont été consultés. Ces consultations, ont eu lieu entre les mois de janvier et avril 2002, et ont porté principalement sur le contenu de l'enquête. Par la suite, les membres du comité directeur se sont réunis pour discuter des priorités du contenu. Cette réunion a permis l'élaboration de l'ébauche d'un instrument d'enquête présenté sous forme de test en groupe et d'entrevue un à un. L'essai qualitatif du contenu a été mené au cours des mois d'été dans l'ensemble du Canada. Les modifications apportées à l'enquête à la suite de l'essai qualitatif ont été suivies d'une enquête pilote en avril 2003. Cette enquête pilote nous a permis de corriger les erreurs inhérentes à l'application informatique et d'améliorer les procédures d'enquête.

L'ECDBP se divise en deux principales catégories de questions : celles qui mesurent les comportements et précisent ce que font les répondants dans leurs activités de dons, de bénévolat et de participation et celles qui mesurent les corrélats de ces comportements. Cette dernière catégorie regroupe les attitudes et les motivations ainsi que les facteurs qui restreignent ou facilitent les dons et le bénévolat.

### **7.2 Supervision et contrôle de qualité**

Tous les intervieweurs de Statistique Canada travaillent sous la supervision d'un groupe d'intervieweurs principaux. Ceux-ci ont pour responsabilité de s'assurer que les intervieweurs connaissent bien les concepts et procédures qui sont associées aux enquêtes auxquelles ils sont attribués. Les intervieweurs principaux sont aussi chargés d'exercer un contrôle périodique des interviews.

Les intervieweurs ont reçu une formation sur l'utilisation de l'application de l'interview téléphonique assistée par ordinateur (ITAO) et sur le contenu de l'enquête. En plus de la formation en classe, les intervieweurs complétaient une série d'interviews simulées enfin de se familiariser avec l'enquête, ses concepts et ses définitions.

### **7.3 Méthodologie de la collecte de données**

#### **7.3.1 Composante provinciale**

Pour les dix provinces, toutes les données ont été recueillies à l'aide des techniques de l'interview téléphonique assistée par ordinateur (ITAO). Le système ITAO contient plusieurs modules génériques qui s'adaptent rapidement à la plupart des types d'enquêtes. Un module frontal contient un ensemble de codes de réponse standards qui s'appliquent à toutes les issues possibles des appels, ainsi que les scénarios correspondants qui sont lus par les intervieweurs. Une approche normalisée a été utilisée pour présenter l'organisme, le nom et le but de l'enquête, les clients de l'enquête, l'utilisation qui sera faite des résultats et la durée de l'interview.

La sélection aléatoire d'une personne parmi le ménage a été effectuée au moment de l'interview. L'intervieweur s'est d'abord renseigné sur l'âge et le sexe de tous les

membres du ménage ainsi que sur les liens unissant chacun d'eux. Après la saisie de ces renseignements, l'application ITAO a choisi au hasard la personne devant répondre à l'ECDBP. Les répondants ont été informés que leur participation à l'enquête était volontaire et que les renseignements fournis resteraient strictement confidentiels.

L'application ITAO garantissait l'entrée des seules réponses valides et le bon enchaînement des questions. Des contrôles intégrés à l'application garantissaient la cohérence des réponses, repéraient et corrigeaient les valeurs aberrantes et déterminaient à qui étaient posées les différentes questions. Ainsi, à la fin du processus de collecte, les données étaient déjà passablement « épurées ».

Les cas ont été distribués dans cinq bureaux régionaux de Statistique Canada. La charge de travail et les intervieweurs de chaque bureau étaient supervisés par un chargé de projet. L'ordonnanceur automatique utilisé dans le système ITAO garantissait que les cas étaient assignés au hasard aux intervieweurs. Un maximum de 20 appels par cas identifié comme un numéro de téléphone résidentiel a été tenté. Une fois le maximum atteint, un intervieweur principal révisait le cas et déterminait si d'autres appels seraient tentés.

### **7.3.2 Composante territoriale**

La collecte de l'ECDBP dans les territoires était très semblable à la collecte dans les provinces, avec les exceptions suivantes :

- Toutes les données ont été recueillies à l'aide d'une application d'interview sur place assistée par ordinateur (IPAO) permettant à l'intervieweur de saisir directement les réponses au moment de l'interview;
- La plupart des interviews ont eu lieu par téléphone (74 %), mais dans le cas des ménages n'ayant pas une ligne téléphonique terrestre, les interviews ont été effectuées sur place (26 %).

## **7.4 Non-réponse**

Les intervieweurs ont reçu comme consigne de faire toute tentative raisonnable pour obtenir une interview complète avec le membre du ménage qui avait été choisi au hasard. On a communiqué de nouveau à au plus deux reprises avec les personnes qui avaient d'abord refusé de participer, afin de leur expliquer l'importance de l'enquête et de les inciter à y participer. Dans les cas où l'intervieweur avait appelé à un mauvais moment, celui-ci prenait rendez-vous en vue de rappeler à un moment plus opportun. Dans les cas où il n'y avait personne à la maison, de nombreux appels ont été faits.



## **8.0 Traitement des données**

Le principal produit de l'Enquête canadienne sur le don, le bénévolat et la participation (ECDBP) est un fichier de microdonnées « épuré ». Ce chapitre présente un bref résumé des phases de traitement inhérentes à la production de ce fichier.

### **8.1 Saisie des données**

Les intervieweurs saisissent directement les réponses aux questions de l'enquête au moment de l'interview à l'aide d'une version automatisée du questionnaire. L'emploi d'un questionnaire automatisé réduit les délais et coûts de traitement associés à la saisie des données, aux erreurs de transcription et à la transmission des données.

Une partie du contrôle se fait au moment de l'interview. Lorsque les renseignements introduits sont hors limites (trop faibles ou trop élevés) des valeurs attendues, ou qu'ils entrent en contradiction avec des renseignements introduits auparavant, l'intervieweur voit paraître à l'écran de l'ordinateur des messages lui demandant de modifier les renseignements. Cependant, pour certaines questions, l'intervieweur a la possibilité de passer outre aux contrôles et de sauter des questions si l'enquêté ne connaît pas la réponse ou refuse de répondre. Pour cette raison, on soumet les données des réponses à d'autres processus de vérification et d'imputation après réception au bureau central.

### **8.2 Vérification**

La première étape du traitement d'enquête effectué au bureau central a été de remplacer les valeurs « hors limites » incluses dans le fichier de données par des blancs. Ce processus a été conçu pour faciliter les vérifications ultérieures.

Les erreurs dans le déroulement du questionnaire, où l'on a relevé des questions qui ne s'appliquaient pas au répondant (et auxquelles on n'aurait donc pas dû répondre) et qui renfermaient des réponses, constituaient le premier type d'erreurs traitées. Dans ces cas, une vérification par ordinateur a éliminé automatiquement les données superflues en suivant l'ordre du questionnaire dicté par les réponses à des questions antérieures et subséquentes, parfois.

Le second type d'erreurs traitées avait trait à un manque d'information dans les questions pour lesquelles le répondant aurait dû répondre. Pour ce type d'erreur, un code de non-réponse ou « non déclaré » était attribué au poste.

### **8.3 Codage des questions ouvertes**

Quelques données élémentaires ont été consignées sur le questionnaire par les intervieweurs sous forme de questions ouvertes. Le questionnaire informatisé renfermait une liste de sélection d'organismes bien connus qui aidait l'intervieweur au moment de saisir l'information relative au type d'organisme pour lequel le répondant avait fait du bénévolat (VD\_Q01) ou auquel il avait fait un don (GS\_Q01). Lorsque l'organisme cité par le répondant ne figurait pas sur la liste de sélection, on demandait au répondant de fournir certains renseignements sur ce que faisait l'organisme, lesquels servaient à coder le type d'organisme selon la Classification internationale des organisations à but non lucratif (CIOBNL), Révision 1 (pour de plus amples renseignements sur ce système de classification, voir le chapitre 4.0).

On a effectué le codage de l'industrie (Système de classification des industries de l'Amérique du Nord de 1997 et 2002) et de la profession (Classification type des professions de 1991 et Classification nationale des professions – statistiques de 2001) liées à l'emploi qu'occupait le

répondant la semaine ayant précédé l'interview en fonction des réponses données aux questions LF\_Q05 à LF\_Q08.

Dans le cas des six questions suivantes du questionnaire de l'ECDBP, le texte inscrit dans la catégorie « Autre - Précisez » a été examiné au bureau central et, lorsque c'était possible, codé dans une catégorie existante.

- FV\_Q16 : autres activités de bénévolat;
- FG\_Q15 : autres méthodes utilisées en faisant des dons à un organisme de bienfaisance ou à but non lucratif;
- PA\_Q11 : autres types d'organismes auxquelles le répondant était membre;
- SD\_Q01 : religion;
- SD\_Q03 : pays de naissance;
- SD\_Q08 : ethnicité des ancêtres.

## **8.4 Imputation**

L'imputation est le processus qui fournit des valeurs valides concernant les variables qui ont été retenues pour être modifiées, soit en raison de renseignements invalides, soit en raison de renseignements manquants. Les nouvelles valeurs sont établies de façon à préserver la structure sous-jacente des données et à garantir que les enregistrements qui en résultent passeront tous les contrôles requis. En d'autres mots, l'objectif n'est pas de reproduire les véritables valeurs des microdonnées mais plutôt d'établir des enregistrements intrinsèquement cohérents qui permettront de produire de bonnes estimations agrégées.

Nous pouvons faire la distinction entre trois types de non-réponse. La non-réponse complète est lorsque le répondant ne fournit pas le nombre minimal de réponses. Ces enregistrements sont supprimés et seront pris en compte lors du processus de pondération (voir le chapitre 12.0). La non-réponse ponctuelle est lorsque le répondant ne fournit pas une réponse à une question mais qu'il passe à la question suivante. Ce type de non-réponse est habituellement traité en utilisant le code « non déclaré » ou en ayant recours à l'imputation. Enfin, la non-réponse partielle est lorsque le répondant fournit le nombre minimal de réponses mais ne termine pas l'interview. Ces enregistrements peuvent être traités comme des cas de non-réponse complète ou ponctuelle.

Dans le cas de l'ECDBP, l'imputation par enregistrement donneur a été utilisée pour fournir les éléments d'information manquants aux enregistrements avec non-réponse ponctuelle ou partielle. De plus amples renseignements sur le processus d'imputation sont donnés à la section 9.2.3.

## **8.5 Création de variables dérivées**

On a calculé un certain nombre de données élémentaires du fichier de microdonnées en combinant des variables du questionnaire pour faciliter l'analyse des données. La plupart des noms des variables dérivées ont un « D » à la position du quatrième caractère du nom. Certaines variables ont un « G » dans la position du quatrième caractère du nom. La plupart de ses variables ont été groupées pour faciliter l'utilisation.

À titre d'exemple de variables dérivées, mentionnons :

- le nombre total d'heures de bénévolat (VD1DHRS);
- le nombre total d'heures de bénévolat pour les 15 catégories d'organismes (VD1DTX01 à VD1DTX15 sur le fichier maître et VD1GTX01 à VD1GTX15 sur le fichier de microdonnées à grande diffusion (FMGD));
- le montant total des dons (GS1DATOT sur le fichier maître et GS1GATOT sur le FMGD)
- le montant total des dons pour les 15 catégories d'organismes (GS1DAX01 à GS1DAX15 sur le fichier maître et GS1GAX01 à GS1GAX15 sur le FMGD);

- le montant total des dons selon la méthode de sollicitation (FG1DA03 à FG1DA15 sur le fichier maître et FG1GA03 à FG1GA15 sur le FMGD).

Les variables dérivées pour les dons ont été dérivées d'après le fichier GS et incorporés au fichier MAIN (voir le chapitre 14.0 pour plus d'information sur la structure des fichiers).

De façon générale, on ne calculait pas de variable dérivée si l'on n'avait pas de réponse à une partie quelconque de l'équation (p. ex. ne sait pas, refus, et non déclaré). Dans ces cas, le code assigné à la variable dérivée était, en général, « non déclaré ».

## **8.6 Pondération**

Le principe qui sous-tend une estimation pour un échantillon probabiliste veut que chacune des personnes incluses dans l'échantillon « représente », en plus d'elle-même, plusieurs autres personnes qui en sont exclues. Par exemple, dans un échantillon aléatoire simple de 2 % de la population, chaque personne incluse dans l'échantillon représente 50 membres de la population.

La phase de la pondération est une étape où l'on calcule ce nombre (ou poids) pour chaque enregistrement. Ce poids, qui figure dans le fichier de microdonnées, **doit** servir à calculer des estimations significatives à partir de l'enquête. Si, par exemple, le nombre de personnes qui avaient fait du bénévolat au cours des 12 mois précédents, doit être estimé, cette opération s'effectue en sélectionnant les enregistrements se référant aux personnes incluses à l'intérieur de l'échantillon qui présentent cette caractéristique et en additionnant les poids inscrits dans ces enregistrements.

Le chapitre 12.0 renferme des détails au sujet de la méthode utilisée pour calculer ces poids.

## **8.7 Suppression de renseignements confidentiels**

Il convient de souligner que les fichiers de microdonnées « à grande diffusion » peuvent différer des fichiers « maîtres » de l'enquête que conserve Statistique Canada. Ces différences sont habituellement le résultat de mesures prises pour protéger l'anonymat des répondants à une enquête. Les mesures les plus courantes sont la suppression de variables du fichier, le regroupement de valeurs en des catégories plus étendues et le codage de valeurs spécifiques à la catégorie « non déclaré ». Les utilisateurs ayant besoin d'avoir accès à de l'information exclue des fichiers de microdonnées peuvent acheter des totalisations spéciales. Les estimations produites seront communiquées à l'utilisateur, sous réserve du respect des lignes directrices pour l'analyse et la diffusion dont le chapitre 10.0 de ce document fournit un aperçu.

Le fichier maître de données de l'enquête comprend des identificateurs géographiques plus explicites que le FMGD, notamment les régions métropolitaines de recensement et les régions urbaines. Le FMGD ne renferme pas d'identificateurs géographiques inférieurs au niveau provincial. Aussi, le fichier maître contient quelques variables démographiques qui sont exclues du FMGD, p. ex. l'ethnicité ancestrale et le statut d'immigrant.

Le fichier maître de l'enquête comprend quelques renseignements détaillés qui sont inclus sur le FMGD en forme groupé seulement. Ceci comprend

- l'âge précis du répondant;
- nombre d'enfants dans le ménage âgés de 0 à 5 ans (groupé sur le FMGD à une variable « oui/non » qui indique la présence d'enfants dans le ménage âgés de 0 à 5 ans);
- une classification détaillée, à 43 groupes, des industries de l'Amérique du Nord qui apparaît seulement sur le FMGD en forme d'un groupement de 18 catégories;
- pays de naissance, qui, pour le FMGD a été groupé à « Canada » et « Extérieur du Canada ».

Pour certaines variables sensibles au risque d'identification d'individus, le FMGD peut avoir subi de la suppression locale, c'est-à-dire que certaines valeurs présentes sur le fichier maître peuvent avoir été codées comme « non déclarées » sur le FMGD.

## 9.0 Qualité des données

### 9.1 Taux de réponse

#### 9.1.1 Réponse à la composante provinciale

Le taux de résolution et le taux de succès pour les numéros de téléphone selon la province sont présentés au tableau suivant.

Le taux de résolution des numéros de téléphone s'entend de la proportion des numéros de téléphone confirmés, soit lors du processus de présélection soit sur le terrain, comme étant des numéros résidentiels ou hors du champ de l'enquête (p. ex. numéros d'affaires ou hors service, numéros de téléphones cellulaires, de locaux non résidentiels ou de logements collectifs) par rapport au nombre total de numéros de téléphone générés.

$$\text{taux de résolution} = \frac{\text{nombre de numéros de téléphone résolus}}{\text{nombre de numéros de téléphone générés}}$$

Le taux de succès s'entend de la proportion des numéros de téléphone résolus confirmés comme étant des numéros de téléphone résidentiels.

$$\text{taux de succès} = \frac{\text{nombre de numéros de téléphone résidentiels}}{\text{nombre de numéros de téléphone résolus}}$$

#### Taux de résolution et de succès des numéros de téléphone selon la province

Province	Numéros de téléphone générés	Numéros de téléphone transmis à la collecte	Numéros de téléphone résolus sur le terrain	Numéros résolus total	Taux de résolution (%)	Numéros de téléphone résidentiels confirmés	Réponses	Taux de succès (%)
Terre-Neuve-et-Labrador	9 003	5 644	4 905	8 264	91,8	2 884	1 990	34,9
Île-du-Prince-Édouard	5 803	3 697	3 242	5 348	92,2	2 090	1 314	39,1
Nouvelle-Écosse	8 619	6 049	5 482	8 052	93,4	3 209	2 182	39,9
Nouveau-Brunswick	9 623	6 449	5 760	8 934	92,8	3 086	2 113	34,5
Québec	17 914	14 883	13 404	16 435	91,7	8 939	4 510	54,4
Ontario	25 217	20 264	17 731	22 684	90,0	10 809	5 421	47,7
Manitoba	8 445	6 592	6 297	8 150	96,5	3 714	2 554	45,6
Saskatchewan	9 034	5 743	5 209	8 500	94,1	3 727	2 272	43,8
Alberta	9 172	7 310	6 760	8 622	94,0	4 367	2 487	50,6
Colombie-Britannique	17 820	14 090	13 094	16 824	94,4	8 452	4 188	50,2
<b>Toutes les provinces</b>	<b>120 650</b>	<b>90 721</b>	<b>81 884</b>	<b>111 813</b>	<b>92,7</b>	<b>51 277</b>	<b>29 031</b>	<b>45,9</b>

Les taux de réponse de la composante provinciale de l'Enquête canadienne sur le don, le bénévolat et la participation (ECDBP) sont présentés au tableau suivant. Le répondant s'entend de toute personne échantillonnée qui a répondu aux 15 questions du module

Bénévolat encadré (FV) du questionnaire permettant de déterminer si la personne a fait du bénévolat ou non. Le taux de réponse s'entend du nombre de personnes échantillonnées ayant répondu au moins à ces questions divisé par le nombre de numéros de téléphone résidentiels confirmés.

$$\text{taux de réponse} = \frac{\text{nombre de répondants}}{\text{nombre de numéros de téléphone résidentiels}}$$

### Taux de réponse selon la province

Province	Numéros de téléphone résidentiels confirmés	Réponses	Taux de réponse (%)
Terre-Neuve-et-Labrador	2 884	1 990	69,0
Île-du-Prince-Édouard	2 090	1 314	62,9
Nouvelle-Écosse	3 209	2 182	68,0
Nouveau-Brunswick	3 086	2 113	68,5
Québec	8 939	4 510	50,5
Ontario	10 809	5 421	50,2
Manitoba	3 714	2 554	68,8
Saskatchewan	3 727	2 272	61,0
Alberta	4 367	2 487	56,9
Colombie-Britannique	8 452	4 188	49,6
<b>Toutes les provinces</b>	<b>51 277</b>	<b>29 031</b>	<b>56,6</b>

### 9.1.2 Réponse à la composante territoriale

Les taux de réponse de la composante territoriale de l'ECDBP sont présentés au tableau suivant. La même définition des répondants s'applique aux composantes provinciale et territoriale.

$$\text{taux de réponse} = \frac{\text{nombre de répondants}}{\text{nombre de répondants} + \text{nombre de non-répondants}}$$

### Taux de réponse selon le territoire

Territoire	Échantillon total	Hors du champ d'enquête	Répondants	Non-répondants	Taux de réponse (%)
Yukon	713	122	508	83	86,0
Territoires du Nord-Ouest	680	102	489	89	84,6
Nunavut	438	67	335	36	90,3
<b>Tous les territoires</b>	<b>1 831</b>	<b>291</b>	<b>1 332</b>	<b>208</b>	<b>86,5</b>

## **9.2 Erreurs relatives à l'enquête**

Les estimations calculées à partir de cette enquête reposent sur un échantillon de ménages. Des estimations légèrement différentes auraient pu être obtenues si un recensement complet avait été effectué en reprenant le même questionnaire et en faisant appel aux mêmes intervieweurs, superviseurs, méthodes de traitement, etc. que ceux effectivement utilisés dans l'enquête. L'écart entre les estimations découlant de l'échantillon et celles que donnerait un dénombrement complet réalisé dans des conditions semblables est appelé erreur d'échantillonnage de l'estimation. Des erreurs qui ne sont pas liées à l'échantillonnage peuvent se produire à presque toutes les étapes des opérations d'enquête. Les intervieweurs peuvent avoir mal compris les instructions, les enquêtés peuvent se tromper en répondant aux questions, les réponses peuvent être mal saisies sur le questionnaire et des erreurs peuvent survenir lors du traitement et de la totalisation des données. Ces erreurs sont toutes des exemples d'erreurs non dues à l'échantillonnage.

Sur un grand nombre d'observations, les erreurs aléatoires auront peu d'effet sur les estimations calculées à partir de l'enquête. Toutefois, les erreurs systématiques contribuent à biaiser les estimations de l'enquête. Énormément de temps et d'efforts ont été consacrés à réduire les erreurs non dues à l'échantillonnage dans l'enquête. Des mesures d'assurance de la qualité ont été prises à chacune des étapes du cycle de collecte et de traitement des données afin de contrôler la qualité des données. Ces mesures comprenaient le recours à des intervieweurs hautement qualifiés, une formation poussée des intervieweurs concernant les procédures d'enquête et du questionnaire, l'observation des intervieweurs en vue de cerner les problèmes liés à la conception du questionnaire ou à une mauvaise compréhension des instructions, des procédures visant à s'assurer que les erreurs de saisie des données étaient réduites au minimum ainsi que des vérifications de la qualité du codage et de contrôle ayant pour but d'attester la logique du traitement.

### **9.2.1 Collecte des données**

Pour sa formation, l'intervieweur devait lire par lui-même le Manuel de l'intervieweur de l'ECDBP et parcourir la publication sommaire intitulée *Canadiens dévoués, Canadiens engagés : Points saillants de l'Enquête nationale de 2000 sur le don, le bénévolat et la participation*, et ensuite suivre deux jours de formation en classe. Les manuels comprennent une description du contexte et des objectifs de l'enquête, de même qu'un glossaire et une série de questions et réponses. La formation en classe comprenait un exposé des objectifs de l'enquête, l'examen des principaux concepts ainsi que des exercices fondés sur des cas de formation (interviews simulées) faits avec l'application d'interview téléphonique assistée par ordinateur (ITAO). C'était aussi une occasion pour les intervieweurs de poser des questions avant le début de la collecte.

### **9.2.2 Traitement des données**

Le traitement des données de l'ECDBP a été fait par étapes, y compris la vérification, le codage, le contrôle, l'imputation, l'estimation, la confidentialité, etc. À chaque étape, une photo des fichiers de sortie a été prise et une vérification a été faite en comparant les fichiers de l'étape en cours avec ceux de l'étape antérieure.

### **9.2.3 Non-réponse et imputation**

L'effet de la non-réponse sur les résultats de l'enquête constitue une source importante d'erreurs non dues à l'échantillonnage dans les enquêtes. L'importance de la non-réponse varie d'une non-réponse partielle (le fait de ne pas répondre à une ou à plusieurs questions) à une non-réponse totale. Il y a non-réponse totale quand l'intervieweur n'a pu communiquer avec le répondant, lorsque aucun membre du ménage

ne pouvait fournir les renseignements ou quand le répondant a refusé de participer à l'enquête. On a traité les cas de non-réponse totale en ajustant le poids des ménages qui ont répondu à l'enquête afin de compenser pour ceux qui n'ont pas répondu.

Dans la majorité des cas, il y a non-réponse ponctuelle ou partielle à l'enquête quand le répondant n'a pas compris ou a mal interprété une question, quand il a refusé de répondre à une question ou quand il ne pouvait se rappeler le renseignement requis. Dans les cas de non-réponse partielle et de non-réponse ponctuelle, une imputation par donneurs a été exécutée pour certaines variables. La plupart de ces imputations ont été faites pour fournir des données complètes permettant de calculer les totaux (p. ex. le nombre total d'heures et la valeur totale de dons). En outre, l'imputation a permis de conserver des enregistrements dans l'échantillon même si une partie de l'information requise n'avait pas été fournie par le répondant.

Toutes les imputations ont été faites à l'aide d'enregistrements donneurs choisis au moyen d'une fonction de « score ». On a comparé certaines caractéristiques apparaissant dans chaque enregistrement comptant une non-réponse ponctuelle ou une non-réponse partielle (également appelé enregistrement receveur) à celles de tous les enregistrements donneurs. Lorsqu'une caractéristique d'un enregistrement donneur concorde avec celle d'un enregistrement receveur, la cote du donneur augmente. L'enregistrement donneur ayant obtenu la cote la plus élevée est considéré comme le donneur « le plus proche », et cet enregistrement est retenu pour fournir les éléments d'information manquants pour le non-répondant. Si la cote la plus élevée est partagée par plus d'un enregistrement donneur, on procède à une sélection aléatoire. Le bassin d'enregistrements donneurs est constitué de telle façon que la valeur imputée au receveur, parallèlement à d'autres éléments d'information non imputés du receveur, est encore acceptée à la vérification.

L'imputation a été effectuée en six étapes. Les trois premières étapes se rapportent à l'imputation de variables dans le fichier principal (voir le chapitre 14.0 pour la structure des fichiers). La première étape a consisté à imputer à la fois le revenu personnel et le revenu du ménage. À la deuxième étape, on a imputé les données sur les heures de bénévolat selon l'activité dans l'organisme principal. La troisième étape a consisté à imputer le nombre total d'heures de bénévolat au sein des deuxièmes et troisièmes organismes et le nombre total d'heures de bénévolat au sein de tous les autres organismes combinés. À la quatrième étape, les données relatives aux variables des contributions versées ont été imputées dans le fichier GS (le fichier sur les dons). On a ensuite créé des enregistrements supplémentaires dans le fichier GS lorsque la valeur « oui » a été imputée à la question GS\_Q07, « *Avez-vous fait d'autres dons en espèces en réponse à cette méthode de sollicitation?* ». À la cinquième étape, on a imputé, dans le fichier principal, les données manquantes aux variables qui indiquent si le répondant avait versé un don en réponse à chacune des 13 méthodes de sollicitation (FG\_Q03 à FG\_Q15 de la section FG du questionnaire (la section « *Dons en espèces à des organismes de charité* »). À cette étape, on a fait de l'imputation seulement pour les répondants qui étaient donateurs (c.-à-d. les cas avec au moins une valeur de « oui » aux questions FG\_Q03 à FG\_Q15). À la dernière étape, on a procédé à l'imputation des enregistrements partiellement remplis ne permettant pas de déterminer la situation du répondant au regard des dons en raison des valeurs manquantes à la section FG du questionnaire. Au total, 88 variables ont fait l'objet d'une imputation. Cette dernière étape comprenait, elle aussi, la création d'enregistrements supplémentaires dans le fichier GS lorsque la valeur « oui » a été imputée à au moins une des questions FG\_Q03 à FG\_Q15.

Le tableau suivant montre le nombre d'enregistrements dans lesquels certaines des variables clés de l'enquête ont fait l'objet d'une imputation. Les taux d'imputation pour les variables du revenu sont élevés, mais dans 45 % des cas d'imputation du revenu personnel, le répondant avait déclaré une fourchette de revenu. Ce pourcentage s'élève à 28 % dans le cas du revenu du ménage.



### Nombre et pourcentage des enregistrements dont certaines variables ont fait l'objet d'une imputation

Variable	Enregistrements imputés	Enregistrements totaux	% imputé
Revenu personnel	8 418	22 164	38
Revenu du ménage	9 131	22 164	41
Heures pour l'organisme 1	599	22 164	3
Heures pour l'organisme 2	285	22 164	1
Heures pour l'organisme 3	236	22 164	1
Heures pour les organismes 4+	188	22 164	1
Dons aux organismes 1 à 10	16 132	93 047	17
Dons aux organismes 11+	4 867	93 047	5

Le tableau suivant montre l'impact de l'imputation sur les estimations réelles.

### Pourcentage des estimations issues de valeurs imputées

Variable	Estimation imputée (millions)	Estimation totale (millions)	% imputé
Heures pour l'organisme 1	70,6	1 467,8	5
Heures pour l'organisme 2	14,6	311,7	5
Heures pour l'organisme 3	7,5	117,0	6
Heures pour les organismes 4+	7,1	86,8	8
Valeur totale des dons	1 732,3	8 882,3	20
Nombre de donateurs	0,9	22,2	4

Le processus d'imputation de l'ECDBP a donné de bons résultats et il a aidé à compléter les réponses incomplètes avec l'expérience d'autres répondants ayant des caractéristiques semblables ou identiques. Il ajoutera au nombre d'unités utilisées dans toute analyse effectuée par des chercheurs.

À noter que le fichier de microdonnées à grande diffusion ne renferme aucun indicateur d'imputation. Ce fait aura comme répercussion, une étape de confidentialité additionnelle.

## 9.2.4 Mesure de l'erreur d'échantillonnage

Puisqu'il est inévitable que des estimations établies à partir d'une enquête-échantillon (ou par sondage) soient sujettes à une erreur d'échantillonnage, une saine pratique de la statistique exige que les chercheurs fournissent aux utilisateurs une certaine indication de l'importance de cette erreur d'échantillonnage. Cette section de la documentation renferme un aperçu des mesures de l'erreur d'échantillonnage dont Statistique Canada se sert couramment et dont le Bureau conseille vivement aux utilisateurs qui produisent des estimations à partir de ce fichier de microdonnées à employer également.

La base pour mesurer l'importance potentielle des erreurs d'échantillonnage est l'erreur-type des estimations calculées à partir des résultats d'une enquête.

En raison, cependant, de la diversité des estimations pouvant être produites à partir d'une enquête, l'erreur-type d'une estimation est habituellement exprimée en fonction de

l'estimation à laquelle elle se rapporte. La mesure résultante, appelée coefficient de variation (CV) d'une estimation, s'obtient en divisant l'erreur-type de l'estimation par l'estimation elle-même et s'exprime en pourcentage de l'estimation.

Par exemple, supposons que, d'après les résultats de l'enquête, l'on estime que 54,8 % des Canadiens âgés de 15 à 24 ans avaient fait du bénévolat pendant l'année précédente et l'on constate que l'erreur-type de cette estimation est de 0,012. Le coefficient de variation de l'estimation est donc calculé comme suit :

$$\left( \frac{0,012}{0,548} \right) \times 100 \% = 2,2 \%$$

De plus amples renseignements sur le calcul du coefficient de variation, se trouvent au chapitre 11.0.

## **10.0 Lignes directrices pour la totalisation, l'analyse et la diffusion de données**

Ce chapitre de la documentation renferme un aperçu des lignes directrices que doivent respecter les utilisateurs qui totalisent, analysent, publient ou autrement diffusent des données calculées à partir des fichiers de microdonnées de l'enquête. Ces lignes directrices devraient permettre aux utilisateurs de microdonnées de produire les mêmes chiffres que ceux produits par Statistique Canada, tout en étant en mesure d'obtenir des chiffres actuellement inédits de façon conforme à ces lignes directrices établies.

### **10.1 Lignes directrices pour l'arrondissement d'estimations**

Afin que les estimations qui sont destinées à la publication ou à toute autre forme de diffusion qui sont calculées à partir de ces fichiers de microdonnées correspondent à celles produites par Statistique Canada, nous conseillons vivement aux utilisateurs de respecter les lignes directrices qui suivent en ce qui concerne l'arrondissement de telles estimations :

- a) Les estimations dans le corps principal d'un tableau statistique doivent être arrondies à la centaine près à l'aide de la technique d'arrondissement normale. 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, le dernier chiffre à conserver est augmenté de 1. Par exemple, selon la technique d'arrondissement normale à la centaine près, si les deux derniers chiffres se situent entre 00 et 49, ils sont remplacés par 00 et le chiffre précédent (le chiffre des centaines) reste inchangé. Si les derniers chiffres se situent entre 50 et 99, ils sont remplacés par 00 et le chiffre précédent est augmenté de 1.
- b) Les totaux partiels marginaux et les totaux marginaux des tableaux statistiques doivent être calculés à partir de leurs composantes non arrondies correspondantes, puis ensuite être arrondis à leur tour à la centaine près à l'aide de la technique d'arrondissement normale.
- c) Les moyennes, les proportions, les taux et les pourcentages doivent être calculés à partir de composantes non arrondies (c'est-à-dire des numérateurs et/ou des dénominateurs), puis être arrondis à leur tour à une décimale à l'aide de la technique d'arrondissement normale. Dans le cas d'un arrondissement normal à un seul chiffre, si le dernier 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, le dernier chiffre à conserver est augmenté de 1.
- d) Les sommes et les différences d'agrégats (ou de rapports) doivent être calculées à partir de leurs composantes non arrondies correspondantes, puis être arrondies à leur tour à la centaine près (ou à la décimale près) à l'aide de la technique d'arrondissement normale.
- e) Dans les cas, où, en raison de limitations d'ordre technique ou de toutes autres limites, une technique d'arrondissement autre que la technique normale est utilisée produisant des estimations à être publiées ou autrement diffusées différentes des estimations correspondantes publiées par Statistique Canada, nous conseillons vivement aux utilisateurs d'indiquer la raison de ces différences dans le ou les documents à publier ou à diffuser.
- f) En aucun cas, les utilisateurs ne doivent publier ou autrement diffuser des estimations non arrondies. Des estimations non arrondies laissent entendre qu'elles sont plus précises qu'elles ne le sont en réalité.

## **10.2 Lignes directrices pour la pondération de l'échantillon en vue de la totalisation**

Le plan d'échantillonnage utilisé pour l'Enquête canadienne sur le don, le bénévolat et la participation (ECDPB) n'était pas autopondéré. Lorsqu'ils produisent des estimations simples, y compris des tableaux statistiques ordinaires, les utilisateurs doivent appliquer le poids d'échantillonnage approprié.

Si les poids appropriés ne sont pas utilisés, les estimations calculées à partir des fichiers de microdonnées ne peuvent être considérées comme représentatives de la population visée par l'enquête et ne correspondront pas à celles produites par Statistique Canada.

Les utilisateurs devraient également prendre note que certains progiciels pourraient peut-être ne pas permettre la production d'estimations correspondant exactement à celles qu'offre Statistique Canada, en raison du mode de traitement du champ du poids par ces progiciels.

## **10.3 Définitions de types d'estimations : catégoriques et quantitatives**

Avant de discuter de la façon dont on peut totaliser et analyser les données de l'ECDBP il est utile de décrire les deux principaux types d'estimations ponctuelles des caractéristiques de la population qui peuvent être produites à partir du fichier de microdonnées créé pour l'ECDBP.

### **10.3.1 Estimations catégoriques**

Les estimations catégoriques sont des estimations du nombre ou du pourcentage de membres de la population visée par l'enquête possédant certaines caractéristiques ou faisant partie d'une catégorie définie. Le nombre de Canadiens qui ont fait du bénévolat ou le nombre de Canadiens qui ont fait des dons constituent des exemples de telles estimations. Une estimation du nombre de personnes possédant une certaine caractéristique peut aussi être désignée une estimation d'un agrégat.

#### Exemples de questions catégoriques :

- Q: Au cours des 12 derniers mois, avez-vous réalisé l'une des activités suivantes sans rémunération pour le compte d'un groupe ou d'un organisme? Veuillez inclure tout aide non rémunérée apportée dans les écoles, les organismes religieux, les associations sportives et communautaires. Avez-vous fait les activités suivantes : ...enseigné, donné de la formation ou agi à titre de mentor?
- R: Oui / Non
- Q: Au cours des 12 derniers mois, avez-vous fait un don en espèces : ... en réponse à une sollicitation effectuée par la poste?
- R: Oui / Non

### **10.3.2 Estimations quantitatives**

Les estimations quantitatives sont des estimations de totaux ou de moyennes, de médianes et d'autres mesures d'une tendance centrale de quantités reposant sur certains ou sur tous les membres de la population visée par l'enquête. Elles comprennent aussi expressément des estimations de la forme  $\hat{X} / \hat{Y}$  où  $\hat{X}$  est une estimation de la quantité totale pour la population visée par l'enquête et  $\hat{Y}$  est une

estimation du nombre de personnes dans la population visée par l'enquête qui contribuent à cette quantité totale.

Un exemple d'estimation quantitative est le nombre moyen d'heures consacrées au bénévolat par des bénévoles. Le numérateur est une estimation du nombre total d'heures consacrées au bénévolat et son dénominateur est le nombre de bénévoles.

Exemples de questions quantitatives :

Q: Au cours des 12 derniers mois, combien d'heures avez-vous consacrées à toutes les activités non rémunérées pour cet organisme?

R: |\_|\_|\_|\_| heures

Q: Quel a été le montant versé à cet organisme?

R: |\_|\_|\_|\_| dollars

### **10.3.3 Totalisation d'estimations catégoriques**

On peut obtenir des estimations du nombre de gens possédant une certaine caractéristique à partir du fichier de microdonnées en additionnant les poids finals de tous les enregistrements possédant la ou les caractéristiques qui nous intéressent. On obtient des proportions et des rapports de la forme  $\hat{X} / \hat{Y}$  en

- additionnant les poids finals des enregistrements présentant la caractéristique qui nous intéresse pour le numérateur ( $\hat{X}$ );
- additionnant les poids finals des enregistrements présentant la caractéristique qui nous intéresse pour le dénominateur ( $\hat{Y}$ ); puis en
- divisant l'estimation a) par celle de b) ( $\hat{X} / \hat{Y}$ ).

### **10.3.4 Totalisation d'estimations quantitatives**

On peut obtenir des estimations de quantités à partir du fichier de microdonnées en multipliant la valeur de la variable qui nous intéresse par le poids final de chaque enregistrement, puis en additionnant cette quantité pour tous les enregistrements qui nous intéressent. Pour obtenir, par exemple, une estimation du nombre total d'heures de bénévolat consacrées par les personnes âgées de 65 ans et plus, multipliez la valeur déclarée à la VD1DHRS (heures de bénévolat) par le poids final de l'enregistrement, puis additionnez cette valeur pour tous les enregistrements où la variable DH1GAGE = 6 (groupe d'âge de 65 ans et plus).

Pour obtenir une moyenne pondérée de la forme  $\hat{X} / \hat{Y}$ , le numérateur ( $\hat{X}$ ) est calculé comme une estimation quantitative et le dénominateur ( $\hat{Y}$ ) est calculé comme une estimation catégorique. Pour estimer, par exemple, le nombre moyen d'heures de bénévolat consacrées par des personnes âgées de 65 ans et plus,

- estimez le nombre total d'heures de bénévolat ( $\hat{X}$ ) tel qu'il est décrit ci-dessus;
- estimez le nombre de personnes ( $\hat{Y}$ ) incluses dans cette catégorie en additionnant les poids finals de tous les enregistrements où la variable DH1GAGE = 6; puis
- divisez l'estimation a) par l'estimation b) ( $\hat{X} / \hat{Y}$ ).

## **10.4 Lignes directrices pour l'analyse statistique**

L'ECDBP repose sur un plan d'échantillonnage complexe comportant une stratification, de multiples étapes de sélection ainsi que des probabilités inégales de sélection des répondants. L'utilisation des données provenant d'enquêtes aussi complexes présente des problèmes pour les analystes, parce que le plan d'enquête et les probabilités de sélection influent sur les procédures d'estimation et de calcul de la variance qui devraient être utilisées. Il faut utiliser les poids de l'enquête pour que les estimations et les analyses des données de l'enquête soient exemptes de biais.

Bien que de nombreuses procédures d'analyse que l'on trouve à l'intérieur de progiciels statistiques permettent d'utiliser des poids, la signification ou la définition du poids inclus dans ces procédures peut différer de ce qui convient dans le contexte d'une enquête-échantillon, de telle sorte que dans bien des cas les estimations produites au moyen de ces progiciels sont correctes, mais que les variances calculées sont piètres. Les variances approximatives pour des estimations simples comme des totaux, des proportions et des rapports (pour des variables qualitatives) peuvent être calculées à partir des tables de variabilité d'échantillonnage approximative qui accompagnent les données.

Pour d'autres techniques d'analyse (de régression linéaire, de régression logistique et de l'analyse de variance, par exemple), il existe une méthode qui peut rendre les variances calculées par l'application des progiciels normalisés plus significatives, en intégrant les probabilités inégales de sélection. L'application de cette méthode entraîne une remise à l'échelle des poids de façon à ce que le poids moyen soit de un.

Supposons, par exemple, qu'il faut effectuer l'analyse de tous les répondants de sexe masculin. Les étapes à suivre pour remettre à l'échelle les poids sont les suivantes :

- 1) sélectionner tous les répondants du fichier qui ont déclaré RESPSEX = homme;
- 2) calculer le poids MOYEN pour ces enregistrements en additionnant les poids originaux des personnes établis à partir du fichier de microdonnées pour ces enregistrements puis diviser cette somme par le nombre de répondants ayant déclaré RESPSEX = homme;
- 3) pour chacun de ces répondants, calculer un poids REMIS À L'ÉCHELLE égal au poids original de la personne divisé par le poids MOYEN;
- 4) effectuer l'analyse portant sur ces répondants en utilisant le poids REMIS À L'ÉCHELLE.

Parce qu'on ne tient toujours compte ni de la stratification ni des grappes du plan d'échantillonnage, les estimations des variances calculées avec cette méthode risquent cependant d'être des sous-estimations.

Il faut connaître les détails du plan d'enquête pour calculer des estimations des variances plus précises. De tels détails ne peuvent être fournis dans le fichier de microdonnées en raison de la confidentialité. Statistique Canada peut, contre remboursement des frais, calculer des variances qui tiennent compte du plan complet d'échantillonnage pour beaucoup de statistiques.

## **10.5 Lignes directrices pour la diffusion de coefficients de variation**

Avant de diffuser et/ou de publier toutes estimations établies à partir de l'ECDBP, les utilisateurs devraient premièrement déterminer le niveau de qualité de cette estimation. Les niveaux de qualité sont *acceptable*, *médiocre* et *inacceptable*. Les erreurs d'échantillonnage et non dues à l'échantillonnage, dont il a été question au chapitre 9.0, influencent la qualité des données. Aux fins du présent document, cependant, on ne déterminera le niveau de qualité d'une estimation

qu'à partir d'une erreur d'échantillonnage dont rend compte le coefficient de variation indiqué à l'intérieur du tableau qui figure ci-dessous. Les utilisateurs devraient néanmoins s'assurer de lire le chapitre 9.0 pour être plus pleinement informés des caractéristiques relatives à la qualité de ces données.

On devrait premièrement déterminer le nombre de répondants retenus pour le calcul de l'estimation. Si ce nombre est inférieur à 30, il faudrait considérer l'estimation pondérée comme étant de qualité inacceptable

Pour les estimations pondérées fondées sur les tailles d'échantillon de 30 ou plus, les utilisateurs devraient déterminer le coefficient de variation de l'estimation et suivre les lignes directrices relatives au niveau de qualité qui figurent ci-dessous. Celles-ci devraient être appliquées, pour la détermination du niveau de qualité d'une estimation, aux estimations pondérées arrondies.

On peut considérer qu'il est possible de divulguer toutes les estimations. Celles d'un niveau de qualité médiocre ou inacceptable doivent cependant être accompagnées d'une mise en garde pour avertir les utilisateurs subséquents.

### Lignes directrices relatives au niveau de qualité de l'estimation

Niveau de qualité de l'estimation	Lignes directrices
1) Acceptable	<p>Les estimations proviennent d'une taille d'échantillon de 30 ou plus, et présentent de faibles coefficients de variation, de l'ordre de 0,0 à 16,5 %.</p> <p>Aucune mise en garde n'est requise.</p>
2) Médiocre	<p>Les estimations proviennent d'une taille d'échantillon de 30 ou plus, et présentent des coefficients de variation élevés, de l'ordre de 16,6 à 33,3 %.</p> <p>Ces estimations devraient être signalées par la lettre M (ou un quelconque identificateur similaire). Elles devraient être accompagnées d'une mise en garde avertissant les utilisateurs subséquents des niveaux élevés d'erreur associés aux estimations.</p>
3) Inacceptable	<p>Les estimations proviennent d'une taille d'échantillon inférieure à 30, ou présentent des coefficients de variation très élevés, supérieurs à 33,3 %.</p> <p>Statistique Canada recommande de ne pas diffuser d'estimations de qualité inacceptable. Si un utilisateur choisit cependant de le faire, ces estimations devraient alors être signalées à l'aide de la lettre I (ou d'un quelconque identificateur similaire) et devraient être accompagnées de la mise en garde suivante :</p> <p>« Nous informons l'utilisateur que ces estimations (désignées avec la lettre I) ne respectent pas les normes de qualité de Statistique Canada. Les conclusions qui reposeront sur ces données ne seront pas fiables et seront très probablement invalides. »</p>

## 10.6 Seuils pour la diffusion des estimations pour l'Enquête canadienne sur le don, le bénévolat et la participation

Le tableau ci-dessous fournit une indication de la précision des estimations des prévisions démographiques ainsi que les seuils de diffusion associés aux trois niveaux de qualité de l'estimation présentés à la section précédente. Ces seuils proviennent des tables de coefficients de variation (CV) dont il sera question au chapitre 11.0.

Par exemple, d'après le tableau, la qualité d'une estimation pondérée de 10 000 personnes possédant une caractéristique donnée à Terre-Neuve-et-Labrador est médiocre.

Veillez noter que ces seuils de diffusion correspondent aux estimations de chiffres de population seulement. Dans le cas d'estimations de rapports, les utilisateurs ne devraient pas utiliser la valeur du numérateur (ni le dénominateur) afin de trouver le niveau de qualité de l'estimation correspondant. La règle 4 à la section 11.1 ainsi que l'exemple 4 à la section 11.1.1 expliquent la bonne procédure à suivre dans le cas d'un rapport.

Province / territoires	CV acceptable 0,0 à 16,5 %	CV médiocre 16,6 à 33,3 %	CV inacceptable > 33,3 %
Terre-Neuve-et-Labrador	15 000 et plus	4 000 à < 15 000	moins de 4 000
Île-du-Prince-Édouard	6 000 et plus	1 500 à < 6 000	moins de 1 500
Nouvelle-Écosse	22 500 et plus	5 500 à < 22 500	moins de 5 500
Nouveau-Brunswick	19 500 et plus	5 000 à < 19 500	moins de 5 000
Québec	113 000 et plus	28 000 à < 113 000	moins de 28 000
Ontario	130 000 et plus	32 000 à < 130 000	moins de 32 000
Manitoba	24 500 et plus	6 000 à < 24 500	moins de 6 000
Saskatchewan	25 000 et plus	6 000 à < 25 000	moins de 6 000
Alberta	73 000 et plus	18 500 à < 73 000	moins de 18 500
Colombie-Britannique	60 000 et plus	15 000 à < 60 000	moins de 15 000
<b>Provinces</b>	<b>96 500 et plus</b>	<b>23 500 à &lt; 96 500</b>	<b>moins de 23 500</b>
Territoires	3 500 et plus	1 000 à < 3 500	moins de 1 000
<b>Canada</b>	<b>96 000 et plus</b>	<b>23 500 à &lt; 96 000</b>	<b>moins de 23 500</b>



## 11.0 Tables de variabilité d'échantillonnage approximative

Afin de fournir des coefficients de variation (CV) qui pourraient s'appliquer à une gamme étendue d'estimations catégoriques produites à partir de ce fichier de microdonnées et auxquels il serait facilement possible pour l'utilisateur d'avoir accès, un ensemble de tables de variabilité d'échantillonnage approximative a été produit. Ces tables de CV permettent à l'utilisateur d'obtenir un coefficient de variation approximatif fondé sur la taille de l'estimation calculée à partir des données de l'enquête.

Les coefficients de variation sont calculés à l'aide de la formule de la variance pour un échantillonnage aléatoire simple et en y incorporant un facteur qui reflète la nature du plan d'échantillonnage, qui est à plusieurs degrés et qui prévoit la formation de grappes. Ce facteur, appelé l'effet du plan, a été déterminé en calculant premièrement les effets du plan pour une gamme étendue de caractéristiques, puis en choisissant parmi ceux-ci une valeur modérée (habituellement le 75<sup>e</sup> percentile) à utiliser à l'intérieur des tables de CV qui s'appliqueraient ensuite à l'ensemble entier des caractéristiques.

Le tableau ci-dessous indique la valeur modérée des effets du plan, ainsi que les tailles de l'échantillon et les chiffres de population selon la province qui ont été utilisés pour produire les tables de variabilité d'échantillonnage approximative de l'Enquête canadienne sur le don, le bénévolat et la participation (ECDBP) de 2004.

Province / territoires	Effet du plan	Taille de l'échantillon	Population
Terre-Neuve-et-Labrador	1,37	1 407	440 863
Île-du-Prince-Édouard	1,39	936	115 184
Nouvelle-Écosse	1,31	1 612	779 570
Nouveau-Brunswick	1,33	1 510	622 946
Québec	1,49	2 948	6 211 020
Ontario	1,45	4 071	10 068 734
Manitoba	1,37	1 834	921 621
Saskatchewan	1,49	1 688	789 055
Alberta	1,44	1 807	2 573 431
Colombie-Britannique	1,44	3 019	3 498 788
<b>Provinces</b>	<b>2,11</b>	<b>20 832</b>	<b>26 021 212</b>
Territoires	1,83	1 332	71 962
<b>Canada</b>	<b>2,23</b>	<b>22 164</b>	<b>26 093 174</b>

Tous les coefficients de variation inclus dans les tables de variabilité d'échantillonnage approximative sont approximatifs et donc non officiels. Des estimations de la variance réelle pour des variables précises peuvent être obtenues auprès de Statistique Canada, contre remboursement des frais. Étant donné que le CV approximatif est une estimation prudente, l'utilisation de la variance réelle estimée pourrait faire passer l'estimation d'un niveau de qualité à un autre. Par exemple, une estimation *médiocre* pourrait devenir *acceptable* si elle était fondée sur le calcul du CV exact.

Rappelez-vous que : Si le nombre d'observations sur lesquelles une estimation est basée est inférieur à 30, l'estimation pondérée est très probablement inacceptable et Statistique Canada recommande de ne pas diffuser une telle estimation, quelle que soit la valeur du coefficient de variation.

## 11.1 Comment utiliser les tables de coefficients de variation pour les estimations catégoriques

Les règles qui suivent devraient permettre à l'utilisateur de déterminer les coefficients de variation approximatifs à partir des tables de variabilité d'échantillonnage approximative pour des estimations du nombre, de la proportion ou du pourcentage de membres de la population visée par l'enquête possédant une certaine caractéristique et pour des rapports et des différences entre de telles estimations.

### Règle 1 : Estimations du nombre de personnes possédant une caractéristique donnée (agrégats)

Le coefficient de variation dépend uniquement de la taille de l'estimation elle-même. Dans le table de variabilité d'échantillonnage approximative pour la région géographique appropriée, repérez le nombre estimé dans la colonne la plus à gauche (intitulée « Numérateur du pourcentage ») et suivez les astérisques (le cas échéant) jusqu'au premier chiffre rencontré. Ce chiffre est le coefficient de variation approximatif.

### Règle 2 : Estimations de proportions ou de pourcentages de personnes possédant une caractéristique donnée

Le coefficient de variation d'une proportion estimée ou d'un pourcentage estimé dépend à la fois de la taille de la proportion ou du pourcentage et de la taille du total sur lequel la proportion ou le pourcentage repose. Les proportions estimées ou les pourcentages estimés sont relativement plus fiables que les estimations correspondantes du numérateur de la proportion ou du pourcentage, lorsque la proportion ou que le pourcentage repose sur un sous-groupe de la population. La proportion, par exemple, de bénévoles qui ont assuré la prestation de soins de santé y compris de l'accompagnement est plus fiable que le nombre estimé de bénévoles qui ont assuré la prestation de soins de santé y compris de l'accompagnement. (Remarquez que dans les tables, la valeur des coefficients de variation diminue lorsqu'on les lit de gauche à droite.)

Lorsque la proportion ou que le pourcentage repose sur la population totale de la région géographique visée par le table, le CV de la proportion ou du pourcentage est le même que le CV du numérateur de la proportion ou du pourcentage. Dans ce cas, la règle 1 peut être appliquée.

Lorsque la proportion ou que le pourcentage repose sur un sous-ensemble de la population totale (p. ex. comme ses membres d'un sexe ou d'un groupe d'âge particulier), on devrait faire référence à la proportion ou au pourcentage (dans le haut de la table) et au numérateur de la proportion ou du pourcentage (dans la colonne de gauche de la table). L'intersection de la rangée et de la colonne appropriées donne le coefficient de variation.

### Règle 3 : Estimations de différences entre des agrégats ou des pourcentages

L'erreur-type d'une différence entre deux estimations est approximativement égale à la racine carrée de la somme des carrés de chaque erreur-type considérée séparément. C'est-à-dire que l'erreur-type d'une différence  $(\hat{d} = \hat{X}_1 - \hat{X}_2)$  est

$$\sigma_{\hat{d}} \sqrt{(\hat{X}_1 \alpha_1)^2 + (\hat{X}_2 \alpha_2)^2}$$

où  $\hat{X}_1$  est l'estimation 1,  $\hat{X}_2$  est l'estimation 2 et  $\alpha_1$  et  $\alpha_2$  sont les coefficients de variation de  $\hat{X}_1$  et  $\hat{X}_2$  respectivement. Le coefficient de variation de  $\hat{d}$  est donné par  $\sigma_{\hat{d}}/\hat{d}$ . Cette formule est exacte pour la différence entre des caractéristiques distinctes et non corrélées, mais n'est autrement qu'approximative.

**Règle 4 : Estimations de rapports**

Dans le cas où le numérateur est un sous-ensemble du dénominateur, le rapport devrait être converti en un pourcentage et la règle 2 appliquée. Cela s'appliquerait, par exemple, au cas où le dénominateur est le nombre de personnes avec un diplôme universitaire et le numérateur, le nombre de bénévoles avec un diplôme universitaire.

Dans le cas où le numérateur n'est pas un sous-ensemble du dénominateur, comme dans l'exemple du rapport du nombre de bénévoles avec un diplôme universitaire comparativement au nombre de bénévoles sans diplôme universitaire, l'erreur-type du rapport des estimations est approximativement égale à la racine carrée de la somme des carrés de chaque coefficient de variation considéré séparément multipliée par  $\hat{R}$ . C'est-à-dire que l'erreur-type d'un rapport ( $\hat{R} = \hat{X}_1 / \hat{X}_2$ ) est

$$\sigma_{\hat{R}} = \hat{R} \sqrt{\alpha_1^2 + \alpha_2^2}$$

où  $\alpha_1$  et  $\alpha_2$  sont les coefficients de variation de  $\hat{X}_1$  et de  $\hat{X}_2$  respectivement. Le coefficient de variation de  $\hat{R}$  est donné par  $\sigma_{\hat{R}} / \hat{R}$ . La formule tendra à surestimer l'erreur si  $\hat{X}_1$  et  $\hat{X}_2$  sont corrélés positivement et à la sous-estimer si  $\hat{X}_1$  et  $\hat{X}_2$  sont corrélés négativement.

**Règle 5 : Estimations de différences entre des rapports**

Dans ce cas, les règles 3 et 4 sont combinées. On détermine premièrement les CV pour les deux rapports à l'aide de la règle 4, puis on trouve le CV de leur différence au moyen de la règle 3.

**11.1.1 Exemples d'utilisation des tables de coefficients de variation pour des estimations catégoriques**

Les exemples ci-dessous utilisent des données du fichier de l'ECDBP de 2004 et sont destinés à aider les utilisateurs à appliquer les règles que nous venons de présenter.

**Exemple 1 : Estimations du nombre de personnes possédant une caractéristique donnée (agrégats)**

Supposons qu'un utilisateur estime que 5 615 215 d'hommes étaient bénévoles durant la période de référence. Comment l'utilisateur détermine-t-il le coefficient de variation de cette estimation?

- 1) Reportez-vous à la table des coefficients de variation pour le CANADA.
- 2) L'agrégat estimé (5 615 215) ne figure pas dans la colonne de gauche (la colonne « Numérateur du pourcentage »); il faut donc utiliser le chiffre qui s'en rapproche le plus, c'est-à-dire 6 000 000.
- 3) On trouve le coefficient de variation pour un agrégat estimé en se reportant à la première entrée autre que des astérisques sur cette rangée, c'est-à-dire 1,8 %.
- 4) Le coefficient de variation approximatif de l'estimation est donc 1,8 %. Le résultat selon lequel il y avait 5 615 215 (à être arrondi selon les lignes directrices pour

l'arrondissement figurant à la section 10.1) d'hommes qui étaient bénévoles durant la période de référence peut être publié sans réserve.

**Exemple 2 : Estimations de proportions ou de pourcentages de personnes possédant une caractéristique donnée**

Supposons qu'un utilisateur estime que  $1\ 605\ 006 / 5\ 615\ 215 = 28,6\ %$  la proportion d'hommes qui ont fait du bénévolat ayant enseigné, donné de la formation ou agi à titre de mentor. Comment l'utilisateur détermine-t-il le coefficient de variation de cette estimation?

- 1) Reportez-vous à la table des coefficients de variation pour le CANADA.
- 2) Parce que l'estimation est un pourcentage fondé sur un sous-ensemble de la population totale (c'est-à-dire les hommes qui étaient bénévoles), il faut utiliser à la fois le pourcentage (28,6 %) et la portion numérateur du pourcentage (1 605 006) pour déterminer le coefficient de variation.
- 3) Le numérateur, 1 605 006, ne figure pas dans la colonne de gauche (la colonne « Numérateur du pourcentage »); il faut donc utiliser le chiffre qui s'en rapproche le plus, soit 1 500 000. De même, l'estimation du pourcentage ne figure dans l'en-tête d'aucune colonne; il faut donc utiliser la proportion qui s'en rapproche le plus, soit 30,0 %.
- 4) Le chiffre indiqué à l'intersection de la rangée et de la colonne utilisées, soit 3,5 % est le coefficient de variation à employer.
- 5) Le coefficient de variation approximatif de l'estimation est donc 3,5 %. Le résultat selon lequel 28,6 % d'hommes qui ont fait du bénévolat ont enseigné, donné de la formation ou agi à titre de mentor peut être publié sans réserve.

**Exemple 3 : Estimations de différences entre des agrégats ou des pourcentages**

Supposons qu'un utilisateur estime que  $1\ 979\ 228 / 6\ 193\ 361 = 32,0\ %$  la proportion de femmes qui ont fait du bénévolat ayant enseigné, donné de la formation ou agi à titre de mentor, par rapport à  $1\ 605\ 006 / 5\ 615\ 215 = 28,6\ %$  la proportion d'hommes qui ont fait du bénévolat ayant enseigné, donné de la formation ou agi à titre de mentor. Comment l'utilisateur détermine-t-il le coefficient de variation de la différence entre ces deux estimations?

- 1) L'utilisation de la table des coefficients de variation du CANADA de la même façon que celle décrite dans l'exemple 2, donne un CV de l'estimation pour les femmes de 3,0 %, et un CV de l'estimation pour les hommes de 3,5 %.
- 2) En utilisant la règle 3, l'erreur-type d'une différence ( $\hat{d} = \hat{X}_1 - \hat{X}_2$ ) est

$$\sigma_{\hat{d}} = \sqrt{(\hat{X}_1 \alpha_1)^2 + (\hat{X}_2 \alpha_2)^2}$$

où  $\hat{X}_1$  est l'estimation 1 (femmes),  $\hat{X}_2$  est l'estimation 2 (hommes), et  $\alpha_1$  et  $\alpha_2$  sont les coefficients de variation de  $\hat{X}_1$  et de  $\hat{X}_2$  respectivement.

C'est-à-dire que l'erreur-type de la différence  $\hat{d} = 0,320 - 0,286 = 0,034$  est

$$\begin{aligned}\sigma_{\hat{d}} &= \sqrt{[(0,320)(0,030)]^2 + [(0,286)(0,035)]^2} \\ &= \sqrt{(0,00009216) + (0,0001002)} \\ &= 0,014\end{aligned}$$

- 3) Le coefficient de variation de  $\hat{d}$  est donné par  $\sigma_{\hat{d}} / \hat{d} = 0,014 / 0,034 = 0,412$
- 4) Le coefficient de variation approximatif de la différence entre les estimations est donc 41,2 %. La différence entre les estimations est considérée inacceptable et Statistique Canada recommande de ne pas publier cette estimation. Cependant, si l'utilisateur choisit de publier cette donnée, elle devra être désignée ainsi en utilisant la lettre I (ou un autre identificateur semblable) et être accompagnée d'un avertissement mettant les utilisateurs subséquents en garde contre les hauts taux d'erreur associés à l'estimation.

#### Exemple 4 : Estimations de rapports

Supposons qu'un utilisateur estime que 1 979 228 de femmes qui ont fait du bénévolat ayant enseigné, donné de la formation ou agi à titre de mentor par l'entremise d'un organisme par rapport à 1 605 006 d'hommes qui ont fait du bénévolat ayant enseigné, donné de la formation ou agi à titre de mentor. L'utilisateur est intéressé à comparer l'estimation des femmes à celle des hommes sous la forme d'un rapport. Comment l'utilisateur détermine-t-il le coefficient de variation de cette estimation?

- 1) Tout d'abord, cette estimation est une estimation d'un rapport, où le numérateur de l'estimation ( $\hat{X}_1$ ) est le nombre de bénévoles de sexe féminin ayant enseigné, donné de la formation ou agi à titre de mentor par l'entremise d'un organisme. Le dénominateur de l'estimation ( $\hat{X}_2$ ) est le nombre de bénévoles de sexe masculin ayant enseigné, donné de la formation ou agi à titre de mentor.
- 2) Reportez-vous à la table des coefficients de variation pour le CANADA.
- 3) Le numérateur de cette estimation de rapport est 1 979 228. Le chiffre qui s'en rapproche le plus est 2 000 000. On trouve le coefficient de variation pour cette estimation en se reportant à la première entrée autre que des astérisques sur cette rangée, soit 3,4 %.
- 4) Le dénominateur de cette estimation de rapport est 1 605 006. Le chiffre qui s'en rapproche le plus est 1 500 000. On trouve le coefficient de variation pour cette estimation en se reportant à la première entrée autre que des astérisques sur cette rangée, soit 4,0 %.
- 5) Le coefficient de variation approximatif de l'estimation du rapport est donc donné par la règle 4, qui est

$$\alpha_{\hat{r}} = \sqrt{\alpha_1^2 + \alpha_2^2}$$

où  $\alpha_1$  et  $\alpha_2$  sont les coefficients de variation de  $\hat{X}_1$  et  $\hat{X}_2$  respectivement.

C'est-à-dire que

$$\begin{aligned}\alpha_{\hat{R}} &= \sqrt{(0,034)^2 + (0,040)^2} \\ &= \sqrt{0,001156 + 0,0016} \\ &= 0,052\end{aligned}$$

- 6) Le rapport obtenu entre les bénévoles de sexe féminin et les bénévoles de sexe masculin ayant enseigné, donné de la formation ou agi à titre de mentor par l'entremise d'un organisme est 1 979 228 / 1 605 006 c'est-à-dire 1,23 : 1 (à être arrondi selon les lignes directrices pour l'arrondissement figurant à la section 10.1). Le coefficient de variation de cette estimation est 5,2 %, ce qui fait qu'on peut la diffuser sans réserve.

### Exemple 5 : Estimations de différences de rapports

Supposons qu'un utilisateur estime que le rapport entre les bénévoles de sexe féminin et les bénévoles de sexe masculin est 1,039 pour les personnes âgées de 15 à 24 ans tandis qu'il est de 1,169 pour les personnes âgées de 55 ans et plus. L'utilisateur est intéressé à comparer les deux rapports pour voir s'il y a une différence statistique entre ceux-ci. Comment l'utilisateur détermine-t-il le coefficient de variation de cette différence?

- 1) Tout d'abord calculez le coefficient de variation approximatif pour le rapport du groupe d'âge de 15 à 24 ans ( $\hat{R}_1$ ) et pour le rapport du groupe d'âge de 55 ans et plus ( $\hat{R}_2$ ) tel qu'il est décrit dans l'exemple 4. Le CV approximatif pour le rapport du groupe d'âge de 15 à 24 ans est 7,07 % et 5,66 % pour le groupe d'âge de 55 ans et plus.
- 2) En utilisant la règle 3, l'erreur-type d'une différence ( $\hat{d} = \hat{R}_1 - \hat{R}_2$ ) est

$$\sigma_{\hat{d}} = \sqrt{(\hat{R}_1 \alpha_1)^2 + (\hat{R}_2 \alpha_2)^2}$$

où  $\alpha_1$  et  $\alpha_2$  sont les coefficients de variation de  $\hat{R}_1$  et  $\hat{R}_2$  respectivement. C'est-à-dire que l'erreur-type de la différence  $\hat{d} = 1,039 - 1,169 = -0,13$  est

$$\begin{aligned}\sigma_{\hat{d}} &= \sqrt{[(1,039)(0,0707)]^2 + [(1,169)(0,0566)]^2} \\ &= \sqrt{(0,005396) + (0,004378)} \\ &= 0,099\end{aligned}$$

- 3) Le coefficient de variation de  $\hat{d}$  est donné par  $\sigma_{\hat{d}} / \hat{d} = 0,099 / (-0,13) = -0,762$ .
- 4) Le coefficient de variation approximatif de la différence entre les estimations est donc 76,2 %. L'estimation de la différence entre les deux estimations est considérée inacceptable et Statistique Canada recommande de ne pas la publier. Cependant, si l'utilisateur choisit de publier cette donnée, elle devra être désignée ainsi en utilisant la lettre I (ou un autre identificateur semblable) et être accompagnée d'un avertissement mettant les utilisateurs subséquents en garde contre les hauts taux d'erreur associés à l'estimation.

## 11.2 Comment utiliser les tables de coefficients de variation pour obtenir des limites de confiance

Bien que les coefficients de variation soient beaucoup utilisés, l'intervalle de confiance d'une estimation est une mesure plus intuitivement significative de l'erreur d'échantillonnage. Un intervalle de confiance constitue une déclaration du niveau de confiance selon laquelle la valeur vraie pour la population se situe à l'intérieur d'une gamme précisée de valeurs. Par exemple, un intervalle de confiance de 95 % peut être décrit comme suit :

Si l'échantillonnage de la population est répété indéfiniment, chaque échantillon menant à un nouvel intervalle de confiance pour une estimation, l'intervalle englobera alors dans 95 % des échantillons la valeur vraie de la population.

En utilisant l'erreur-type d'une estimation, des intervalles de confiance pour des estimations peuvent être obtenues en partant de l'hypothèse qu'aux termes d'un échantillonnage répété de la population, les diverses estimations obtenues pour une caractéristique donnée de la population se répartiront normalement autour de la valeur vraie de la population. Selon cette hypothèse, il y a environ 68 chances sur 100 que l'écart entre une estimation de l'échantillon et la valeur vraie pour la population soit inférieur à une erreur-type, environ 95 chances sur 100 que l'écart soit inférieur à deux erreurs-types et environ 99 chances sur 100 que l'écart soit inférieur à trois erreurs-types. Ces différents degrés de confiance sont désignés sous le nom de niveaux de confiance.

Des intervalles de confiance pour une estimation,  $\hat{X}$ , sont généralement exprimés sous forme de deux chiffres, un inférieur et un supérieur à l'estimation, comme étant  $(\hat{X} - k, \hat{X} + k)$  où  $k$  est déterminé suivant le niveau de confiance désiré et l'erreur d'échantillonnage de l'estimation.

Des intervalles de confiance pour une estimation peuvent être calculés directement à partir des tables de variabilité d'échantillonnage approximative, en déterminant d'abord à partir de la table appropriée le coefficient de variation de l'estimation  $\hat{X}$ , puis en utilisant la formule suivante pour le convertir à un intervalle de confiance ( $CI_{\hat{X}}$ ) :

$$CI_{\hat{X}} = (\hat{X} - t\hat{X}\alpha_{\hat{X}}, \hat{X} + t\hat{X}\alpha_{\hat{X}})$$

où  $\alpha_{\hat{X}}$  est le coefficient de variation déterminé de  $\hat{X}$ , et

- $t = 1$  si l'on désire un intervalle de confiance de 68 %;
- $t = 1,6$  si l'on désire un intervalle de confiance de 90 %;
- $t = 2$  si l'on désire un intervalle de confiance de 95 %;
- $t = 2,6$  si l'on désire un intervalle de confiance de 99 %.

**Nota :** Les lignes directrices pour la diffusion des estimations s'appliquent également aux intervalles de confiance. S'il est impossible, par exemple, de diffuser une estimation, on ne peut alors pas non plus communiquer un intervalle de confiance.

### 11.2.1 Exemple d'utilisation des tables de coefficients de variation pour obtenir des limites de confiance

Un intervalle de confiance de 95 % pour la proportion estimée de bénévoles de sexe masculin qui ont fait du bénévolat ayant enseigné, donné de la formation ou agi à titre de mentor (d'après l'exemple 2 à la section 11.1.1) serait calculé comme suit :

$$\hat{X} = 28,6 \% \text{ (ou exprimé sous forme de proportion } 0,286)$$

$$t = 2$$

$\alpha_{\hat{x}}$  = 3,5 % (0,035 exprimé sous forme de proportion) est le coefficient de variation de cette estimation, tel que déterminé à partir des tables.

$$CI_{\hat{x}} = \{0,286 - (2) (0,286) (0,035), 0,286 + (2) (0,286) (0,035)\}$$

$$CI_{\hat{x}} = \{0,286 - 0,020, 0,286 + 0,020\}$$

$$CI_{\hat{x}} = \{0,266, 0,306\}$$

Avec un intervalle de confiance de 95 %, on peut dire qu'entre 26,6 % et 30,6 % des bénévoles de sexe masculin ont enseigné, donné de la formation ou agi à titre de mentor.

### 11.3 Comment utiliser les tables de coefficients de variation pour effectuer un test t

Des erreurs-types peuvent aussi être utilisés pour effectuer des tests d'hypothèses, une procédure destinée à distinguer des paramètres d'une population à l'aide d'estimations d'un échantillon. Ces estimations peuvent être des chiffres, des moyennes, des pourcentages, des rapports, etc. Les tests peuvent être effectués à divers niveaux de signification, où un niveau de signification est la probabilité de conclure que les caractéristiques sont différentes quand, en fait, elles sont identiques.

Supposons que  $\hat{X}_1$  et  $\hat{X}_2$  sont des estimations d'un échantillon pour deux caractéristiques qui nous intéressent. Supposons également que l'erreur-type de la différence  $\hat{X}_1 - \hat{X}_2$  et  $\sigma_{\hat{d}}$ .

Si  $t = \frac{\hat{X}_1 - \hat{X}_2}{\sigma_{\hat{d}}}$  se situe entre -2 et 2, aucune conclusion à propos de la différence entre les

caractéristiques n'est alors justifiée au niveau de signification de 5 %. Si, cependant, ce rapport est inférieur à -2 ou supérieur à +2, la différence observée est significative au niveau de 0,05. C'est-à-dire que la différence entre les estimations est significative.



### **11.3.1 Exemple d'utilisation des tables de coefficients de variation pour effectuer un test *t***

Supposons que l'utilisateur désire tester, au niveau de signification de 5 %, l'hypothèse selon laquelle il n'y a pas de différence entre la proportion de bénévoles de sexe féminin ayant enseigné, donné de la formation ou agi à titre de mentor et la proportion de bénévoles de sexe masculin travailleurs rémunérés de sexe masculin ayant enseigné, donné de la formation ou agi à titre de mentor. D'après l'exemple 3 à la section 11.1.1, il s'est avéré que l'erreur-type de la différence entre ces deux estimations était 0,014. Par conséquent,

$$t = \frac{\hat{X}_1 - \hat{X}_2}{\sigma_{\hat{d}}} = \frac{0,320 - 0,286}{0,014} = \frac{0,034}{0,014} = 2,43$$

Puisque  $t = 2,43$  est supérieur à 2, il faut en conclure qu'il existe une différence significative entre les deux estimations au niveau de signification de 0,05.

## **11.4 Coefficients de variation pour des estimations quantitatives**

Il faudrait produire des tables spéciales afin de déterminer l'erreur d'échantillonnage d'estimations quantitatives, ce qui n'a pas été fait, parce que la plupart des variables pour l'ECDBP sont principalement de nature catégorique.

En général cependant, le coefficient de variation d'un total quantitatif sera supérieur au coefficient de variation de l'estimation de la catégorie correspondante (c'est-à-dire l'estimation du nombre de personnes retenues dans l'estimation quantitative). S'il est impossible de diffuser l'estimation de la catégorie correspondante, on ne pourra pas non plus communiquer l'estimation quantitative. Par exemple, le coefficient de variation du nombre d'heures de bénévolat consacrées à des organismes d'arts et de culture serait supérieur au coefficient de variation de la proportion correspondante de bénévoles qui ont fait du bénévolat au sein d'un organisme d'arts ou de culture. Si, par conséquent, le coefficient de variation de la proportion est inacceptable (rendant la proportion non diffusable), il en sera de même du coefficient de variation de l'estimation quantitative correspondante (rendant cette estimation quantitative non diffusable).

Des coefficients de variation de telles estimations peuvent être calculés, au besoin, pour une estimation précise à l'aide d'une technique appelée pseudo-répétition, ce qui veut dire diviser les enregistrements inclus dans les fichiers de microdonnées en sous-groupes (ou répétitions) et déterminer la variation à l'intérieur de l'estimation de répétition en répétition. Les utilisateurs qui désirent calculer des coefficients de variation pour des estimations quantitatives peuvent communiquer avec Statistique Canada afin d'obtenir des conseils sur l'allocation d'enregistrements à des répétitions appropriées et sur les formules à employer à l'intérieur de ces calculs.

## **11.5 Tables des coefficients de variation**

Consulter le fichier ECDBP2004\_CVTabF.pdf pour les tables de coefficients de variation.



## 12.0 Pondération

Un poids statistique a été placé sur chaque enregistrement du fichier des microdonnées. Ce poids indique le nombre de personnes dans la population représenté par chacune des unités de l'échantillon.

L'Enquête canadienne sur le don, le bénévolat et la participation (ECDBP) de 2004 a été menée dans les dix provinces à titre d'enquête téléphonique à composition aléatoire (CA) alors que dans les trois territoires, l'ECDBP a été réalisée auprès d'un sous-échantillon de répondants à l'Enquête sur la population active (EPA). Deux procédures de pondération distinctes ont donc dû être utilisées.

### 12.1 Pondération de la composante provinciale

La pondération de la composante provinciale comporte quelques étapes :

- le calcul du poids de base du numéro de téléphone;
- un ajustement pour tenir compte des numéros de téléphone non résolus;
- l'élimination des enregistrements hors du champ d'enquête;
- un ajustement pour tenir compte du nombre de lignes téléphoniques dans le ménage;
- les ajustements pour tenir compte de la non-réponse (au niveau du ménage et au niveau de la personne);
- un ajustement pour tenir compte de la sélection d'une seule personne au sein du ménage;
- un ajustement pour tenir compte du sous-échantillonnage des non-bénévoles;
- un ajustement pour tenir compte des valeurs aberrantes;
- un ajustement pour assurer la concordance entre les estimations de la population et les totaux connus selon le province, le groupe d'âge et le sexe tirés des chiffres de population projetés du recensement pour les personnes âgées de 15 ans et plus.

Ces étapes sont décrites ci-dessous.

#### 1. Calcul du poids de base du numéro de téléphone

Le poids initial, qui correspond à l'inverse de la probabilité de sélection du numéro de téléphone, est calculé dans chaque strate comme suit :

$$w_i = \left( \frac{\text{nombre total de numéros de téléphone possibles dans les banques valides}}{\text{nombre de numéros de téléphone échantillonnés}} \right)$$

120 650 numéros de téléphone ont été sélectionnés dans l'échantillon.

#### 2. Ajustement pour tenir compte des numéros de téléphone non résolus

Avant la collecte des données, les 120 650 numéros de téléphone ont été soumis à un processus de filtrage qui a permis d'éliminer 9 154 numéros d'affaires et 20 775 numéros hors service, de sorte qu'il restait 90 721 numéros de téléphone pour la collecte des données. On a attribué à chaque enregistrement retenu un statut initial « résidentiel » ou « indéterminé ».

Au terme de la période de collecte des données, on a utilisé l'information portant sur l'historique des appels pour déterminer le statut final des enregistrements. Chaque enregistrement a été classé comme étant hors du champ d'enquête, dans le champ d'enquête ou non résolu. Les poids des enregistrements résolus et hors du champ d'enquête ont été corrigés pour tenir compte des enregistrements non résolus, et ces derniers ont été éliminés. L'ajustement a été effectué au niveau de la strate et de façon distincte pour les enregistrements ayant un statut initial résidentiel et pour ceux ayant un statut initial indéterminé (voir la section 5.2 pour la description des strates).

Au total, 8 837 enregistrements non résolus ont été supprimés, ce qui laissait 81 884 enregistrements.

On a ajusté les poids comme suit, dans chaque strate et selon le statut initial :

$$w_2 = w_1 * \left( \frac{\sum w_1 \text{ pour les numéros de téléphone résolus} + \sum w_1 \text{ pour les numéros de téléphone non résolus}}{\sum w_1 \text{ pour les numéros de téléphone résolus}} \right)$$

### 3. Élimination des numéros de téléphone hors du champ d'enquête

On a éliminé, au terme de la collecte, les numéros de téléphone hors service ou jugés hors du champ d'enquête (numéros d'affaires, de téléphones cellulaires, de locaux non résidentiels, de logements collectifs, etc.). Il restait 51 277 enregistrements au terme de cette étape.

### 4. Ajustement pour tenir compte des données manquantes sur le nombre de lignes téléphoniques (premier ajustement pour tenir compte de la non-réponse au niveau du ménage)

Pour convertir le poids du numéro de téléphone calculé à l'étape 2 en poids du ménage, on doit diviser le poids du numéro de téléphone par le nombre de lignes téléphoniques associées au ménage. Dans certains cas, on ne peut déterminer le nombre de lignes soit en raison d'une non-réponse ponctuelle, soit en raison d'une non-réponse totale pour le ménage. Dans le cas d'une non-réponse ponctuelle, la valeur un a été imputée au nombre de lignes. Dans les autres cas où il n'a pas été possible de déterminer le nombre de lignes, les enregistrements ont été supprimés, et les poids des unités retenues ont été gonflés pour tenir compte des enregistrements éliminés.

L'examen de la non-réponse a révélé que les enregistrements des ménages ayant complétés l'entrevue mais comportant au moins un code de refus ou de barrière linguistique dans l'historique des appels se caractérisent par des taux de bénévoles nettement inférieurs à ceux des autres enregistrements. On a donc créé des groupes d'ajustement en subdivisant chaque strate en groupes, en fonction de la présence ou non d'un code de refus et/ou de barrière linguistique.

On a ajusté les poids comme suit, dans chaque strate et selon les groupes d'ajustement (refus/barrière linguistique) :

$$w_3 = w_2 * \left( \frac{\sum w_2 \text{ pour les ménages avec nombre de lignes déterminé} + \sum w_2 \text{ pour les ménages avec nombre de lignes indéterminé}}{\sum w_2 \text{ pour les ménages avec nombre de lignes déterminé}} \right)$$

Il restait 33 714 enregistrements au terme de cette étape.

### 5. Ajustement pour tenir compte du nombre de lignes dans le ménage

On a ajusté à la baisse les poids des ménages ayant plus d'une ligne téléphonique (numéros de téléphone distincts) pour tenir compte de la probabilité de sélection plus élevée de ces ménages. Le poids du numéro de téléphone a été divisé par le nombre de lignes dans le ménage. La valeur maximale du nombre de lignes téléphoniques a été fixée à quatre pour éviter les valeurs aberrantes. À cette étape, le poids du numéro de téléphone correspond au poids du ménage.

L'ajustement des poids a été le suivant :

$$w_4 = \left( \frac{w_3}{\text{nombre de lignes téléphoniques admissibles dans le ménage}} \right)$$

## 6. Ajustement pour tenir compte de la non-réponse au niveau du ménage (deuxième ajustement pour tenir compte de la non-réponse au niveau du ménage)

Cette étape vise les autres ménages non répondants, c'est-à-dire ceux pour lesquels il a été possible de déterminer le nombre de lignes téléphoniques. Les poids ont été ajustés à la hausse, dans chaque strate, pour tenir compte des ménages non répondants. On a supprimé ces derniers, de sorte qu'il restait alors 32 464 enregistrements au terme de cette étape.

On a ajusté les poids comme suit, dans chaque strate :

$$w_5 = w_4 * \left( \frac{\sum w_4 \text{ pour les ménages répondants} + \sum w_4 \text{ pour les ménages non répondants}}{\sum w_4 \text{ pour les ménages répondants}} \right)$$

## 7. Ajustement pour tenir compte de la sélection d'une seule personne dans le ménage (âgée de 15 ans ou plus)

Le poids des ménages calculé à l'étape 6 a été multiplié par le nombre de personnes âgées de 15 ans et plus au sein du ménage. La valeur maximale du nombre de membres a été fixée à cinq pour éviter les valeurs aberrantes. À cette étape, les poids ne représentent plus les ménages, mais représentent plutôt les personnes.

L'ajustement des poids a été le suivant :

$$w_6 = w_5 * (\text{nombre de membres âgés de 15 ans et plus au sein du ménage})$$

## 8. Correction pour tenir compte de la non-réponse au niveau de la personne

Les poids ont été ajustés à la hausse pour tenir compte des personnes non répondantes. Cet ajustement a été effectué dans chaque strate, selon le groupe d'âge et le sexe, et les personnes non répondantes ont été éliminées. Il restait 29 031 enregistrements au terme de cette étape.

On a ajusté les poids comme suit, dans chaque strate et selon le groupe d'âge et le sexe :

$$w_7 = w_6 * \left( \frac{\sum w_6 \text{ pour les personnes répondantes} + \sum w_6 \text{ pour les personnes non répondantes}}{\sum w_6 \text{ pour les personnes répondantes}} \right)$$

## 9. Ajustement pour tenir compte du sous-échantillonnage des non-bénévoles

Le taux pondéré de sous-échantillonnage des non-bénévoles a été calculé comme suit, dans chaque strate, à partir des données pondérées de l'étape précédente :

$$\text{taux pondéré de sous-échantillonnage} = \left( \frac{\sum w_7 \text{ pour les non-bénévoles sélectionnés}}{\sum w_7 \text{ pour les non-bénévoles sélectionnés} + \sum w_7 \text{ pour les non-bénévoles non sélectionnés}} \right)$$

L'inverse du taux obtenu a été multiplié par les poids des non-bénévoles sélectionnés, et les non-bénévoles non sélectionnés ont été éliminés. En fait, les poids des non-bénévoles sélectionnés ont à peu près doublé pour tenir compte des non-bénévoles qui n'ont pas été sélectionnés. Le taux théorique du sous-échantillonnage des non-bénévoles est de 50 %, mais le taux réel pondéré du sous-échantillonnage dans chaque strate varie de 46,5 % à 59 %.

Dans le cas des non-bénévoles, les poids ont été ajustés comme suit, dans chaque strate :

$$w_8 = \left( \frac{w_7}{\text{taux pondéré de sous-échantillonnage}} \right)$$

Dans le cas des bénévoles,  $w_8 = w_7$

Le nombre final d'enregistrements s'est établi à 20 832.

### 10. Calage en fonction de totaux connus

On a ajusté les poids de manière à assurer la concordance entre les estimations de la population et les chiffres de population pour les personnes âgées de 15 ans et plus tirés de sources externes. On a utilisé les totaux de contrôle suivants issus de sources externes :

- Population totale de chaque strate définie selon la province/région métropolitaine de recensement (RMR), et
- Population totale selon la province, le sexe et les groupes d'âge suivants : 15 à 19 ans, 20 à 24 ans, 25 à 29 ans, 30 à 34 ans, 35 à 39 ans, 40 à 44 ans, 45 à 49 ans, 50 à 54 ans, 55 à 59 ans, 60 à 64 ans, 65 à 69 ans, 70 ans et plus.

Cette opération de calage a simplement servi de procédure d'ajustement temporaire avant le repérage des valeurs aberrantes. Une fois les valeurs aberrantes repérées, on a mis de côté ce calage.

### 11. Détection et traitement des valeurs aberrantes

Le traitement des valeurs aberrantes est un processus qui permet de réduire l'impact des valeurs extrêmes pondérées. Les valeurs aberrantes de deux variables ont été détectées : le nombre total d'heures de bénévolat (VD1DHRS) et la valeur totale des dons (GS1DATOT). Une fois les valeurs aberrantes repérées, on a limité leur impact sur les estimations totales en réduisant le poids ( $w_8$ ) de l'étape 9 au moyen de la technique de Winsor. On a réduit le poids de l'observation aberrante de manière à ce que la valeur pondérée ajustée de l'observation aberrante corresponde à la valeur pondérée de la valeur non aberrante la plus élevée.

Le poids résultant de cette opération est  $w_9$ .

### 12. Calage en fonction des chiffres de population connus

À cette étape, on a effectué le calage de la même manière qu'à l'étape 10, la seule différence étant les poids utilisés dans la procédure. Dans cette opération de calage, on a employé les poids  $w_9$  issus de l'étape 11 après l'ajustement pour tenir compte des valeurs aberrantes. Une fois le calage effectué, on a répété l'opération de détection des valeurs aberrantes pour s'assurer qu'il n'en restait plus.

Le poids,  $w_{10}$ , obtenu à cette étape est le poids final, WTPM, figurant dans le fichier maître des microdonnées et également le poids WTPP dans le fichier de microdonnées à grande diffusion.

## 12.2 Pondération de la composante territoriale

Les étapes du calcul des poids pour la composante territoriale sont décrites ci-dessous.

### 1. Calcul des poids initiaux

Comme l'échantillon de la composante territoriale de l'ECDBP a été sélectionné de l'échantillon de l'Enquête sur la population active (EPA), le poids initial,  $w_1$ , a été calculé conformément au cadre conceptuel de l'EPA. Le poids initial correspond à l'inverse de la probabilité initiale de sélection.

### 2. Ajustement des poids initiaux pour tenir compte de la non-réponse

L'échantillon de l'ECDBP peut être réparti en quatre groupes :

- 1) les répondants;
- 2) les unités jugées hors du champ d'enquête;
- 3) les non-répondants considérés comme faisant partie du champ d'enquête;
- 4) les non-répondants dont le statut (hors du champs d'enquête ou dans le champ d'enquête) reste indéterminé.

On a attribué à chacune des 1 831 unités de l'échantillon de la composante territoriale un statut final selon ces quatre groupes, en fonction du code de résultat de l'application de collecte des données. Dans la mesure où les poids des 1 332 répondants doivent représenter toute la population dans le champ d'enquête, il a fallu gonfler les poids des répondants dans le champ d'enquête pour tenir compte des non-répondants. On doit aussi ajuster les poids pour tenir compte du fait que le quatrième groupe comprend à la fois des unités dans le champ d'enquête et hors du champ d'enquête. Partant de l'hypothèse que la proportion des unités hors du champ d'enquête parmi les unités dont le statut reste indéterminé correspond à celle des unités hors du champ d'enquête parmi les unités dont le statut est déterminé, on peut ajuster le poids des répondants pour tenir compte de la non-réponse à l'aide de la formule suivante :

$$w_{2gi} = w_{1gi} * \frac{(\sum_{Rép} w_{1gi} + \sum_{NR,dét} w_{1gi})}{\sum_{Rép} w_{1gi}} * \frac{(\sum_{Rép} w_{1gi} + \sum_{NR,dét} w_{1gi} + \sum_{HCE} w_{1gi} + \sum_{NR,indét} w_{1gi})}{(\sum_{Rép} w_{1gi} + \sum_{NR,dét} w_{1gi} + \sum_{HCE} w_{1gi})}$$

où :

$g$  représente le niveau auquel l'ajustement est effectué

$w_{1gi}$  = le poids initial de l'unité  $i$  dans le groupe d'ajustement  $g$

$\sum_{Rép} w_{1gi}$  = la somme des poids initiaux de tous les répondants du groupe d'ajustement  $g$

$\sum_{NR,dét} w_{1gi}$  = la somme des poids initiaux de tous les non-répondants au statut déterminé du groupe d'ajustement  $g$

$\sum_{NR,indét} w_{1gi}$  = la somme des poids initiaux de tous les non-répondants au statut indéterminé du groupe d'ajustement  $g$

$\sum_{HCE} w_{1gi}$  = la somme des poids initiaux de toutes les unités hors du champ d'enquête du groupe d'ajustement  $g$

On a effectué cet ajustement pour chacune des strates dans tous les cas où le nombre total de répondants et non-répondants était supérieur à 30 et le facteur d'ajustement, inférieur à deux. Lorsque ces conditions n'étaient pas remplies, les strates ont été fusionnées pour les fins de la procédure d'ajustement. Dans quatre cas, la taille de l'échantillon a nécessité une fusion des strates.

### **3. Ajustement pour tenir compte de la sélection d'une seule personne au sein du ménage**

Le poids calculé à l'étape 2 a été multiplié par le nombre de personnes âgées de 15 ans et plus au sein du ménage. Pour éviter les problèmes relatifs aux valeurs aberrantes et pour assurer la concordance avec la procédure de pondération utilisée pour la composante provinciale, on a fixé à cinq le nombre maximal de membres de 15 ans et plus dans le ménage. On a recensé 11 cas où le ménage comptait plus de cinq personnes âgées de 15 ans et plus.

$$w_3 = w_2 * (\text{nombre de membres du ménage âgés de 15 ans et plus})$$

### **4. Calage en fonction des chiffres de population connus**

L'opération de calage permet d'assurer une correspondance entre la somme des poids des répondants et les chiffres de population connus. Le calage a été effectué en fonction des totaux de contrôle des groupes d'âge-sexe selon le territoire, les trois groupes d'âge utilisés étant les suivants : 15 à 24 ans, 25 à 54 ans, 55 ans et plus. En outre, au Nunavut, le calage s'est aussi appuyé sur un total de contrôle pour la population inuite âgée de 15 ans et plus. Les totaux de contrôle retenus sont ceux du mois de référence d'octobre 2004.

Cette opération de calage a simplement servi de procédure de correction temporaire avant le repérage des valeurs aberrantes. Une fois les valeurs aberrantes repérées, on a mis de côté ce calage.

### **5. Repérage et traitement des valeurs aberrantes**

Le traitement des valeurs aberrantes est un processus qui permet de réduire l'impact des valeurs extrêmes pondérées. Les valeurs aberrantes de deux variables ont été détectées : le nombre total d'heures de bénévolat (VD1DHRS) et la valeur totale des dons (GS1DATOT). Une fois les valeurs aberrantes repérées, on a limité leur impact sur les estimations totales en réduisant le poids ( $w_3$ ) de l'étape 3 au moyen de la technique de Winsor. On a réduit le poids de l'observation aberrante de manière à ce que la valeur pondérée corrigée de l'observation aberrante corresponde à la valeur pondérée de la valeur non aberrante la plus élevée.

Le poids résultant de cette opération est  $w_4$ .

### **6. Calage en fonction des chiffres de population connus**

À cette étape, on a effectué le calage de la même manière qu'à l'étape 4, la seule différence étant les poids utilisés dans la procédure. Dans cette opération de calage, on a employé les poids ( $w_4$ ) issus de l'étape 5 après la correction pour tenir compte des valeurs aberrantes. Une fois le calage effectué, on a répété l'opération de détection des valeurs aberrantes pour s'assurer qu'il n'en restait plus.

Le poids,  $w_5$ , obtenu à cette étape est le poids final, WTPM, figurant dans le fichier maître des microdonnées et également le poids WTPP dans le fichier de microdonnées à grande diffusion.



## **13.0 Questionnaires**

Consulter le fichier ECDBP2004\_QuestF.pdf renferme le questionnaire français pour l'Enquête canadienne sur le don, le bénévolat et la participation (ECDBP) de 2004.



## **14.0 Structure des fichiers**

Les données de l'Enquête canadienne sur le don, le bénévolat et la participation (ECDBP) de 2004 sont réparties entre deux fichiers, soit le fichier principal des réponses (MAIN.TXT), et le fichier sur les donateurs (GS.TXT). On se sert de la variable MASTERID pour coupler les deux fichiers maîtres et de la variable PUMFID pour coupler les deux fichiers de microdonnées à grande diffusion.

### **MAIN.TXT**

Il s'agit du fichier principal qui contient un enregistrement par répondant. Il renferme les réponses à toutes les questions, sauf celles qui figurent dans le fichier GS. En outre, des variables dérivées sommaires ont été créées à partir du fichier GS, et elles ont été incorporées à ce fichier.

### **GS.TXT**

Il s'agit du fichier des réponses sur les dons de charité. Il contient un ou plusieurs enregistrements pour chaque personne qui a fait un don en argent : un enregistrement pour chacun des organismes de bienfaisance, jusqu'à concurrence de 10, auquel le répondant a fait un don au cours de la période de référence de 12 mois, en réponse à une méthode de sollicitation donnée. Pour chacune des 13 méthodes de sollicitation énumérées dans le questionnaire, un donateur peut donc avoir jusqu'à 10 enregistrements contenant chacun des données sur le type d'organisme, de même que sur la valeur totale de tous les dons versés à cette organisme en réponse à cette méthode de sollicitation. Dans les cas où le répondant a fait un don à plus de 10 organismes en réponse à une méthode de sollicitation donnée, la valeur totale de tous les dons faits aux autres organismes figure dans le dixième enregistrement comme variable dérivée GS1D08.



## 15.0 Désignations des variables

L'Enquête canadienne sur le don, le bénévolat et la participation (ECDBP) de 2004 a adopté une désignation standard de huit caractères pour les noms des variables sur les fichiers de microdonnées.

### Structure de la composante des noms de variables

- Les **deux premiers** caractères sont une combinaison de lettres qui identifie la section du questionnaire dans laquelle la variable a été recueillie ou de laquelle les données utilisées pour la dérivation de la variable ont été venues.

Positions 1 et 2	Nom de la section du questionnaire
FV	Bénévolat encadré
HV	Historique des activités de bénévolat
VS	Spécifiques de bénévolat
VD	Détails du bénévolat
MV	Principales activités de bénévolat
RV	Raisons de faire du bénévolat
GV	Bénévolat en général
ES	Appui de l'employeur
NV	Raisons pour ne pas faire (plus) de bénévolat
IV	Activités bénévoles non encadrées
FG	Dons en espèces à des organismes de charité

Positions 1 et 2	Nom de la section du questionnaire
GS	Détails sur les dons
DG	Décisions de donner
RG	Raisons pour faire des dons
NG	Raisons pour ne pas faire (plus) de dons
OG	Autres dons
HG	Santé en général
PA	Participation
ED	Éducation
LF	Population active
SD	Données socio-démographiques
IN	Revenu

- Le **troisième** caractère indique la « vague » ou itération d'une enquête longitudinale. Ça prend toujours la valeur « 1 » pour l'ECDBP de 2004.
- Le **quatrième** caractère du nom de la variable fait référence au type de variable.

Position 4	Type de variable	Description
-	Variable recueillie	Une variable qui apparaît directement sur le questionnaire
C	Variable codée	Une variable codée d'après l'une ou plusieurs des variables recueillies (p. ex. la Classification nationale des professions - statistiques)
D	Variable dérivée	Une variable calculée à partir de l'une ou de plusieurs des variables recueillies ou codées, normalement calculée pendant le traitement au bureau central (p. ex. nombre total d'heures de bénévolat)
F	Variable indicatrice	Une variable calculée à partir de l'une ou de plusieurs des variables recueillies (comme une variable dérivée) mais normalement calculée par l'application informatique pour être utilisée plus tard pendant l'entrevue (p. ex. indicateur de bénévole)
G	Variable groupée	Variations recueillies, codées, supprimées ou dérivées réunies en groupes (p. ex. groupes d'âge)
I	Indicateur d'imputation	Indique si une variable sur le fichier a été imputée (non compris sur le fichier de microdonnées à grande diffusion)

- Les **cinquième, sixième, septième et huitième caractères** identifient la variable ou numéro de la question du questionnaire. Normalement, les quatre dernières positions suivent la désignation sur le questionnaire. Des chiffres sont utilisés lorsque cela est possible (p.ex. Q01 devient 01). Les questions à plusieurs réponses utilisent des lettres pour chaque catégorie de réponse possible (p.ex. Q01 (cocher toutes les réponses qui s'appliquent) devient 01A, 01B, 01C, etc.).

**Exemples de nom de variable :**

<b>MV1_02A : Nombre d'heures consacrées à faire du porte-à-porte pour l'organisme de bénévolat principal</b>	
<b>MV</b>	Section du questionnaire « Principales activités de bénévolat »
<b>1</b>	ECDBP de 2004
<b>_</b>	Variable recueillie
<b>02</b>	Numéro de la question du questionnaire
<b>A</b>	Première catégorie pour une question de type « cocher toutes les réponses qui s'appliquent »

<b>FV1FVOL: Indicateur de bénévolat</b>	
<b>FV</b>	Section du questionnaire « Bénévolat encadré »
<b>1</b>	ECDBP de 2004
<b>F</b>	Indicateur
<b>VOL</b>	Nom de variable

Nota : Quelques variables importantes ne portent pas la désignation (p. ex. MASTERID, PUMFID, PROVCODE, WTPM et WTPP.)

## **16.0 Cliché d'enregistrement à valeurs univariées**

Consulter le fichier ECDBP2004\_Principal\_Maître\_LvCds.pdf pour le cliché d'enregistrement français à chiffres univariés pour le fichier principal maître.

Consulter le fichier ECDBP2004\_GS\_Maître\_LvCds.pdf pour le cliché d'enregistrement français à chiffres univariés pour le fichier GS maître.

Consulter le fichier ECDBP2004\_Principal\_FMGD\_LvCds.pdf pour le cliché d'enregistrement français à chiffres univariés pour le fichier principal de microdonnées à grande diffusion.

Consulter le fichier ECDBP2004\_GS\_FMGD\_LvCds.pdf pour le cliché d'enregistrement français à chiffres univariés pour le fichier GS de microdonnées à grande diffusion.