Nº 89-653-X au catalogue — Nº 002
ISBN 978-0-660-21292-0

Concepts et méthodes

Enquête auprès des peuples autochtones, 2012 : Guide des concepts et méthodes

Elisabeth Cloutier et Éric Langlet

Février 2014

Statistique Canada Statistics Canada

Canada

Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à www.statcan.gc.ca.

Vous pouvez également communiquer avec nous par :

Courriel à infostats@statcan.gc.ca

Téléphone entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros sans frais suivants :

- Service de renseignements statistiques 1-800-263-1136
- Service national d'appareils de télécommunications pour les malentendants 1-800-363-7629
- Télécopieur 1-877-287-4369

Programme des services de dépôt

- Service de renseignements 1-800-635-7943
- Télécopieur 1-800-565-7757

Comment accéder à ce produit

Le produit nº 89-653-X au catalogue est disponible gratuitement sous format électronique. Pour obtenir un exemplaire, il suffit de visiter notre site Web à www.statcan.gc.ca et de parcourir par « Ressource clé » > « Publications ».

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site www.statcan.gc.ca sous « À propos de nous » > « Notre organisme » > « Offrir des services aux Canadiens ».

Publication autorisée par le ministre responsable de Statistique Canada

© Ministre de l'Industrie, 2014

Tous droits réservés. L'utilisation de la présente publication est assujettie aux modalités de l'entente de licence ouverte de Statistique Canada (http://www.statcan.gc.ca/reference/copyright-droit-auteur-fra.htm).

This publication is also available in English.

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, ses entreprises, ses administrations et les autres établissements. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

Signes conventionnels

Les signes conventionnels suivants sont employés dans les publications de Statistique Canada :

.	indisponible pour toute période de référence
..	indisponible pour une période de référence précise
...	n'ayant pas lieu de figurer
0	zéro absolu ou valeur arrondie à zéro
0^s	valeur arrondie à 0 (zéro) là où il y a une distinction importante entre le zéro absolu et la valeur arrondie
^{p}	provisoire
^{r}	révisé
x	confidentiel en vertu des dispositions de la *Loi sur la statistique*
^{E}	à utiliser avec prudence
F	trop peu fiable pour être publié
*	valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,05$)

Table des matières

1. Introduction 7

1.1 Aperçu de l'enquête 7

1.2 Objectifs du *Guide des concepts et méthodes* 7

2. Contenu de l'enquête : concepts et questions 8

2.1 Processus d'élaboration du contenu 8

2.2 Groupes d'identité autochtone 9

Définition de l'identité autochtone 9

Questions relatives à l'identification autochtone 10

Variables de l'identité autochtone à la disposition des utilisateurs de données 11

2.3 Niveaux géographiques 11

2.4 Thèmes du contenu de l'enquête 13

Composante de l'éducation 13

Composante de l'emploi 14

Supplément sur la santé 14

Autres indicateurs de base 14

2.5 Modules du questionnaire 14

2.6 Contenu couplé aux données de l'Enquête nationale auprès des ménages 16

2.7 Contenu tiré d'autres enquêtes 17

Identification autochtone 17

Éducation 17

Emploi 18

Santé 18

3. Conception de l'enquête 18

3.1 Population cible et couverture 18

3.1.1 Identification de la population autochtone 19

3.1.2 Date de référence de l'enquête 21

3.1.3 Base de données de l'Enquête nationale auprès des ménages 21

3.2 Plan d'échantillonnage 22

3.2.1 Domaines d'estimation 22

3.2.2 Stratification 24

3.2.3 Plan d'échantillonnage et répartition de l'échantillon 26

3.2.4 Ajustements de la taille de l'échantillon 27

3.3 Taille de l'échantillon et taux de réponse 28

Table des matières (suite)

4. Collecte des données 30

4.1 Échéancier 30
4.2 Mode de collecte 30
4.3 Supervision et contrôle de la qualité 30
4.4 Interviews par personne interposée 30
4.5 Stratégie de communication 31
4.6 Problèmes particuliers sur le terrain 31
4.7 Taux de réponse finaux 31

5. Traitement des données 32

5.1 Saisie des données 32
5.2 Étapes de traitement des enquêtes sociales 32
5.3 Réception des données brutes et épuration des dossiers 32
5.3.1 Définition d'un répondant 33
5.4 Nouveau codage des variables et questions à réponses multiples 33
5.5 Enchaînements : cheminements de réponse, sauts valides et non-réponse à des questions 34
5.6 Codage 35
5.6.1 Questions « Autre – Précisez » 35
5.6.2 Questions ouvertes et classifications types 35
5.7 Contrôle et imputation 36
5.8 Variables dérivées et couplage de l'ENM 36
5.9 Création de fichiers de données finaux et d'un dictionnaire de données 37

6. Pondération 38

6.1 Poids initiaux 38
6.2 Ajustement pour les unités non envoyées à la collecte 38
6.3 Ajustement pour tenir compte de la non-réponse 39
6.4 Ajustement pour les répondants partiels 40
6.5 Poststratification 40
6.6 Ajustement des poids extrêmes – méthode de l'écart-sigma 41

7. Qualité des données 42

7.1 Aperçu de l'évaluation de la qualité des données 42
7.2 Erreurs d'échantillonnage et méthode bootstrap 42

Table des matières (suite)

7.3 Erreurs non dues à l'échantillonnage 45
7.3.1 Erreurs de non-réponse 45
7.3.2 Erreurs de couverture 46
7.3.3 Erreurs de mesure 46
7.3.4 Erreurs de traitement 47

8. Différences entre l'Enquête auprès des peuples autochtones et d'autres sources de données 47

8.1 Différences entre l'Enquête auprès des peuples autochtones et l'Enquête nationale auprès des ménages 48
8.1.1 Méthodes d'interview différentes et effet de la déclaration par personne interposée 48
8.1.2 Questionnaires différents 48
8.1.3 Contextes différents 49
8.1.4 Effet du temps 49
8.1.5 Traitement 49
8.1.6 Différences dans les estimations 49
8.2 Différences entre l'EAPA de 2006 et de celle de 2012 51
8.2.1 Différences sur le plan méthodologique 51
8.2.2 Différences en ce qui a trait au contenu 52
8.2.3 Modifications des questions sur l'identité autochtone 52

9. Diffusion des données 54

9.1 Aperçu de la diffusion des données de l'Enquête auprès des peuples autochtones de 2012 54
9.2 Produits de données et services 54
9.3 Produits analytiques 55
9.4 Produits de référence 56
9.5 Contrôle de la divulgation 56

Annexe A Indicateurs de l'Enquête auprès des peuples autochtones de 2012 selon le module du questionnaire 57

Introduction 57
Identification autochtone 57
Composition du ménage et état matrimonial 57
Mobilité 57
Éducation 58
Langues autochtones 60
Pensionnat 60
Emploi (répondants âgés de 15 ans et plus) 60
Activités traditionnelles 61

Table des matières (fin)

Revenu 62
Santé 62
Logement 64
Géographie 64

Annexe B Catégories de réponses supplémentaires créées pour les questions « Autre — Précisez » 65

Annexe C Classifications types utilisées pour l'Enquête auprès des peuples autochtones de 2012 . 66

Principal domaine d'études : Classification des programmes d'enseignement 66
Industrie et profession : Système de classification des industries de l'Amérique du Nord et Classification nationale des professions 66
Langues autochtones 67

Annexe D Glossaire des termes de l'enquête 68

Bibliographie 78

1. Introduction

1.1 Aperçu de l'enquête

L'Enquête auprès des peuples autochtones (EAPA) est une enquête nationale sur la situation socioéconomique des Premières Nations vivant hors réserve, des Métis et des Inuit âgés de 6 ans et plus. L'EAPA vise à déterminer les besoins de ces groupes autochtones et à guider les politiques et les programmes destinés à améliorer leur bien-être. Elle fournit des données pertinentes et à jour à divers intervenants, y compris les organismes des Premières Nations, métis et inuits, les communautés, les fournisseurs de services, les chercheurs, les gouvernements et le grand public.

L'EAPA est réalisée par Statistique Canada depuis 1991, et elle fournit une gamme d'indicateurs sociaux et économiques relatifs aux peuples autochtones. Il s'agit d'une enquête postcensitaire conçue pour compléter les données du recensement et celles de l'Enquête nationale auprès des ménages (ENM). L'EAPA de 2012 représente le quatrième cycle de l'enquête et le premier à adopter une approche thématique ciblée. L'enquête de 2012 porte sur l'éducation, l'emploi et la santé. Celle-ci continuera de fournir des indicateurs de base, notamment sur la langue, le revenu, le logement et la mobilité. Trois ministères fédéraux ont contribué au financement de l'enquête : Affaires autochtones et Développement du Nord Canada, Santé Canada et Emploi et Développement social Canada (anciennement connu sous le nom de Ressources humaines et Développement des compétences Canada).

Le présent cycle de l'EAPA a été mené du 6 février au 30 juillet 2012. Plus de 50 000 personnes ont été sélectionnées en vue d'y participer, et le taux de réponse final a été de 76 %. Le plan d'enquête a permis de produire des données fiables pour chaque province et territoire (provinces de l'Atlantique regroupées) ainsi que pour chacune des quatre régions inuites : le Nunatsiavut (côte nord du Labrador), le Nunavik (Nord du Québec), le territoire du Nunavut et la région inuvialuite des Territoires du Nord-Ouest. L'enquête visait aussi quatre groupes de scolarité : les élèves actuels de la 1re à la 6e année; les élèves actuels de la 7e à la 12e année; les finissants de l'école secondaire (diplômés); ainsi que les décrocheurs de l'école secondaire. Des analyses distinctes de ces aspects sont possibles pour chaque groupe autochtone, soit les Premières Nations vivant hors réserve, les Métis et les Inuit âgés de 6 ans et plus.

Jusqu'à présent, les données de l'EAPA ont largement été utilisées et ont permis de combler les lacunes statistiques pour bon nombre de questions. Ces données ont servi à éclairer la prise de décisions relatives à la planification et à l'élaboration de programmes et de politiques, à améliorer les services offerts aux peuples autochtones et à appuyer les projets de recherche universitaire. La Commission royale sur les peuples autochtones a utilisé les données de l'EAPA de 1991 comme principale source de renseignements démographiques, sociaux et économiques dans son rapport final et ses études de recherche. Le rapport final de la Commission recommandait que l'EAPA soit menée de façon régulière en vue de faire le suivi des conditions sociodémographiques des peuples autochtones.

L'EAPA de 2012 continuera d'orienter l'élaboration de politiques et de programmes visant à améliorer le bien-être des peuples autochtones. Les nouvelles données thématiques de l'enquête permettront d'avoir une connaissance plus approfondie des possibilités et des défis menant à la réussite au chapitre de l'éducation et de l'emploi. Des données sur la santé, la langue, le revenu, le logement et la mobilité seront aussi disponibles aux fins de l'analyse. L'EAPA continuera d'être une source importante de renseignements pour les organismes des Premières Nations, métis et inuits, les planificateurs communautaires, les fournisseurs de services, les gouvernements et les chercheurs.

1.2 Objectifs du *Guide des concepts et méthodes*

Ce *Guide des concepts et méthodes* présente un examen détaillé des approches spécialisées et méthodologiques de l'EAPA de 2012. Conçu pour les utilisateurs de données de l'EAPA, ce guide décrit les concepts et les mesures de l'enquête, tout en abordant les détails techniques de la conception, les travaux sur le terrain et le traitement des données. Ce document a aussi pour objectif de fournir des renseignements sur la façon d'utiliser et d'interpréter les résultats de l'enquête. De plus, une analyse de la qualité des données permet aux utilisateurs d'examiner les forces et les limites des données en fonction de leurs besoins particuliers.

La section 1 du présent guide donne un aperçu de l'EAPA de 2012. Le contexte et les objectifs de l'enquête y sont également abordés. Le chapitre 2 traite des thèmes de l'EAPA et explique les concepts et les définitions clés utilisés dans le cadre de l'enquête. Les chapitres 3 à 6 portent sur des aspects importants de la méthodologie d'enquête de l'EAPA, allant de la conception de l'échantillon à la collecte et au traitement des données. Dans les chapitres 7 et 8, il est question des problèmes liés à la qualité des données, et l'on met en garde les utilisateurs en ce qui a trait à la comparaison des données de l'EAPA de 2012 et celles d'autres sources. Le chapitre 9 présente les produits de l'enquête qui sont offerts au public, y compris les tableaux de données, les articles analytiques et les documents de référence. Les annexes fournissent une liste exhaustive des indicateurs de l'enquête, des catégories supplémentaires de codage et des classifications types utilisées dans le cadre de l'EAPA. On y trouve également un glossaire des termes relatifs à l'enquête.

2. Contenu de l'enquête : concepts et questions

2.1 Processus d'élaboration du contenu

Le contenu de l'Enquête auprès des peuples autochtones (EAPA) de 2012 a été élaboré par Statistique Canada en collaboration avec trois ministères fédéraux bailleurs de fonds : Affaires autochtones et Développement du Nord Canada, Santé Canada et Emploi et Développement social Canada (anciennement connu sous le nom de Ressources humaines et Développement des compétences Canada). L'initiative de 2007 intitulée « Redéfinir le mode d'évaluation de la réussite de l'apprentissage chez les Autochtones », qui a été menée par le Conseil canadien sur l'apprentissage (CCA), de concert avec des communautés et des organismes des Premières Nations, inuits et métis provenant de partout au Canada, a servi de cadre important pour guider l'élaboration du contenu. L'enquête de 2012 a aussi tiré parti de nombreux indicateurs clés de cycles précédents de l'EAPA, qui ont été élaborés en collaboration avec des organismes autochtones nationaux.

Par suite de la diffusion des données de l'EAPA de 2006, on a procédé à un examen approfondi du contenu afin d'assurer la pertinence future des questions de l'EAPA pour les principaux intervenants et de déterminer les lacunes statistiques possibles. Cet examen a bénéficié de l'expertise d'un groupe diversifié de chercheurs et de spécialistes du domaine provenant de Statistique Canada et de l'extérieur de celui-ci. Il a inclus des analyses exhaustives des forces et des limites des indicateurs de 2006 et des domaines devant être améliorés. Les indicateurs ont été passés en revue en rapport avec les activités de collecte de données, le codage, les étapes de traitement des données et les résultats analytiques de 2006.

En ce qui a trait aux nouvelles exigences de contenu pour 2012, des mesures pertinentes uniformisées et bien établies qui sont utilisées dans d'autres enquêtes de Statistique Canada ont été recueillies et examinées en tant qu'indicateurs possibles. Ces indicateurs améliorent les possibilités de comparaison des réponses entre l'EAPA et d'autres enquêtes de Statistique Canada. En outre, dans le contexte du nouveau cadre du CCA pour l'apprentissage des Autochtones, on visait l'établissement d'un nouveau contenu pour l'EAPA afin de permettre la mesure d'une vaste gamme d'expériences au chapitre de l'éducation chez les Inuit, les Métis et les Premières Nations vivant hors réserve. Le nouveau cadre du CCA reconnaît que l'apprentissage chez les Autochtones est un processus holistique, qui se poursuit toute la vie, émane de nombreuses sources différentes et est ancré dans les langues et les cultures autochtones (CCA, 2009).

Ces activités d'élaboration de contenu ont mené à une série de recommandations qui ont été mises en œuvre pour l'EAPA de 2012. Le questionnaire de l'enquête de 2012 a par la suite fait l'objet de plusieurs essais qualitatifs auprès des Premières Nations, des Métis et des Inuit au Canada afin d'évaluer plus particulièrement la compréhension culturelle des questions. Les essais qualitatifs ont été menés par le Centre de ressources en conception de questionnaires (CRCQ) dans huit emplacements au pays, y compris le Nord. À partir de ces résultats, des ajustements ont été apportés au libellé des questions, et plus particulièrement en ce qui a trait au nouveau contenu et à l'enchaînement des questions.

Avant 2012, on utilisait un questionnaire papier. Pour la première fois, les questions de l'enquête de 2012 ont été conçues pour être posées au moyen d'une application d'interview assistée par ordinateur (IAO). L'IAO comporte de nombreuses fonctions qui ont servi à maximiser la qualité de la collecte des données. Elle a permis l'élaboration d'enchaînements de questions plus complexes, ainsi que des contrôles en ligne servant à déterminer les incohérences logiques pouvant être corrigées par les intervieweurs, en

collaboration avec les répondants, au moment de l'interview. Deux questionnaires d'IAO ont été élaborés pour cette enquête : l'interview téléphonique assistée par ordinateur (ITAO) et l'interview sur place assistée par ordinateur (IPAO). Ces nouveaux questionnaires d'ITAO et d'IPAO ont fait l'objet d'essais par module et d'essais de bout en bout exhaustifs.

Une dernière mesure a été prise en ce qui a trait à l'élaboration du contenu de l'EAPA de 2012. Comme l'échantillon de l'EAPA de 2012 est tiré de l'Enquête nationale auprès des ménages (ENM) de 2011 (voir la section 3.1), il a été décidé que les données pertinentes de l'ENM seraient combinées aux renseignements fournis dans le cadre de l'interview de l'EAPA de 2012, grâce à un processus de couplage des données. Cette approche devait permettre de réduire le nombre de questions de l'EAPA et de fournir aux utilisateurs un ensemble de données plus exhaustives. Ce processus est décrit à la section 2.6.

2.2 Groupes d'identité autochtone

Les Premières Nations vivant hors réserve, les Métis et les Inuit constituent la population cible de l'EAPA de 2012. Les données d'enquête serviront à appuyer des analyses distinctes pour chacun de ces groupes.

L'EAPA de 2012 ne comprenait pas la population vivant dans les réserves indiennes ou les établissements indiens, de même que la population vivant dans certaines communautés des Premières Nations au Yukon et dans les Territoires du Nord-Ouest au moment de la collecte des données de l'ENM. Pour obtenir des renseignements à jour sur les Premières Nations vivant dans les réserves, veuillez consulter l'Enquête nationale auprès des ménages de 2011 et l'Enquête régionale sur l'éducation, l'emploi et la petite enfance chez les Premières Nations (EREE), laquelle n'est pas réalisée par Statistique Canada. L'EREE sera menée par le Centre de gouvernance de l'information des Premières Nations à l'automne 2013. Pour en savoir davantage sur l'EREE, veuillez consulter le site web du Centre de gouvernance de l'information des Premières Nations.

Aux fins de l'EAPA de 2012, les Premières Nations vivant hors réserve, les Métis et les Inuit ont été définis en fonction du concept d'« identité autochtone ». Tant du point de vue conceptuel que méthodologique, l'identité autochtone comporte de nombreux aspects complexes. La présente section du guide porte sur la définition de l'identité autochtone et sur la façon dont elle a été mesurée dans le cadre de l'EAPA de 2012. Une description plus technique des méthodes d'échantillonnage de l'enquête ayant servi à déterminer la population d'identité autochtone pour l'EAPA de 2012 est présentée au chapitre 3. Celle-ci explique comment les répondants de l'EAPA ont été sélectionnés parmi ceux de l'ENM de 2011.

Définition de l'identité autochtone

La population d'identité autochtone comprend toute personne ayant déclaré appartenir à au moins l'un des groupes suivants lors de l'EAPA de 2012 :

- Autochtone, c'est-à-dire Première Nation (Indien de l'Amérique du Nord), Métis ou Inuk (Inuit);
- Indien avec statut, soit un Indien inscrit ou des traités en vertu de la *Loi sur les Indiens* du Canada;
- membre d'une Première Nation ou d'une bande indienne.

Le premier critère ci-dessus est désigné comme l'« autodéclaration autochtone ». Un répondant peut se considérer comme Autochtone ou se percevoir comme appartenant à l'un ou plusieurs des groupes autochtones mentionnés : Première Nation, Métis ou Inuit. Les Premières Nations comprennent les Indiens avec statut et les Indiens sans statut. Il convient aussi de souligner que certains répondants utilisent le terme *Premières Nations*, alors que d'autres utilisent *Indiens de l'Amérique du Nord*.

Comme le laisse entendre la définition de l'EAPA, une personne n'a pas à se déclarer comme Autochtone (Première Nation, Métis ou Inuit) pour être considérée comme faisant partie de la population d'identité autochtone. Si une personne a déclaré être un Indien avec statut ou être membre d'une Première Nation ou d'une bande indienne, elle est considérée comme appartenant à la population d'identité autochtone, et ce, peu importe ses réponses aux autres questions.

Pour ce qui est de la mesure relative aux « Indiens avec statut », l'EAPA de 2012 comprenait toutes les personnes ayant indiqué être un Indien inscrit ou un Indien des traités aux termes de la *Loi sur les Indiens*. Cette définition englobe les personnes qui étaient inscrites en vertu des projets de loi C-31 et C-3, des modifications législatives ayant été apportées à la *Loi sur les Indiens* en 1985 et 2011. La modification la plus récente, qui remonte à 2011, est le projet de loi C3 : *Loi sur l'équité entre les sexes relativement à l'inscription au registre des Indiens*.

Questions relatives à l'identification autochtone

Le tableau 1 ci-dessous présente les six questions relatives à l'identification autochtone qui ont été posées dans le cadre de l'EAPA de 2012 et montre la classification de l'identité autochtone établie à partir des réponses fournies à ces questions. Comme il est indiqué, pour qu'une personne fasse partie de la population d'identité autochtone, elle doit avoir répondu OUI à la question ID_Q01 (autodéclaration autochtone) ou OUI à au moins l'une des questions ID_Q03 (Indien avec statut), ID_Q04B (Indien inscrit en vertu de la Loi C-31 et de la Loi C-3), ou ID_Q05 (membre d'une Première Nation ou d'une bande indienne). Fait à noter, les personnes qui ont répondu NON à toutes ces questions ont été classées comme étant non-Autochtones et considérées comme ne faisant pas partie du champ de l'enquête. Les poids d'enquête ont été ajustés en conséquence.

Comme le montre le tableau 1, la mesure de l'autodéclaration autochtone a été répartie entre deux questions : ID_Q01 et ID_Q02. Cela a permis aux répondants de se déclarer comme Autochtones à la question ID_Q01, même s'ils ne se déclaraient pas comme appartenant à un groupe particulier désigné à la question ID_Q02 (Premières Nations, Métis ou Inuit). Dans le cadre de l'EAPA de 2012, les répondants qui ont indiqué OUI à ID_Q01, mais NON à ID_Q02, ont été classés dans un groupe particulier selon leurs réponses à l'ENM de 2011. En outre, les répondants qui ont indiqué être un Indien avec statut ou un membre d'une Première Nation ou d'une bande indienne, sans se déclarer comme autochtone, ont été classés comme ayant une identité des Premières Nations aux fins de l'EAPA de 2012.

Pour obtenir plus de renseignements sur la collecte et le traitement des données relatives à l'identification autochtone dans le cadre de l'EAPA de 2012, veuillez consulter la section 5.7 du présent guide. Par ailleurs, les différences entre les questions de l'EAPA de 2012 et celles de l'EAPA de 2006 concernant l'identité autochtone sont présentées à la sous-section 8.2.3.

Tableau 1 Enchaînement des questions et classification de l'identité autochtone

Enchaînement du questionnaire et classification de l'identité autochtone	Premières Nations	Métis	Inuit	Non-autochtones (hors du champ de l'enquête)
ID_Q01 Êtes-vous un(e) Autochtone, c'est-à-dire, Première Nation, Métis(se) ou Inuk (Inuit)? Première Nation comprend les Indiens(nes) avec statut et les Indiens(nes) sans statut				
Oui... voir les questions subséquentes				
ID_Q02 – Êtes-vous Première Nation, Métis(se) ou Inuk (Inuit)?				
Première Nation (Indien de l'Amérique du Nord)	X			
Métis		X		
Inuk			X	
Non – voir les questions subséquentes				
ID_Q03 – Êtes-vous un(e) Indien(ne) avec statut, soit un(e) Indien(ne) inscrit(e) ou des traités, aux termes de la *Loi sur les Indiens* du Canada?				
Oui	X			
Non - classification fondée sur les réponses globales				
ID_Q04A Avez-vous déjà présenté une demande au ministère des Affaires autochtones et Développement du Nord Canada (précédemment nommé AINC) afin d'être inscrit(e) comme Indien(ne) de plein droit en vertu de la Loi C-31 ou de la Loi C-3?				
Oui ...				
ID_Q04B - Êtes-vous inscrit(e) comme Indien(ne) de plein droit en vertu de la Loi C-31 ou de la Loi C-3?				
Oui, Loi C-31	X			
Oui, Loi C-3	X			
Non - classification fondée sur les réponses globales				
Non - classification fondée sur les réponses globales				
ID_Q05 - Êtes-vous membre d'une Première Nation ou d'une bande indienne?				
Oui	X			
Non - classification fondée sur les réponses globales				
Non à toutes les questions : ID_Q01-Q02, ID_Q03, ID_Q04A-Q04B et ID_Q05				X

Variables de l'identité autochtone à la disposition des utilisateurs de données

Les données diffusées à partir de l'EAPA de 2012 fourniront des variables analytiques pour chacun des aspects de l'identité autochtone. Ces variables seront essentielles pour les utilisateurs des données qui souhaitent effectuer des analyses au sujet de thèmes précis pour chacun des groupes suivants : Premières Nations vivant hors réserve, Métis et Inuit. Les variables de l'identité autochtone comprendront des indicateurs d'identité simple et multiple (p. ex. les personnes ayant déclaré appartenir à la fois aux Premières Nations et aux Métis). Les variables concernant les Indiens avec statut (Indiens inscrits ou des traités) seront aussi disponibles aux fins de l'analyse. Une variable sur l'ascendance autochtone sera fournie à partir de l'ENM, l'EAPA de 2012 ne comportant pas de mesure directe de cet aspect.

2.3 Niveaux géographiques

L'EAPA de 2012 assure la couverture de certaines régions géographiques de base, y compris les provinces et les territoires (les provinces de l'Atlantique regroupées) et les quatre régions inuites de l'Inuit Nunangat.

L'Inuit Nunangat est la patrie des Inuit du Canada. Elle comprend les communautés incluses dans les quatre régions inuites : Nunatsiavut (côte nord du Labrador), Nunavik (Nord du Québec), le territoire du Nunavut et la région inuvialuite des Territoires du Nord-Ouest. Ces régions englobent le territoire traditionnellement habité par les Inuit au Canada (voir la carte 1).

Comme il a été décrit à la section 3.2 portant sur le plan d'échantillonnage, ces régions géographiques ont été ciblées par l'EAPA de 2012 afin de faire en sorte que des estimations de données appropriées soient disponibles à ces niveaux géographiques. D'autres variables géographiques figurent dans la base de données de l'EAPA de 2012, à partir des niveaux géographiques de l'ENM. Il s'agit des régions métropolitaines de recensement (RMR), des subdivisions de recensement (SDR) et des centres de population. En outre, les niveaux géographiques comprennent les régions sociosanitaires du Canada, qui correspondent à des régions administratives qui présentent un intérêt pour les autorités en matière de santé.

Cependant, les utilisateurs doivent savoir que ce ne sont pas toutes les données de l'EAPA qui peuvent faire l'objet de totalisations croisées ou qui peuvent être analysées selon des niveaux géographiques détaillés. Certains tableaux de données seront disponibles pour des niveaux géographiques plus détaillés; toutefois, la fiabilité des estimations de données pour chaque niveau devra être examinée au cas par cas.

Les utilisateurs peuvent consulter le dictionnaire de données de l'EAPA de 2012 pour obtenir la liste complète des niveaux géographiques tirés de l'enquête (veuillez communiquer avec les Services à la clientèle de Statistique Canada, à sasd-dssea@statcan.gc.ca, ou au 1-800-263-1136). Le *Dictionnaire de l'Enquête nationale auprès des ménages* définit également les niveaux géographiques relatifs à l'EAPA.

Carte 1 Les quatre régions de l'Inuit Nunangat

Source : Inuit Tapiriit Kanatami.

2.4 Thèmes du contenu de l'enquête

En raison de l'approche thématique adoptée, l'EAPA de 2012 a permis de recueillir des données pour les quatre composantes figurant ci-dessous :

- Composante de l'éducation
- Composante de l'emploi
- Supplément sur la santé
- Autres indicateurs de base

La présente section décrit les quatre composantes. La section 2.5 donne un aperçu des modules du questionnaire de l'EAPA de 2012. L'annexe A comprend une liste détaillée des indicateurs mesurés dans chacun de ces modules. Pour obtenir plus de renseignements, veuillez consulter le questionnaire de l'EAPA de 2012.

Par ailleurs, une description exhaustive de toutes les variables disponibles à partir des données de l'enquête figure dans le dictionnaire de données de l'EAPA de 2012. Pour obtenir le dictionnaire de données de l'enquête, veuillez communiquer avec les Services à la clientèle de Statistique Canada, à sasd-dssea@statcan.gc.ca, ou au 1-800-263-1136.

Composante de l'éducation

Le contenu de la composante de l'éducation de l'EAPA de 2012 a été conceptualisé selon différents groupes ou cheminements en matière d'éducation. Les questions ont été formulées pour en apprendre davantage sur les expériences vécues à l'école par les personnes faisant partie des quatre groupes de scolarité suivants :

- les élèves actuels, de la 1re à la 6e année;
- les élèves actuels, de la 7e à la 12e année;
- les finissants de l'école secondaire, avec un diplôme d'études secondaires (ou l'équivalent); et
- les décrocheurs de l'école secondaire, sans diplôme d'études secondaires et ne fréquentant pas l'école de la 1re à la 12e année.

Les finissants de l'école secondaire sont définis comme des personnes qui respectaient les exigences minimales pour l'obtention d'un diplôme d'études secondaires au moment de la tenue de l'enquête. Ce groupe comprend les personnes qui ont terminé un programme d'équivalence d'études secondaires. Les décrocheurs sont définis comme ceux qui ne fréquentaient pas l'école primaire ou secondaire au moment de la tenue de l'enquête et qui ne respectaient pas les exigences minimales pour l'obtention d'un diplôme d'études secondaires. Les personnes inscrites à un programme d'équivalence d'études secondaires qui n'avaient pas terminé ce programme au moment de la tenue de l'enquête ont été regroupées dans le groupe des « décrocheurs ».

Dans le cadre de l'enquête, divers enchaînements de questions ont été utilisés pour les différents groupes de scolarité, fournissant ainsi des données sur certains thèmes pour chaque groupe. Les thèmes étaient répartis comme suit :

- Pour les élèves actuels : expériences vécues à l'école par les enfants et les jeunes, caractéristiques de l'école, participation des parents, aspirations en matière d'études.
- Pour les finissants et les décrocheurs de l'école secondaire : expériences vécues à l'école primaire et à l'école secondaire, particulièrement pendant la dernière année de fréquentation; cheminements éducatifs dans le système scolaire, plus haut niveau de scolarité, obstacles rencontrés et expériences des études postsecondaires.

Le contenu de chacun de ces groupes a donné lieu à un riche ensemble de données portant sur une vaste gamme de questions de recherche et de politique possibles ainsi que sur les besoins du programme d'enseignement. L'enquête fournit des données sur les besoins actuels des élèves autochtones et de leurs familles, et permet aussi de mieux comprendre les facteurs liés à la réussite des études secondaires. Les

données permettent une analyse approfondie des différences entre les finissants et les décrocheurs de l'école secondaire, du point de vue de la qualité des expériences des études et des difficultés auxquelles les élèves font face, y compris les types de soutien qu'ils reçoivent à l'école, la participation de la famille à l'école, les relations avec les pairs, la raison du changement d'école et celles du décrochage scolaire. Les données sur l'enseignement couvrent aussi d'autres indicateurs importants, comme le plus haut niveau de scolarité, les diplômes d'études postsecondaires et le principal domaine d'études.

Il convient aussi de souligner que même si l'EAPA cible les Premières Nations vivant hors réserve, certains de ces répondants ont, à un moment donné, fréquenté l'école dans une réserve. Dans le cadre de l'EAPA de 2012, on demandait aux répondants s'ils avaient déjà fréquenté une école primaire ou secondaire dans une communauté des Premières Nations (dans une réserve).

Composante de l'emploi

La composante de l'emploi de l'EAPA de 2012 est axée sur les personnes de 15 ans et plus. Elle explore les expériences d'emploi et de chômage, en englobant des indicateurs clés du marché du travail, comme les activités sur le marché du travail au cours d'une période de référence particulière, la profession et l'industrie, ainsi que les conditions de travail atypiques, comme le travail temporaire, le travail à temps partiel et les emplois multiples. Le contenu comprend aussi des données approfondies sur les difficultés possibles sur le marché du travail, y compris les difficultés à trouver du travail, les raisons de travailler à temps partiel, les méthodes utilisées pour chercher du travail et les obstacles à la mobilité de la main-d'œuvre. Des renseignements sur la participation à des activités traditionnelles sont aussi recueillis.

Ce contenu fournit des données précieuses pour une gamme de questions de recherche et de politique dans le domaine de l'emploi des Premières Nations vivant hors réserve, des Métis et des Inuit. Ces indicateurs viennent aussi enrichir les analyses des réussites et des défis des finissants et des décrocheurs de l'école secondaire.

Supplément sur la santé

Le supplément sur la santé de l'EAPA de 2012 porte sur des sujets importants, dont l'accès aux soins de santé, les mesures de la santé physique, la santé mentale, les blessures, l'usage du tabac, la consommation d'alcool et de drogues, l'insécurité alimentaire et le soutien de la communauté. Ces indicateurs permettent de faire des analyses sur des questions liées à l'accès aux soins de santé, à l'état de santé physique et mentale et aux risques pour la santé. En outre, ces indicateurs permettront d'enrichir l'analyse des réussites et des obstacles dans les domaines de l'éducation et de l'emploi.

Autres indicateurs de base

Une vaste gamme d'indicateurs de base a été ajoutée à l'EAPA de 2012, notamment les langues autochtones, le logement, la mobilité, la composition du ménage et la fréquentation d'un pensionnat. Ces variables permettront de faire le suivi continu de nombreux enjeux importants touchant les Premières Nations vivant hors réserve, les Métis et les Inuit.

2.5 Modules du questionnaire

La présente section dresse la liste des modules du questionnaire de l'EAPA de 2012. L'annexe A comprend une liste détaillée des indicateurs mesurés dans chacun de ces modules. L'annexe B comprend une liste de catégories additionnelles qui ont été créées pendant le codage des données. L'annexe C décrit les classifications types utilisées en vue de créer des indicateurs pour les questions d'enquête ouvertes.

- Introduction
- Identification autochtone
- Composition du ménage et état matrimonial
- Mobilité
- Éducation
  - Situation par rapport aux études
  - Éducation 1 : Élèves actuels, de la 1re à la 6e année
  - Éducation 2 : Élèves actuels, de la 7e à la 12e année
  - Éducation 3 : Finissants et décrocheurs de l'école secondaire âgés de moins de 45 ans
  - Éducation 4 : Finissants et décrocheurs de l'école secondaire âgés de 45 ans et plus
- Langues autochtones
- Pensionnat
- Emploi (répondants de 15 ans et plus)
  - Activité sur le marché du travail
  - Situation vis-à-vis de l'activité
  - Recherche d'emploi
  - Travailleurs découragés
  - Lien à l'emploi précédent
  - Emplois multiples
  - Catégorie de travailleurs
  - Entreprise constituée en société
  - Durée d'occupation de l'emploi
  - Industrie
  - Professions
  - Nombre habituel d'heures travaillées
  - Emploi à temps partiel
  - Emploi permanent
  - Mobilité de la main-d'œuvre
- Activités traditionnelles
- Revenu
- Santé
  - Santé générale 1 — État de santé
  - Grossesse et accouchement
  - Taille et poids
  - Santé générale 2 — Consultation de différents professionnels de la santé
  - Problèmes de santé chroniques
  - Blessures
  - Santé mentale
  - Détresse
  - Suicide
  - Usage du tabac
  - Consommation d'alcool
  - Consommation de drogues
  - Sécurité alimentaire
  - Soutien de la communauté
- Logement

2.6 Contenu couplé aux données de l'Enquête nationale auprès des ménages

Dans le passé, l'EAPA, une enquête postcensitaire, tirait son échantillon des répondants autochtones du recensement. En 1991, 2001 et 2006, les données de l'EAPA ont été couplées à celles de chacun des recensements correspondants. En 2011, le Recensement de la population a été modifié, et les données très détaillées recueillies précédemment à l'aide du questionnaire complet l'ont été au moyen de la nouvelle Enquête nationale auprès des ménages de 2011 (ENM), y compris les questions permettant d'identifier la population autochtone. L'échantillon de l'EAPA a été sélectionné à partir des réponses fournies à l'ENM de 2011. En outre, le fichier de microdonnées finales de l'EAPA a été couplé à la base de données de la diffusion de l'ENM de 2011 (voir le chapitre 3 pour en savoir plus sur la sélection de l'échantillon de l'EAPA).

Parmi les avantages du couplage d'enregistrements de l'EAPA et de l'ENM figurent la réduction du fardeau de réponse pour la population cible de l'EAPA, la dérivation de poids d'enquête permettant de fournir des estimations valides, et la création d'un fichier de microdonnées exhaustif, qui peut être utilisé par les analystes pour élargir leurs connaissances et éclairer l'élaboration de politiques et de programmes pour les populations autochtones au Canada. Ensemble, les données des deux sources permettront de brosser un portrait statistique détaillé des populations des Premières Nations vivant hors réserve, des Métis et des Inuit au Canada, données qu'il n'est pas possible d'obtenir d'aucune autre source.

Au moment de la collecte des données, tous les répondants de l'ENM ont été informés que les renseignements fournis pouvaient être utilisés à l'appui d'autres enquêtes de Statistique Canada. Le message figurant sur la page couverture du questionnaire de l'ENM de 2011 indiquait ceci : « Il se peut que Statistique Canada utilise vos renseignements pour appuyer ses autres enquêtes ou pour effectuer des analyses. » Par ailleurs, au début de l'interview de l'EAPA de 2012, les répondants ont été informés de l'objectif de l'enquête, de son caractère volontaire et de ce qui suit :

> Afin de réduire le nombre de questions, Statistique Canada prévoit combiner l'information recueillie lors de l'Enquête nationale auprès des ménages de 2011 à l'information que vous fournirez dans cette interview. Il se peut également que nous ajoutions de l'information provenant d'autres enquêtes ou sources de données administratives. Toutes les données demeureront confidentielles et ne serviront qu'à des fins statistiques.

Tous les produits renfermant des données couplées sont diffusés en conformité avec les politiques, lignes directrices et normes de Statistique Canada. Seules les estimations statistiques agrégées qui sont conformes aux dispositions en matière de confidentialité de la *Loi sur la statistique* peuvent être diffusées à l'extérieur de Statistique Canada. Toutes les demandes de données pour des totalisations de l'EAPA font l'objet d'une vérification de confidentialité, et les données agrégées sont arrondies avant d'être diffusées aux clients.

Plus de 100 variables de l'ENM ont été couplées au fichier final de l'EAPA de 2012. La liste ci-dessous indique le type de variables de l'ENM qui ont été annexées au fichier analytique de l'EAPA. Fait à noter, ces variables de l'ENM ont trait à la situation du répondant le jour de la tenue de l'ENM, soit le 10 mai 2011. Elles ont été obtenues à partir des réponses fournies à l'ENM de 2011 par les répondants de l'EAPA de 2012. Les utilisateurs doivent être conscients que, dans certains cas, le répondant peut avoir déménagé, avoir connu un changement dans la composition de son ménage, ou avoir connu un changement d'emploi entre la date de l'ENM et celle de l'EAPA. Autrement dit, certains des renseignements fournis par l'ENM peuvent ne pas rendre compte de la situation du répondant au moment de l'interview de l'EAPA.

Une liste complète des variables couplées de l'ENM et des notes sont fournies dans le dictionnaire de données de l'EAPA de 2012, qui accompagne le fichier analytique de l'enquête.

- **Variables relatives au ménage**
  - Niveau géographique, y compris les régions métropolitaines de recensement et les régions de résidence des Inuit
  - Logement, y compris le mode d'occupation du logement, le nombre de pièces du logement et les besoins en réparations
  - Famille, y compris la présence d'un conjoint dans le ménage, la présence et le nombre d'enfants dans la famille de recensement et des données sur les parents ou les tuteurs pour les jeunes répondants

- Famille, y compris la présence d'un conjoint dans le ménage, la présence et le nombre d'enfants dans la famille de recensement, le nombre de personnes dans la famille de recensement et la situation et la structure de la famille de recensement. En outre, une série de variables dérivées est incluse afin de fournir les renseignements de l'ENM concernant le ou les parents ou le ou les tuteurs des répondants de l'EAPA qui auraient moins de 18 ans. Ces variables comprennent plusieurs caractéristiques de l'ENM concernant la première personne la mieux renseignée au sujet du répondant (PMR1) et, au besoin, les mêmes caractéristiques pour la deuxième personne la mieux renseignée au sujet du répondant (PMR2, généralement le conjoint de la PMR1). Les caractéristiques de la PMR1 et de la PMR2 dans l'ENM comprenaient l'identité et l'ascendance autochtone, le statut d'Indien inscrit ou d'Indien des traités, l'âge et le sexe, l'état matrimonial et l'union libre, la langue maternelle autochtone et la langue autochtone parlée à la maison, la situation vis-à-vis de l'activité et le revenu total.

- **Variables relatives à la personne**
  - Origine ethnique et religion
  - Emploi, y compris la situation vis-à-vis de l'activité, le nombre de semaines travaillées en 2010, la catégorie de travailleurs, le travail à temps plein ou à temps partiel
  - Lieu de travail, y compris la situation du lieu de travail, le type de navettage et la distance
  - Mobilité, y compris où cette personne habitait il y a un an et cinq ans
  - Revenu, y compris le revenu de la famille, le revenu d'emploi et la situation de faible revenu
  - Langue, y compris la connaissance des langues officielles, la langue maternelle, la langue parlée à la maison et la langue de travail
  - Limitations d'activités, y compris les difficultés ou réductions d'activités à la maison, au travail et à l'école

2.7 Contenu tiré d'autres enquêtes

Le contenu du questionnaire de l'EAPA de 2012 provient de plusieurs sources. L'EAPA de 2006 a servi de point de départ et de source principale pour ce qui est des questions bien établies devant être utilisées auprès des Premières Nations vivant hors réserve, des Métis et des Inuit. Des questions ont aussi été tirées d'autres enquêtes, dans la plus large mesure possible, pour améliorer la rigueur des données et permettre la comparabilité entre les sources de données, le cas échéant. Parfois, ces questions ont été modifiées afin de mieux convenir à l'EAPA. Dans certains cas, les données ne sont pas directement comparables.

La présente section fait état des principales sources utilisées pour l'élaboration du questionnaire de l'EAPA de 2012, en rapport avec les principaux thèmes de l'enquête. Il convient de souligner que ces sources ne représentent pas une liste exhaustive de toutes les enquêtes de Statistique Canada qui comprennent des indicateurs que l'on trouve dans l'EAPA.

Identification autochtone

- Enquête nationale auprès des ménages (ENM) de 2011
- Enquête auprès des peuples autochtones (EAPA) de 2006

Éducation

- EAPA de 2006, 2001 et 1991
- Enquête sur les enfants autochtones (EEA) de 2006
- Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ)
- Enquête auprès des jeunes en transition (EJET)
- Programme international pour le suivi des acquis des élèves (PISA)

- Programme d'enquêtes nationales sur le niveau de scolarité des ménages
- Enquête sur l'accès et le soutien à l'éducation et à la formation (EASEF)
- Enquête auprès des sortants (EAS)

Emploi

- EAPA de 2006
- Enquête sur la population active (EPA)
- Enquête sur les travailleurs âgés (ETA)

Santé

- EAPA de 2006
- Enquête régionale sur la santé des Premières Nations (ERS)
- Enquête sur la santé dans les collectivités canadiennes (ESCC)
- Enquête sur les jeunes Canadiens (EJC)
- Enquête canadienne sur les mesures de la santé (ECMS)

Fait à noter, quatre enquêtes sociales principales de Statistique Canada comportent des données selon le groupe autochtone : l'ENM, l'EPA, l'Enquête sociale générale (ESG) et l'ESCC.

3. Conception de l'enquête

3.1 Population cible et couverture

La population cible de l'EAPA de 2012 comprenait la population d'identité autochtone du Canada âgée de 6 ans et plus au 1er février 2012, vivant dans les logements privés à l'extérieur des réserves indiennes ou des établissements indiens, de même qu'à l'extérieur de certaines communautés des Premières Nations au Yukon et dans les Territoires du Nord-Ouest. Ces exclusions visaient à éviter le chevauchement avec l'Enquête régionale sur l'éducation, l'emploi et la petite enfance chez les Premières Nations, enquête qui sera menée à l'automne 2013 par le Centre de gouvernance de l'information des Premières Nations (voir la section 2.2). La liste des subdivisions de recensement (SDR) au Yukon et dans les Territoires du Nord-Ouest est présentée au tableau 2.

Tableau 2 Liste des subdivisions de recensement au Yukon et dans les Territoires du Nord-Ouest

Territoire	Nº de SDR	Nom de SDR	Genre de SDR
Yukon	6001003	Watson Lake	T
	6001004	Faro	T
	6001009	Whitehorse	CY
	6001029	Dawson	T
	6001044	Mt. Lorne	HAM
	6001045	Yukon, Unorganized	NO
	6001046	Swift River	SÉ
	6001049	Destruction Bay	SÉ
	6001050	Stewart Crossing	SÉ
	6001052	Keno Hill	SÉ
	6001055	Ibex Valley	HAM
	6001058	Marsh Lake	NO
	6001059	Macpherson-Grizzly Valley	NO
	6001060	Whitehorse, Unorganized	NO
Territoires du Nord-Ouest	6101014	Paulatuk	HAM
	6101017	Inuvik	T
	6101025	Aklavik	HAM
	6101036	Tuktoyaktuk	HAM
	6101041	Sachs Harbour	HAM
	6101063	Region 1, Unorganized	NO
	6101095	Ulukhaktok	HAM
	6102007	Norman Wells	T
	6102063	Region 2, Unorganized	NO
	6103097	Region 3, Unorganized	NO
	6104097	Region 4, Unorganized	NO
	6105003	Enterprise	SET
	6105016	Hay River	T
	6105026	Reliance	SET
	6105097	Region 5, Unorganized	NO
	6106023	Yellowknife	CY
	6106097	Region 6, Unorganized	NO

Les genres de SDR figurent dans le tableau 5 du *Dictionnaire du recensement*.
La « population d'identité autochtone » est définie à la sous-section 3.1.1 ci-dessous.

3.1.1 Identification de la population autochtone

L'échantillon de l'EAPA est tiré à partir des répondants ayant déclaré soit une identité autochtone ou une ascendance autochtone dans le questionnaire de l'Enquête nationale auprès des ménages (ENM) de 2011. Bien que l'ENM ait remplacé le questionnaire complet du Recensement de 2006, l'EAPA est toujours considérée comme une enquête postcensitaire étant donné le lien entre l'ENM et le recensement. Plus précisément, l'échantillon de l'EAPA a été sélectionné à partir de répondants ayant fourni des réponses spécifiques aux quatre questions de sélection figurant dans le questionnaire de l'ENM, qui comportait deux versions principales, soit le questionnaire N1 et le questionnaire N2.

Le questionnaire N1 était rempli par auto-dénombrement et était distribué à environ 1 ménage sur 3 dans la plupart des régions du Canada (régions N1). Outre les questions démographiques de base du recensement (nom, sexe, date de naissance, état matrimonial légal, union libre, lien avec la personne 1, langues et question sur le consentement en vue de divulguer les données dans 92 ans), le questionnaire N1 de l'ENM contenait notamment des questions sur l'activité sur le marché du travail, le revenu, l'éducation, les limitations d'activités, la citoyenneté, le logement et l'origine ethnique.

Le questionnaire N2, dont le contenu était identique à celui du questionnaire N1, sauf pour certains exemples adaptés et certaines questions exclues[1], était rempli au moyen d'interviews sur place par tous les ménages des régions éloignées, des communautés inuites, des réserves indiennes et des établissements indiens (régions N2).

Les quatre questions de sélection utilisées pour identifier la population autochtone étaient les suivantes : celle sur l'origine ethnique (question 17); celle sur l'autodéclaration autochtone (question 18); celle visant à déterminer si la personne est un Indien avec statut (Indien inscrit ou des traités, question 20); et celle visant à déterminer si la personne est membre d'une Première Nation ou d'une bande indienne (question 21).

La déclaration d'une origine autochtone à la question 17 définit la population d'ascendance autochtone (ou la population d'ascendance).

La population d'identité autochtone dans le cadre de l'ENM est établie en fonction de trois questions (questions 18, 20 et 21 de l'ENM). Le concept d'identité autochtone fait référence aux personnes ayant : (a) déclaré appartenir à au moins un groupe autochtone, c'est-à-dire Première Nation (Indien de l'Amérique du Nord), Métis, Inuit; (b) indiqué être un Indien avec statut (Indien inscrit ou des traités aux termes de la *Loi sur les Indiens* du Canada) ou (c) déclaré être membre d'une Première Nation ou d'une bande indienne.

Dans le cadre de l'EAPA, les personnes ayant une ascendance autochtone mais ne possédant pas d'identité autochtone dans l'ENM composent la population d'ascendance autochtone seulement (ou population d'ascendance seulement). Aux fins d'échantillonnage de l'EAPA, la population autochtone comprend la population d'identité et la population d'ascendance seulement. Bien que, contrairement à l'EAPA de 2006, la population d'ascendance seulement ne faisait pas partie de la population cible de l'EAPA de 2012, cette population était tout de même échantillonnée à l'ENM. On a procédé ainsi après avoir constaté qu'en 2006, selon les données recueillies à l'aide du questionnaire complet, un peu moins du tiers de la population d'ascendance seulement du Recensement de 2006 avait déclaré l'identité autochtone lors de l'EAPA de 2006.

Le concept d'identité autochtone de l'EAPA est le même que pour l'ENM, mais celui-ci est défini selon un ensemble de questions légèrement différentes (voir le tableau 1 à la section 2.2).

Enquête nationale auprès des ménages de 2011 – N1 - question 17

Quelles étaient les origines ethniques ou culturelles des ancêtres de cette personne?

Cette question recueille des données sur les origines ancestrales afin de connaître la diversité de la population du Canada.

Habituellement, un ancêtre est plus éloigné que les grands-parents

Par exemple, canadien, anglais, français, chinois, indien de l'Inde, italien, allemand, écossais, irlandais, cri, mi'kmaq, salish, métis, inuit, philippin, hollandais, ukrainien, polonais, portugais, grec, coréen, vietnamien, jamaïquain, juif, libanais, salvadorien, somalien, colombien, etc.

1: Précisez toutes les origines qui s'appliquent en lettres majuscules.

Enquête nationale auprès des ménages de 2011 – N1 - question 18

Cette personne est-elle un Autochtone, c'est-à-dire Première Nation (Indien de l'Amérique du Nord), Métis ou Inuk (Inuit)?

Nota : Première Nation (Indien de l'Amérique du Nord) comprend les Indiens avec statut et les Indiens sans statut.

Si «Oui » cochez le ou les cercles qui décrivent le mieux cette personne maintenant et passez à la question 20.

1: Non, pas un Autochtone

2: Oui, Première Nation (Indien de l'Amérique du Nord)

1. Les questions sur la citoyenneté (question 10), le statut d'immigrant reçu (question 11) et l'année d'immigration (question 12) ne sont pas posées aux personnes vivant dans les réserves indiennes et dans les établissements indiens qui ont été dénombrées au moyen du questionnaire N2.

3: Oui, Métis

4: Oui, Inuk (Inuit)

Enquête nationale auprès des ménages de 2011 – N1 - question 20

Cette personne est-elle un Indien avec statut (Indien inscrit ou des traités aux termes de la Loi sur les Indiens du Canada)?

1: Non

2: Oui, Indien avec statut (Indien inscrit ou des traités)

Enquête nationale auprès des ménages de 2011 – N1 - question 21

Cette personne est-elle membre d'une Première Nation/bande indienne?

Si «Oui » de quelle Première Nation/bande indienne?

Par exemple, Atikamekw de Manawan, Première Nation de Sturgeon Lake, bande indienne Musqueam.

1: Non

2: Oui, membre d'une Première Nation/bande indienne (Précisez la Première Nation/bande indienne.)

3.1.2 Date de référence de l'enquête

Le 1er février 2012 a été retenu comme date de référence de l'EAPA. Cette date correspond approximativement au début de la collecte des données de l'enquête. L'âge est établi en fonction de cette date de référence et détermine l'enchaînement des questions à utiliser.

3.1.3 Base de données de l'Enquête nationale auprès des ménages

L'échantillon de l'EAPA a été tiré à partir de la base de données non contrôlées et non imputées de l'ENM. Il s'agit de la base de données de l'ENM appelée base de données des réponses (BDR), en anglais *response database* ou *RDB*. Avant de sélectionner l'échantillon, l'équipe de méthodologie d'enquête a développé une stratégie de contrôle et d'imputation pour traiter les valeurs manquantes aux quatre questions de sélection (ou « questions filtres ») de l'ENM ou les variables utilisées lors de la stratification (dont l'âge et certaines variables portant sur l'éducation), de même que pour les personnes présentant certaines caractéristiques qui semblaient incohérentes avec le fait d'être Autochtone.

La première étape du choix de l'échantillon consistait à inclure dans la base de sondage toutes les personnes de 6 ans et plus au 1er février 2012 ayant déclaré avoir une identité ou une ascendance autochtone.

À la deuxième étape, toutes les personnes appartenant aux mêmes ménages que les unités de la base initiale ont été ajoutées à la base. Ces autres personnes ont été conservées dans la base uniquement si leurs réponses aux questions filtres étaient manquantes et que les caractéristiques des membres du ménage ayant répondu aux questions filtres indiquaient qu'ils avaient une forte chance d'avoir une identité ou une ascendance autochtone. Ainsi, une personne dont les réponses étaient manquantes aux questions filtres sur l'identité serait habituellement imputée comme ayant l'identité autochtone si au moins 50 % des membres du même ménage ayant répondu aux questions filtres avaient une identité autochtones. De même, une personne n'ayant pas répondu à la question filtre sur l'ascendance serait habituellement imputée comme ayant une ascendance autochtone si au moins 50 % des membres du même ménage ayant répondu à la question filtre sur l'ascendance avaient une ascendance autochtone.

Il était évidemment très important de stratifier l'échantillon de l'EAPA de 2012 selon les groupes de scolarité décrits au chapitre 2 (les élèves actuels, de la 1re à la 6e année; les élèves actuels, de la 7e à la 12e année; les finissants de l'école secondaire; et les décrocheurs). Des estimations étaient ciblées pour chacun de ces quatre groupes. Étant donné que la situation d'une personne par rapport aux études peut avoir changé entre le moment de l'ENM de 2011 et le 1er février 2012, une méthode d'imputation déterministe a été élaborée pour déterminer le groupe de scolarité le plus probable pour chaque répondant en date du 1er février 2012.

Une fois le traitement complété, on a identifié et exclu de la base de sondage les personnes âgées de moins de 6 ans au 1er février 2012, celles vivant dans les réserves ou dans certaines communautés du Yukon et des Territoires du Nord-Ouest ainsi que les répondants devenus non-Autochtones à la suite du traitement des données.

3.2 Plan d'échantillonnage

3.2.1 Domaines d'estimation

Pour être efficace, la stratification doit utiliser des domaines d'estimation. Les domaines d'estimation sont les groupes d'unités pour lesquels des estimations sont ciblées. Dans le cas de l'EAPA, des domaines d'estimation spécifiques à la stratification ont été utilisés. Ces domaines d'estimation correspondaient à des régions géographiques pour lesquelles des estimations d'un niveau de précision « acceptable » étaient visées pour un groupe autochtone en particulier (c.-à-d. Première Nation, Métis ou Inuit) et un groupe de scolarité précis.

Un domaine d'estimation peut être, par exemple, les Métis de l'Alberta fréquentant l'école primaire. Lors de la stratification, le groupe autochtone composé des Métis était formé de personnes ayant déclaré une identité métisse unique à la question 18 de l'ENM ou, chez les répondants ne possédant pas l'identité autochtone (population d'ascendance autochtone seulement), une ascendance métisse unique à la question 17 (avec ou sans ascendance non autochtone). En réalité, les personnes d'ascendance autochtone seulement ne faisaient pas partie de la population cible de l'enquête. Toutefois, ceux-ci ont été échantillonnés, car ils avaient une chance relativement importantede déclarer avoir une identité autochtone dans l'enquête, tel qu'il est décrit à la sous-section 3.1.1. C'est la raison pour laquelle on parle ici de « domaines d'estimation spécifiques à la stratification » plutôt que de « domaines d'estimation spécifiques à l'enquête ».

Plus précisément, les domaines d'estimation spécifiques à la stratification sont définis par le croisement des variables suivantes :

- Géographie
  - Régions inuites,
  - Hors régions inuites
    - Province ou territoire
    - Provinces de l'Atlantique regroupées
- Groupe de scolarité
  - Élèves actuels, de la 1re à la 6e année
  - Élèves actuels, de la 7e à la 12e année
  - Finissants de l'école secondaire, avec un diplôme d'études secondaires (ou l'équivalent)
  - Décrocheurs de l'école secondaire, sans diplôme d'études secondaires et ne fréquentant pas l'école de la 1re à la 12e année
- Groupe autochtone
  - Inuit dans les régions inuites
  - Inuit hors régions inuites (reste du Canada)
  - Groupes autochtones combinés pour le Canada atlantique, le Québec (hors Nunavik), le Yukon et les Territoires du Nord-Ouest (hors Inuvialuit)
  - Pour l'Ontario, le Manitoba, la Saskatchewan, l'Alberta et la Colombie-Britannique
    - Première Nation avec statut vivant hors réserve
    - Première Nation sans statut vivant hors réserve
    - Métis

Voici la façon dont les groupes autochtones ont été définis lors de la stratification en fonction des réponses indiquées dans l'ENM :

- Première Nation avec statut vivant hors réserve — personnes répondant uniquement Première Nation à la question 18 de l'ENM et ayant répondu OUI à la question 20 de l'ENM
- Première Nation sans statut vivant hors réserve — personnes répondant uniquement Première Nation à la question 18 et ayant répondu NON à la question 20 ou personnes n'ayant pas une identité autochtone, mais répondant Première Nation uniquement à la question 17 (avec ou sans ascendance non autochtone)
- Métis seulement — personnes répondant Métis seulement à la question 18 ou personnes n'ayant pas une identité autochtone, mais répondant Métis uniquement à la question 17 (avec ou sans ascendance non autochtone)
- Inuit seulement — personnes répondant Inuit seulement à la question 18 ou personnes n'ayant pas une identité autochtone, mais répondant Inuit uniquement à la question 17 (avec ou sans ascendance non autochtone)
- Groupe autochtone multiple — personnes déclarant plus d'une identité à la question 18 ou personnes n'ayant pas une identité autochtone, mais déclarant plus d'une ascendance autochtone à la question 17 (avec ou sans ascendance non autochtone)
- Indien avec statut ou membre d'une Première Nation ou d'une bande indienne seulement — personnes ayant répondu OUI à la question 20 ou à la question 21, mais NON à la question 18.

Afin de s'assurer de la fiabilité des estimations, certains groupes de scolarité ou groupes autochtones ont été combinés pour des régions données. Le fait de combiner les groupes de scolarité a permis de réduire leur nombre, qui est passé de quatre à deux groupes; les personnes fréquentant l'école primaire et celles fréquentant l'école secondaire ont été regroupées, et les finissants (diplômés) l'ont été avec les décrocheurs. Au Canada atlantique excluant le Nunatsiavut, au Québec excluant le Nunavik, au Yukon et dans les Territoires du Nord-Ouest, le nombre de Métis est généralement trop faible pour produire des estimations distinctes pour les Premières Nations et les Métis selon le groupe de scolarité. C'est la raison pour laquelle tous les groupes autochtones ont été combinés lors de la stratification.

Bien que les estimations ciblaient les Inuit hors régions inuites à l'échelle nationale, des domaines d'estimation ont été créés « artificiellement »pour les Inuit vivant dans le groupe formé des provinces allant de l'Ontario jusqu'à la Colombie-Britannique (pour les autres provinces et territoires hors régions inuites, tous les groupes autochtones sont combinés). Étant donné que les Inuit vivant dans les provinces de l'Ontario jusqu'à la Colombie-Britannique représentent un petit groupe, des estimations relativement imprécises ont été ciblées pour deux groupes de scolarité seulement. L'ajout des Inuit des autres provinces et territoires hors régions inuites devait permettre d'obtenir des estimations plus précises pour les Inuit hors régions inuites à l'échelle nationale.

En tout, 88 domaines d'estimation cibles et 92 domaines d'estimation complémentaires ont été créés, soit au total 180 domaines d'estimation. Les domaines d'estimation complémentaires comprenaient des domaines pour lesquels la production d'estimations n'était pas prévue dans le cadre de l'EAPA mais comportant un groupe particulier qui devait tout de même être représenté dans l'échantillon. Ainsi, les non-Inuit vivant dans les régions inuites et les personnes de 45 ans et plus constituent des exemples de domaines d'estimation complémentaires.

L'objectif pour chaque domaine cible était d'estimer une caractéristique présente pour au moins 10 % des personnes dans les groupes de scolarité associés aux enfants et aux jeunes (élèves actuels de la 1re à la 6e année et ceux de la 7e à la 12e année) et pour au moins 7,5 % des répondants faisant partie des groupes de scolarité associés aux adultes (finissants et décrocheurs), avec un coefficient de variation (c.v.) donné. Le coefficient de variation est une mesure de précision de l'estimation qui est décrite à la section 7.2 (*Erreurs d'échantillonnage et méthode bootstrap*). La proportion minimale visée est appelée *min-p*. La capacité d'atteindre les c.v. ciblés pour une valeur de *min-p* donnée dépendait de facteurs comme la taille de la population, le nombre de répondants disponibles à l'ENM, le taux de réponse prévu, le nombre prévu de faux positifs (personnes ayant déclaré dans l'ENM avoir l'identité ou l'ascendance autochtone mais sans avoir l'identité autochtone selon l'EAPA, soit un type très particulier de répondants hors de la population cible), la perte prévue de l'échantillon en raison de la contrainte de sélection d'au maximum trois personnes par ménage ainsi que la perte prévue découlant de la réduction du chevauchement avec d'autres enquêtes. Selon la taille des domaines, les valeurs des c.v. ont été fixées à 20 %, 22,5 %, 25 % ou 33 %.

Le tableau ci-dessous présente les c.v. et *min-p* visés pour chaque domaine d'estimation cible.

Tableau 3 Coefficients de variation et valeurs de *min-p* selon le domaine d'estimation

Région	Groupe autochtone					Groupe d'éducation					
	PN avec statut	PN sans statut	Métis	Inuit	Tous	1e à 6e année (*min-p*=10 %)	7e à 12e année (*min-p*=10 %)	Diplômé (*min-p*=7,5 %)	Décrocheur (*min-p*=7,5 %)	1e à 12e année (*min-p*=10%)	Diplômé ou décrocheur (*min-p*=7,5 %)
						pourcentage					
Nunatsiavut				x						20	20
Nunavik				x		20	20	20	20		
Inuvialuit				x						20	20
Nunavut				x		20	20	20	20		
Inuit Ont-C.B.				x						33	33
Atlantique hors Nunatsiavut					x	20	20	20	20		
Québec hors Nunavik					x	20	20	20	20		
Ontario	x				x	25	25	25	25		
		x				25	25	25	25		
			x			22,50	22,50	22,50	22,50		
Manitoba	x					25	25	25	25		
		x				33	33	33	40		
			x			22,50	22,50	22,50	22,50		
Saskatchewan	x					25	25	25	25		
		x				33	33	33	40		
			x			22,50	22,50	22,50	22,50		
Alberta	x					25	25	25	25		
		x				25	25	25	33		
			x			22,50	22,50	22,50	22,50		
C.-B.	x					25	25	25	25		
		x				25	25	25	33		
			x			22,50	22,50	22,50	22,50		
Yukon					x					25	25
T.N.-O. hors Inuvialuit					x	25	25	25	25		

À titre d'exemple, au Nunatsiavut, les estimations étaient ciblées chez les Inuit seulement dans deux groupes de scolarité : les élèves actuels de la 1re à la 12e année et les finissants et les décrocheurs combinés. On espérait pouvoir estimer, avec un c.v. de 20 %, une proportion minimale de 10 % dans le premier groupe et une de 7,5 % dans le deuxième groupe. En Ontario, on ciblait des estimations pour les Premières Nations avec et sans statut séparément, de même que pour les Métis, et ce, dans les quatre groupes de scolarité. Chez les Premières Nations avec et sans statut, on visait des c.v. de 25 %, alors que chez les Métis faisant partie de chacun des groupes de scolarité, les c.v. visés étaient de 22,5 %.

Évidemment, tout regroupement de domaines entraînera des c.v. plus petits. En raison des estimations distinctes pour les Premières Nations avec et sans statut vivant dans les provinces de l'Ontario jusqu'à la Colombie-Britannique, des c.v. de 25 % étaient visés pour ces deux groupes, ce qui devrait produire des c.v. inférieurs à 20 % pour les Premières Nations avec ou sans statut. Des c.v. de 22,5 % étaient visés dans le cas des Métis vivant dans les provinces de l'Ontario jusqu'à la Colombie-Britannique.

3.2.2 Stratification

Une stratification permettra de produire des estimations plus précises si les unités sont homogènes à l'intérieur des strates et hétérogènes entre les strates. De plus, il est idéalement souhaitable que les poids d'estimation associés aux répondants de l'enquête soient les plus rapprochés possibles à l'intérieur des strates

Le plan de l'ENM en est un à deux phases dans les régions où le questionnaire N1 a été distribué (dans la plupart des régions du Canada, voir la sous-section 3.1.1). Dans ces régions, un échantillon systématique d'environ 1 ménage sur 3 a été tiré lors de la première phase. Le 14 juillet 2011, ceux qui avaient déjà répondu formaient les répondants initiaux. Les non-répondants à cette date constituaient le groupe des *non-répondants initiaux*. Un sous-échantillon d'environ 1 ménage sur 3 de ces *non-répondants initiaux* (pour une fraction de sondage combinée d'environ 1 sur 9) a été sélectionné pour le suivi des cas de non-réponse (SCNR). Environ 40 % de ces personnes ont répondu dans le cadre du SCNR. Par conséquent, les poids associés aux répondants de l'ENM (ou plus simplement poids de l'ENM) dans les régions où le questionnaire N1 a été distribué varient beaucoup. Par contre, dans les régions où le questionnaire N2 a été distribué (ce questionnaire a été distribué à tous les ménages dans les régions sans sous-échantillon du SCNR), seule la non-réponse (et non l'échantillonnage) a fait varier les poids de l'ENM. Ainsi, même avec un taux de réponse de 50 % dans une région N2 donnée, le poids de l'ENM ne serait que de 2. Une description plus détaillée du plan d'échantillonnage de l'ENM se trouve au chapitre 3 du *Guide de l'utilisateur de l'Enquête nationale auprès des ménages*.

L'une des hypothèses utilisées dans la méthode de répartition de l'échantillon de l'EAPA était que les poids de l'ENM varieraient le moins possible à l'intérieur d'une strate de l'EAPA. Pour ce faire, seulement une partie de la stratification de l'EAPA provenait de la stratification de l'ENM. Comme on ne pouvait tenir compte de toute la stratification de l'ENM, il fallait choisir les variables de stratification qui causaient la plus grande variation possible des poids entre les strates.

Comme il a été mentionné précédemment, le type de région (N1 ou N2) associé au type de questionnaire de l'ENM et à la méthode de collecte ont créé une différence significative dans les poids de l'ENM. Une répartition optimale entre les régions N1 et N2 aurait été de tirer proportionnellement plus de répondants de l'EAPA dans les régions N1 que dans les régions N2 afin d'équilibrer les poids combinant l'échantillonnage de l'ENM et celui de l'EAPA.

Un autre facteur qui fait varier les poids de l'ENM dans les régions N1 seulement est le fait d'appartenir aux répondants initiaux ou aux répondants de l'échantillon du SCNR. On peut démontrer que les poids des répondants au SCNR sont en moyenne 7,5 fois plus élevés que ceux des répondants initiaux. Une des façons de contrebalancer cet effet serait de tirer en proportion 7,5 fois plus de personnes dans les strates de répondants au SCNR que dans les strates de répondants initiaux. Évidemment, une telle méthode pourrait présenter de nombreux désavantages, tels que de ne pas pouvoir tirer assez de répondants au SCNR et le fait que les probabilités de réponse à l'EAPA sont certainement moindres chez les répondants au SCNR que chez les répondants initiaux. Certains compromis ont dû être effectués lors de la répartition de l'échantillon.

Un autre facteur, qui n'est pas lié à la variabilité des poids de l'ENM mais qui était très important à considérer comme facteur de stratification, était le type d'identification autochtone (ou plus simplement *type d'Autochtone*) des répondants sélectionnés à partir de l'ENM. Ce facteur correspond au fait que les répondants à l'ENM ont indiqué soit l'identité autochtone (ID) ou l'ascendance autochtone seulement (AS) dans l'ENM. Ces deux groupes ont des caractéristiques très différentes. De plus, la répartition doit utiliser la probabilité de chaque unité d'appartenir à la population cible (avoir l'identité autochtone à l'EAPA). À l'échelle nationale, selon les données de l'EAPA de 2006, les personnes ayant l'identité à l'ENM ont une probabilité d'environ 88 % d'avoir l'identité autochtone à l'EAPA, alors que les personnes ayant l'ascendance seulement à l'ENM ont une probabilité d'environ 32 % d'avoir l'identité à l'EAPA.

En combinant les facteurs type de région (N1 ou N2), type de répondant à l'ENM (répondant initial ou de l'échantillon du SCNR) et type d'Autochtone (AS ou ID), on obtenait un maximum de six strates par domaine d'estimation :

- N1, répondant initial, AS
- N1, répondant initial, ID
- N1, répondant SCNR, AS
- N1, répondant SCNR, ID
- N2, AS
- N2, ID

Pour certains domaines comme les Premières Nations avec statut, il n'y avait que trois strates possibles, puisque par définition les personnes incluses avaient une identité autochtone. Il peut également y avoir quelques strates vides, surtout dans les domaines complémentaires. Par exemple, pour une région inuite donnée, il était possible qu'aucune des personnes classifiées comme ayant une ascendance autochtone seulement ait une ascendance autochtone autre qu'inuite.

3.2.3 Plan d'échantillonnage et répartition de l'échantillon

L'échantillon de l'EAPA de 2011 a été sélectionné à partir des répondants de l'ENM de 2011. À la sous-section 3.2.2, il a été mentionné que l'ENM utilisait un plan d'échantillonnage à deux phases dans les régions N1 ainsi qu'un plan d'échantillonnage à une seule phase dans les régions N2. Ainsi, celui de l'EAPA peut être considéré comme un plan d'échantillonnage à trois phases dans les régions N1 et à deux phases dans les régions N2.

Une fois la base de sondage élaborée, elle a été stratifiée selon les domaines d'estimation puis sous-stratifiée en fonction du type de région (N1 vs N2), du type de répondant à l'ENM (répondant initial vs répondant au SCNR) et du type d'Autochtone (ID vs AS). Un échantillon aléatoire systématique a été par la suite sélectionné à l'intérieur de chacune des strates, la base ayant été préalablement triée par numéro de ménage et de personne. Cette mesure visait à bien répartir l'échantillon géographiquement à l'intérieur des strates de même qu'à répartir l'échantillon sur le plus grand nombre de ménages possible. Cette procédure était toutefois limitée du fait que les membres d'un même ménage pouvaient se trouver dans des strates différentes.

Une méthode de répartition optimale entre les sous-strates d'un domaine particulier a été utilisée en tenant compte de différentes formes de pertes de taille d'échantillon telles qu'elles sont décrites à la sous-section 3.2.4, de même que la probabilité de chaque unité d'appartenir à la population cible dans une strate donnée. Cette répartition dépendait, entre autres, des poids de l'ENM. Il est à noter qu'au moment de faire la répartition, ces poids n'avaient pas encore été calculés. Des poids préliminaires ont donc été calculés pour la répartition seulement. Les poids définitifs calculés par l'équipe de méthodologie travaillant sur les estimations de l'ENM ont été utilisés lors de la pondération (voir le chapitre 6). Pour en savoir plus à ce sujet, veuillez consulter l'article de Verret (2013)[2].

La répartition a d'abord été effectuée dans les domaines cibles de l'enquête. Pour les domaines non ciblés ou domaines complémentaires (p. ex. les non-Inuit des régions inuites ou les personnes de 45 ans et plus chez les finissants et les décrocheurs), la taille de l'échantillon dans une strate donnée est calculée en prenant une fraction de sondage égale ou inférieure à la fraction de sondage de la strate correspondante du domaine cible correspondant. Ces domaines complémentaires devaient être retenus pour couvrir toute la population cible mais ne nécessitaient pas une taille de l'échantillon déterminée puisqu'aucune estimation n'était visée pour ces domaines. Pour les groupes autochtones non ciblés (p. ex. les non-Inuit des régions inuites), on a utilisé la même fraction de sondage que chez les Inuit. Dans le cas des personnes de 45 ans et plus, la moitié de la fraction de sondage des personnes de moins de 45 ans a été utilisée.

Bien qu'on ait prévu sélectionner l'échantillon de l'EAPA de 2012 en une seule vague, après quelques semaines de collecte de données, il a été décidé de sélectionner un échantillon supplémentaire (appelé échantillon de vague 2). Trois raisons ont mené à cette décision :

- Comme la BDR (base de sondage de l'EAPA) n'était pas entièrement révisée au moment de la sélection de l'échantillon pour l'EAPA, un certain nombre de personnes apparaissaient par erreur comme Indien avec statut ou membre d'une Première Nation ou d'une bande indienne seulement (voir la sous-section 3.2.1). L'échantillon initial de l'EAPA comportait 1 254 de ces cas, appelés « faux Autochtones » dans le présent document;
- Le taux de réponse des décrocheurs était inférieur à celui des autres groupes et aussi plus faible que prévu;
- La proportion de l'échantillon du Québec faisant partie de la population cible (taux d'identité autochtone à l'enquête) était moins élevée que prévu aussi bien pour les personnes ayant déclaré une identité autochtone à l'ENM que pour celles ayant déclaré une ascendance autochtone seulement.

2. Verret, François. 2013. *The estimation methodology of the 2011 National Household Survey*. Article présenté au Joint Statistical Meeting de 2013.

Pour calculer la taille de l'échantillon supplémentaire, les « faux Autochtones » ont été exclus de la base de sondage. Les taux de réponse chez les décrocheurs ont été ajustés à la baisse, et les taux d'identité à l'enquête ont été revus pour le Québec en fonction des résultats observés au début de la collecte des données. La même méthode de répartition a ensuite été utilisée afin de déterminer de nouvelles tailles d'échantillon. La taille de l'échantillon ainsi obtenue représente celle qui aurait été utilisée si les « faux Autochtones » n'avaient jamais existés dans la base, et si l'on avait pu estimer de façon plus précise les taux de réponse chez les décrocheurs de même que les taux d'identité au Québec. La différence entre cette taille de l'échantillon et celle de l'échantillon initial constitue la taille de l'échantillon supplémentaire. Aucune unité additionnelle n'a été sélectionnée dans les domaines complémentaires. Après le retrait des « faux Autochtones », l'échantillon initial comportait 49 287 unités et l'échantillon supplémentaire, 1 737 unités, ce qui donne un total de 51 024 unités.

3.2.4 Ajustements de la taille de l'échantillon

a) Probabilité d'appartenir à la population cible

Selon l'EAPA de 2006, 88 % des personnes ayant déclaré une identité autochtone dans le cadre du Recensement de 2006 ont déclaré une identité autochtone à l'EAPA de 2006, et 32 % des personnes ayant déclaré une ascendance autochtone seulement au Recensement de 2006 ont déclaré une identité autochtone dans le cadre de l'EAPA de 2006. En raison des changements apportés à la méthodologie et au questionnaire et des changements liés au contexte social, on s'attendait à ce que les probabilités d'appartenance à la population cible en 2012 soient inférieures à celles de 2006. C'est pourquoi les probabilités d'appartenir à la population cible de 2012 ont été ajustées à la baisse, en multipliant les chiffres de 2006 par 75 % pour les strates d'ascendance autochtone seulement de l'ENM et par 95 % pour les strates d'identité autochtones de l'ENM.

b) Taux de réponse

Les taux de réponse de l'EAPA de 2006 ont été utilisés pour répartir l'échantillon de l'EAPA de 2012. À titre préventif, ces taux ont été réduits en les multipliant par un facteur d'ajustement de 90 %. Mentionnons que le fait d'augmenter la taille de l'échantillon en fonction d'un taux de réponse prévu constitue une mesure de protection visant à diminuer l'erreur d'échantillonnage. Toutefois, cela ne joue pas un rôle pour ce qui est de réduire le biais de non-réponse, un type particulier d'erreur non due à l'échantillonnage (voir la section 7.3). Il est évidemment préférable d'obtenir un taux de réponse élevé pour un plus petit échantillon que d'obtenir un taux de réponse plus faible pour un échantillon plus grand.

c) Transition entre les groupes de scolarité

Les estimations étont été ciblées pour quatre groupes de scolarité (ou parfois deux) selon la région et le groupe autochtone. Cependant, la classification des répondants entre ces groupes a été effectuée à partir des variables de la scolarité et de l'âge, qui sont parfois inexactes ou manquantes dans l'ENM. Dans de nombreux cas, particulièrement en ce qui concerne les enfants, le niveau de scolarité ne peut être estimé que de façon approximative 'à partir des données de l'ENM (en fait, seule la date de naissance est utilisée pour déterminer le groupe de scolarité des enfants âgés de moins de 15 ans). Par ailleurs, l'évolution au fil du temps fait en sorte que les données de l'ENM ne permettent pas de prédire exactement le groupe de scolarité des répondants à l'EAPA au 1er février 2012, date de référence de l'enquête. Des facteurs de perte liés aux transitions entre les groupes ont été estimés pour l'EAPA de 2012, toujours à l'aide des données de 2006. Pour chacun des répondants de l'EAPA de 2006, le groupe de scolarité a été calculé de façon approximative à partir de la BDR du Recensement de 2006 et l'établissement des catégories a été comparé aux groupes de scolarité obtenus lors de l'EAPA de 2006.

Les résultats de l'EAPA de 2006 avaient révélé qu'un peu plus de répondants que prévu fréquentaient l'école primaire (2,5 % de plus) et que moins de répondants que prévu allaient à l'école secondaire (8,5 % de moins). Selon ces résultats, il y avait un peu moins de finissants que prévu (5,8 % de moins) et plus

de décrocheurs que prévu (15 % de plus). Des facteurs d'ajustement des tailles ont été calculés selon la région, le groupe autochtone et le groupe de scolarité. Ces facteurs ont été multipliés par 1,05 par mesure de précaution. Si un facteur demeurait plus petit que 1, on lui donnait la valeur de 1 pour ne pas diminuer les tailles d'échantillon. Par exemple, même si l'on s'attendait à obtenir plus de décrocheurs à l'EAPA qu'à l'ENM, on n'a pas diminué la taille du groupe de décrocheurs par souci de précaution. Ces facteurs ont été inclus dans le processus de répartition afin d'augmenter ou non la taille de l'échantillon en conséquence.

d) Ménages comprenant plus de trois membres sélectionnés

Étant donné que l'unité d'échantillonnage était la personne et non le ménage, plusieurs membres d'un même ménage pouvaient faire partie de l'échantillon. De façon à limiter le fardeau de réponse, l'une des contraintes imposées pendant la collecte était de ne pas sélectionner plus de trois personnes par ménage. Pour estimer cette perte, on a tiré un échantillon préliminaire sur lequel on a calculé le nombre de personnes qui auraient été exclues en raison de cette contrainte. La perte relative obtenue a été multipliée par 1,1 pour chaque combinaison de domaines et strates et a été incorporée au facteur d'ajustement de la taille de l'échantillon.

e) Chevauchement avec d'autres enquêtes

La période de collecte des données de l'EAPA chevauche celle de plusieurs autres enquêtes menées par Statistique Canada. On considère généralement qu'une personne contactée pour répondre à une enquête est moins susceptible d'accepter d'y participer si elle vient tout juste d'être interviewée pour une autre enquête. De façon à compenser pour cette perte éventuelle, la taille totale de l'échantillon de l'EAPA de 2012 a été augmentée d'environ 5 % de façon uniforme. Parmi ces enquêtes, plusieurs couvraient aussi les territoires où la population autochtone représente une fraction élevée de la population totale. Les possibilités de chevauchement dans les territoires entre les différentes enquêtes étaient donc plus élevées. Une attention particulière a été accordée à ce sujet, afin de réduire le plus possible le chevauchement et, implicitement, le fardeau potentiel imposé aux répondants. Le chevauchement dans les territoires a été étudié pour l'Enquête sur la santé dans les collectivités canadiennes (ESCC), l'Enquête sur la population active (EPA) et le Programme pour l'évaluation internationale des compétences des adultes (PEICA); des mesures ont été prises afin de réduire le fardeau de réponse lorsque c'était possible.

Ainsi, le chevauchement avec l'ESCC était particulièrement préoccupant. Certaines questions sur la santé de l'EAPA sont directement tirées du questionnaire de l'ESCC, dont plusieurs questions de nature délicate sur la santé mentale. Aucune mesure n'a été prise pour les interviews de l'ESCC menées avant et après la collecte des données de l'EAPA, soit avant janvier 2012 et après juin 2012. Par contre, le chevauchement identifiable des ménages devant être interviewés entre janvier et juin 2012 pour l'ESCC a été éliminé. Mentionnons qu'un appariement entre la base aréolaire de l'EPA (duquel l'échantillon de l'ESCC est tiré) avec le recensement avait été fait et qu'environ 70 % des ménages de la base aréolaire ont pu être appariés avec le recensement. Comme l'ESCC tire un échantillon aléatoire systématique de ménages à l'intérieur de chaque communauté couverte, les ménages non sélectionnés forment aussi un échantillon aléatoire systématique de ménages. Les ménages sélectionnés de janvier à juin à l'ESCC ayant répondu à l'ENM ont été exclus de la base de sondage de l'EAPA. À l'intérieur de chaque communauté couverte par l'ESCC, les poids des ménages restants de l'ENM ont été ajustés pour représenter tous les ménages de la communauté. Puisque ces ménages sont exclus de la base de sondage de l'EAPA, ils n'ont donc aucune chance d'être choisis.

3.3 Taille de l'échantillon et taux de réponse

Les échantillons initial et supplémentaire de l'EAPA de 2012 contenaient en tout 51 024 unités, dont 376 unités retirées avant d'être envoyées à la collecte (voir la section 6.2). La répartition des 50 648 unités échantillonnées selon le domaine géographique de même que les taux de réponse correspondants sont présentés au tableau 4.

Un taux de réponse est défini comme étant le nombre de répondants admissibles divisé par le nombre d'unités admissibles dans l'échantillon. Les personnes demeurant à l'extérieur du Canada ou vivant en établissement au moment de la tenue de l'enquête ainsi que celles âgées de moins de 6 ans au 1er février 2012 sont des exemples d'unités inadmissibles pour l'EAPA.

Dans le cadre de l'EAPA, les unités admissibles étaient définies de deux façons. Dans la première définition, les personnes n'ayant pas l'identité autochtone à l'EAPA étaient considérées comme inadmissibles (en plus des autres types d'unités inadmissibles). Dans la deuxième définition, ces personnes étaient considérées comme admissibles et faisaient partie des répondants. En effet, ces personnes ont accepté de participer à l'enquête et ont répondu à toutes les questions qui leur ont été posées, soit les questions filtres déterminant s'ils avaient ou non l'identité autochtone.

Ces deux définitions d'unités et de répondants admissibles donnent donc lieu à deux façons de mesurer le taux de réponse. Le premier taux de réponse (70,4 %) est généralement celui qui est utilisé au moment de la collecte. Il s'agit du taux de réponse établi en fonction des unités échantillonnées qui font partie de la population d'identité autochtone selon l'EAPA. Le second taux de réponse (76,2 %) est celui dont se sert l'équipe de la Méthodologie; il correspond davantage au taux de réponse établi en fonction de toutes les unités échantillonnées (avec certaines exclusions, cependant). Bien sûr, dans les deux cas, les unités non répondantes ne peuvent pas toujours être définies comme étant « admissibles » ou « inadmissibles ». La probabilité de déclaration d'une identité autochtone à l'EAPA diffère considérablement selon si les répondants ont déclaré une identité autochtone à l'ENM ou l'ascendance autochtone seulement à l'ENM. Par conséquent, le taux de réponse utilisé durant la collecte est très influencé par la répartition des unités d'échantillon entre ces deux groupes, ce qui n'est pas le cas pour le taux de réponse utilisé par l'équipe de méthodologie. Dans les régions inuites, comme la plupart des répondants ont l'identité autochtone, les deux taux de réponse sont très similaires, ce qui n'est pas le cas pour les autres régions. La différence maximale entre les deux taux est observée au Québec excluant le Nunavik. Dans cette région, la population d'ascendance autochtone seulement fforme une grande partie de la population autochtone totale. Selon les données de l'EAPA de 2006, les probabilités d'avoir l'identité autochtone à l'enquête sont les plus faibles du pays, et ce, peu importe que le répondant ait l'ascendance autochtone seulement (20 %) ou qu'il ait l'identité autochtone à l'ENM (66 %). En conséquence, même si le taux de réponse utilisé par la Méthodologie est le plus élevé au Québec, le taux de réponse utilisé durant la collecte correspondant est relativement bas.

Tableau 4 Taille de l'échantillon et taux de réponse par domaine géographique, selon deux définitions

	Total	Éligibles 1	Éligibles 2	Répondants 1	Répondants 2	Taux de réponse 1	Taux de réponse 2
Domaine géographique	nombre d'unités					pourcentage	
Nunatsiavut	777	752	759	604	611	80.3	80.5
Nunavik	2,084	2,000	2,021	1,543	1,564	77.2	77.4
Inuvialuit	1,153	1,117	1,139	791	813	70.8	71.4
Nunavut	2,181	2,128	2,156	1,581	1,609	74.3	74.6
Nunangat (total)	**6,195**	**5,997**	**6,075**	**4,519**	**4,597**	**75.4**	**75.7**
Atlantique excluant le Nunatsiavut	3,352	2,640	3,336	2,025	2,721	76.7	81.6
Québec excluant le Nunavik	6,776	3,155	6,724	2,317	5,886	73.4	87.5
Ontario	8,366	6,299	8,306	4,286	6,293	68.0	75.8
Manitoba	5,190	4,589	5,136	3,006	3,553	65.5	69.2
Saskatchewan	4,946	4,412	4,860	3,124	3,572	70.8	73.5
Alberta	6,603	5,477	6,547	3,765	4,835	68.7	73.9
Colombie-Britannique	6,963	5,707	6,901	3,918	5,112	68.7	74.1
Yukon	786	712	764	489	541	68.7	70.8
Territoires du Nord-Ouest excluant l'Inuvialuit	1,471	1,365	1,440	960	1,035	70.3	71.9
Reste du Canada (total)	**44,453**	**34,356**	**44,014**	**23,890**	**33,548**	**69.5**	**76.2**
Canada (total)	**50,648**	**40,353**	**50,089**	**28,409**	**38,145**	**70.4**	**76.2**

Éligibles 1 Taille d'échantillon excluant les hors champs à l'EAPA (les non Autochtones sont hors champs sous cette définition)
Éligibles 2 Taille d'échantillon excluant les hors champs à l'EAPA (les non Autochtones à l'EAPA ne sont pas hors champs sous cette définition)
Répondants 1 Nombre de répondants excluant les non Autochtones à l'EAPA
Répondants 2 Nombre de répondants incluant les non Autochtones à l'EAPA
Taux de réponse 1 Taux de réponse excluant les non Autochtones à l'EAPA
Taux de réponse 2 Taux de réponse incluant les non Autochtones à l'EAPA

4. Collecte des données

4.1 Échéancier

L'Enquête auprès des peuples autochtones de 2012 est une enquête postcensitaire/post-ENM qui a été menée du 6 février au 30 juillet 2012. Quant à l'Enquête nationale auprès des ménages de 2011, elle a eu lieu de mai à août 2011.

4.2 Mode de collecte

Les questions de l'EAPA de 2012 ont été posées au moyen d'une application d'interview assistée par ordinateur (IAO). Deux méthodes d'interview assistée par ordinateur ont été utilisées : l'interview téléphonique assistée par ordinateur (ITAO) et l'interview sur place assistée par ordinateur (IPAO). Dans la plupart des régions, l'ITAO a été utilisée dans le cas des personnes pour lesquelles un numéro de téléphone figurait dans le fichier de l'échantillon. On a eu recours à l'IPAO pour les personnes qui n'avaient pas de numéro de téléphone ou qu'il était impossible de joindre par téléphone, et ce, même si elles avaient un numéro. Dans les territoires, les régions nordiques de plusieurs provinces et certaines communautés inuites, très peu de numéros de téléphone étaient disponibles. Dans ces situations, des interviews sur place ont été menées.

Les répondants ont été interviewés dans la langue officielle de leur choix. Dans les régions inuites, le questionnaire a été traduit sous forme de copie papier en inuktitut (dialecte de Baffin). De plus, un enregistrement audio en inuktitut du questionnaire a été préparé afin d'aider les intervieweurs qui pouvaient faire face à des barrières linguistiques sur le terrain.

Le temps nécessaire pour répondre au questionnaire variait d'une personne à l'autre. Dans certains cas, l'interview de l'EAPA de 2012 a duré plus d'une heure. Toutefois, il fallait en moyenne environ 40 minutes pour répondre à l'enquête.

4.3 Supervision et contrôle de la qualité

Tous les intervieweurs de Statistique Canada travaillaient sous la supervision d'intervieweurs principaux. Ceux-ci doivent s'assurer que les intervieweurs connaissent bien les concepts de l'enquête et les procédures qui y sont associées. Les intervieweurs principaux ont aussi pour fonction de faire le suivi régulier du travail des intervieweurs.

Les intervieweurs ont reçu de la formation sur le contenu de l'enquête et l'application d'IAO. En plus de la formation en classe, les intervieweurs ont mené une série d'interviews simulées afin de se familiariser avec l'enquête, ses concepts et ses définitions.

4.4 Interviews par personne interposée

Les questions destinées aux enfants et aux jeunes sélectionnés de 6 à 14 ans ont été posées au moyen d'interviews par personne interposée, c'est-à-dire la « personne la mieux renseignée » (PMR) au sujet de la personne sélectionnée. La PMR était généralement le parent ou le tuteur de l'enfant. Dans le cas des jeunes de 15 à 17 ans, les interviews ont été faites directement auprès d'eux seulement s'il y avait consentement préalable du parent ou du tuteur. En l'absence de consentement, les données ont été recueillies auprès du parent ou du tuteur.

Les questionnaires destinés aux adultes ont été conçus pour obtenir les réponses des répondants eux-mêmes. Les interviews par personne interposée ont été permises dans certains cas, notamment lorsque la personne sélectionnée n'était pas en mesure de répondre en raison d'un problème de santé mentale ou physique, d'une barrière linguistique ou parce qu'elle était absente au moment de la tenue de l'enquête.

4.5 Stratégie de communication

Au cours des mois qui ont mené à la collecte des données, des efforts ont été déployés afin de mieux faire connaître l'EAPA de 2012 et de favoriser la participation à l'enquête. La stratégie de communication comprenait la conception d'une brochure et d'affiches de l'EAPA. Les 11 conseillers du Programme de liaison avec les Autochtones de Statistique Canada ont servi de personnes-ressources pour l'EAPA dans les diverses régions du pays. Ils ont mis des affiches, distribué des brochures et fait circuler l'information concernant l'enquête au moyen de courriels, de bulletins d'information et de sites Web. Les conseillers ont également rencontré les organismes locaux et provinciaux et ont visité des Centres d'amitié partout au Canada. Les 119 Centres d'amitié ont exposé des affiches de l'EAPA, distribué des brochures et maintenu à jour leurs sites Web concernant les données de l'EAPA. Le Ralliement national des Métis (RNM) et le Congrès des Peuples Autochtones (CPA) ont aussi collaboré en vue de promouvoir l'enquête.

Avant la collecte des données, des lettres et des brochures ont été envoyées par la poste à certains répondants pour présenter l'objectif de l'enquête et souligner l'importance d'y participer. D'autres lettres en inuktitut et en anglais ont été rédigées et imprimées afin d'être remises en main propre dans les foyers du Nord qui ne pouvaient être joints autrement. Des messages d'intérêt public ont été préparés en vue d'être diffusés dans les stations de radio locales. Ceux-ci visaient à informer les auditeurs de la tenue de l'enquête dans leur communauté et l'on y mentionnait qu'un intervieweur de Statistique Canada pourrait se présenter chez eux. Ces messages ont été traduits dans les quatre principales langues inuites dans chacune des quatre régions inuites.

Une fois la collecte terminée, des lettres de remerciements ainsi qu'un calendrier de l'EAPA ont été envoyés aux répondants qui ont participé à l'enquête.

4.6 Problèmes particuliers sur le terrain

Les intervieweurs ont reçu comme consigne de faire toute tentative raisonnable pour obtenir une interview complète avec le membre du ménage sélectionné. On a communiqué de nouveau à deux reprises au plus 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 téléphoné à un mauvais moment, il 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.

Des problèmes particuliers sont survenus durant la collecte des données de l'EAPA. Ceux-ci ont été résolus grâce à une coordination additionnelle sur le terrain et à des ajustements correctifs aux méthodes d'enquête. Par exemple, la collecte des données de l'EAPA de 2012 s'est déroulée en même temps que celle de plusieurs autres enquêtes de Statistique Canada, ce qui a pu créer un fardeau de réponse important pour les personnes sélectionnées devant participer à plus d'une enquête. On a procédé à une planification soigneuse et à des ajustements dans le cadre de la conception de l'enquête pour résoudre et réduire au minimum ce type de fardeau de réponse (pour de plus amples renseignements, veuillez consulter la section 3.2.4(e)).

4.7 Taux de réponse finaux

Environ 38 150 personnes ont participé à l'EAPA de 2012, incluant celles qui ont déclaré être non-Autochtones, ce qui représente un taux de réponse final de 76 %. En excluant quelque 9 740 répondants non autochtones, le nombre total de répondants autochtones compris dans la base de données de l'EAPA de 2012 était d'environ 28 410.

Une description détaillée du plan d'échantillonnage utilisé et des taux de réponse obtenus figurent au chapitre 3 du présent guide. Le tableau 4 de la section 3.3 présente les taux de réponse finaux et les tailles de l'échantillon pour chacun des domaines géographiques couverts par l'enquête (par exemple les provinces et les territoires).

5. Traitement des données

5.1 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'utilisation d'un questionnaire automatisé réduit les délais et les coûts de traitement associés à la saisie des données, aux erreurs de transcription et à la transmission des données.

Une partie du contrôle a été fait directement au moment de l'interview. De façon plus particulière, lorsqu'une réponse semblait incohérente par rapport aux réponses précédentes ou se situait à l'extérieur des valeurs attendues, on demandait à l'intervieweur, grâce à des messages à l'écran, de confirmer les réponses auprès du répondant et, au besoin, de modifier l'information.

5.2 Étapes de traitement des enquêtes sociales

Le traitement des données comportent une série d'étapes pour convertir les réponses au questionnaire électronique de leur format brut à une base de données conviviale de grande qualité comprenant un ensemble exhaustif de variables pour l'analyse. Plusieurs opérations sont exécutées pour supprimer les erreurs accidentelles dans les fichiers, vérifier rigoureusement les données pour en assurer la cohérence, coder les questions ouvertes, créer des variables utiles pour l'analyse des données et, enfin, systématiser et documenter les variables pour faciliter leur utilisation à des fins analytiques.

Dans le cadre de l'Enquête auprès des peuples autochtones (EAPA) de 2012, on a utilisé un nouvel ensemble d'outils de traitement des enquêtes sociales élaboré à Statistique Canada et appelé « Environnement pour le traitement des enquêtes sociales » (ETES). L'ETES fait intervenir des programmes SAS (Statistical Analysis System/système d'analyse statistique), des applications personnalisées et des processus manuels pour l'exécution des étapes systématiques suivantes :

Étapes de traitement :

- Réception des données brutes
- Épuration
- Nouveau codage
- Enchaînements
- Codage
- Contrôle et imputation
- Variables dérivées
- Création d'un fichier de traitement final
- Création de fichiers de diffusion

5.3 Réception des données brutes et épuration des dossiers

À partir des données brutes recueillies dans les applications du questionnaire électronique de l'EAPA de 2012, un certain nombre de procédures d'épuration préliminaires ont été mises en œuvre pour chacun des dossiers individuels. Cela comprenait la suppression de tous les identificateurs personnels des fichiers, comme les noms et les adresses, dans le cadre d'un ensemble rigoureux de mécanismes permanents visant à assurer la protection de la confidentialité des répondants. Les enregistrements en double ont été réduits à cette étape. Par ailleurs, dans le cadre des procédures d'épuration, on a passé en revue tous les dossiers des répondants, afin de s'assurer que chacun faisait partie du champ de l'enquête et avait un questionnaire suffisamment rempli. (À noter que les critères pour déterminer si un répondant fait partie ou non du champ de l'enquête ont été appliqués avant tout contrôle ou imputation). Les critères particuliers

pour déterminer qui est un répondant final à l'EAPA et qui n'est pas un répondant final sont fournis ci-après.

5.3.1 Définition d'un répondant

1. Pour être dans le champ de l'enquête, les répondants devaient être âgés de 6 ans et plus au 1er février 2012 et respecter au moins un des critères d'identité autochtone (voir la section 2.2 pour les critères complets).
2. Pour que leur questionnaire soit « complet », les répondants âgés de 6 à 14 ans devaient avoir fourni des réponses valides (c.-à-d. pas de « Ne sais pas » ou de « Refus ») à des questions clés particulières dans les domaines de l'éducation ou de la santé.
3. Pour que leur questionnaire soit « complet », les répondants de 15 ans et plus devaient avoir fourni des réponses valides (c.-à-d. pas de « Ne sais pas » ou de « Refus ») à des questions clés particulières relatives à la scolarité ou au travail **et** à la santé.

Les personnes qui ne respectaient pas ces critères ont été retirées de la base de données. Conformément aux règles qui précèdent, tous les répondants « partiels » qui faisaient partie du champ de l'enquête selon la partie 1 de la définition, mais qui ne respectaient pas les exigences de contenu de la partie 2 ou de la partie 3 de la définition, ont fait partie de ceux retirés de la base de données finale. Veuillez vous reporter à la section 6.4 du présent document pour plus de renseignements sur les répondants partiels.

5.4 Nouveau codage des variables et questions à réponses multiples

Cette étape du traitement a nécessité des changements dans les variables individuelles. Des variables pouvaient être supprimées, codées à nouveau, ajustées ou laissées telles quelles. Les modifications du formatage visaient à faciliter le traitement, ainsi que l'analyse des données par les utilisateurs. Parmi ces modifications figurait la conversion des questions à réponses multiples (questions « Choisissez toutes les réponses appropriées ») à des ensembles correspondants de variables à réponse unique, qui sont plus faciles à utiliser. Pour chaque catégorie de réponse associée à la question originale, une variable a été créée avec des valeurs de réponse de oui/non. Un exemple est fourni ci-après.

Question originale à réponses multiples :

ED4_Q11A - Pour quelles raisons n'avez-vous pas terminé vos études postsecondaires?

INTERVIEWER: Choisissez toutes les réponses appropriées

01	Grossesse/S'occuper de son (ses) enfant(s)
02	Autres responsabilités familiales
03	Maladie/Incapacité
04	Raisons financières (pas assez d'argent)
05	Perte d'intérêt/Manque de motivation
06	A obtenu un emploi/Voulait travailler
07	Trop âgé(e) ou trop tard maintenant
08	Cours trop difficiles/Mauvais résultats
09	Trop difficile d'être loin de la maison
10	Préjugés et racisme
11	Déménagement
12	Autre – Précisez
NSP, RF	

Variables finales en format de réponse unique/oui/non :

ED4_Q11AA - Pour quelles raisons n'avez-vous pas terminé vos études postsecondaires?

- Grossesse/S'occuper de son (ses) enfant(s)

1 Oui
2 Non
NSP, RF

ED4_Q11AB - Pour quelles raisons n'avez-vous pas terminé vos études postsecondaires?

- Autres responsabilités familiales

1 Oui
2 Non
NSP, RF

ED4_Q11AC - Pour quelles raisons n'avez-vous pas terminé vos études postsecondaires?

- Maladie/Incapacité

1 Oui
2 Non
NSP, RF

... questions additionnelles Oui-Non pour chaque catégorie de réponse, comme indiqué, de « Raisons financières (pas assez d'argent) » à « Déménagement »... et en incluant la dernière catégorie :

ED4_Q11A_L - Pour quelles raisons n'avez-vous pas terminé vos études postsecondaires?

- Autre - Précisez

1 Oui
2 Non
NSP, RF

5.5 Enchaînements : cheminements de réponse, sauts valides et non-réponse à des questions

Un autre ensemble de procédures de traitement des données pour l'EAPA de 2012 comprenait la vérification des enchaînements des questions. Tous les cheminements de réponse et enchaînements de questions intégrés au questionnaire ont été vérifiés, afin de s'assurer que l'univers ou la population cible pour chaque question avaient été saisis correctement lors du traitement. Une attention spéciale a été accordée aux distinctions entre les sauts valides et la non-réponse, une distinction importante pour l'analyse statistique. Ces concepts sont expliqués ci-après, afin d'aider les utilisateurs à mieux comprendre les univers des questions, ainsi que les statistiques issues des variables de l'EAPA.

Réponse – réponse directement pertinente au contenu de la question qui peut être catégorisée en catégories de réponses préétablies, y compris « Autre – Précisez ».

Saut valide – indique que la question a été sautée parce qu'elle ne s'appliquait pas à la situation du répondant, compte tenu des réponses valides aux questions précédentes. Dans de tels cas, le répondant n'est pas considéré comme faisant partie de la population cible ou de l'univers pour cette question. Comme il est noté ci-après, lorsqu'une question a été sautée en raison d'un cheminement indéterminé (c.-à-d. un « Ne sais pas » ou un « Refus » à une question précédente, qui a entraîné le saut), un code « Non déclaré » est attribué au répondant pour cette question.

Ne sais pas – le répondant n'a pas été en mesure de fournir une réponse pour une ou plusieurs raisons (parce qu'il ne se rappelait pas ou parce qu'il répondait pour quelqu'un d'autre, par exemple).

Refus – le répondant a refusé de répondre, peut-être en raison de la nature délicate de la question.

Non déclaré – ceci indique que la réponse à la question est manquante et que le cheminement du répondant est indéterminé, par exemple, lorsqu'un répondant n'a pas répondu à la question filtre précédente ou lorsqu'une incohérence a été trouvée dans une série de réponses.

Des codes spéciaux ont été désignés pour chacun de ces types de réponses, afin de faciliter la reconnaissance et l'analyse des données par l'utilisateur. Par exemple, le dernier chiffre des codes « saut valide » est « 6 », et les chiffres précédents « 9 » (p. ex. le code serait « 996 » pour une variable à trois chiffres). Toutes les réponses « Ne sais pas » se terminent par un « 7 », précédé par des « 9 » (p. ex. « 997 »). Les refus se terminent par un « 8 », précédé par des « 9 » (p. ex. « 998 »); et les valeurs « Non déclaré » se terminent par un 9, précédé aussi par des « 9 » (p. ex. « 999 »).

5.6 Codage

5.6.1 Questions « Autre – Précisez »

Le traitement des données comprend aussi le codage des réponses aux questions « Autre – Précisez », aussi appelées « réponses en toutes lettres ». Pour la plupart des questions du questionnaire de l'EAPA, des catégories de réponses codées au préalable ont été fournies, et les intervieweurs ont été formés pour attribuer les réponses des répondants à la catégorie appropriée. Toutefois, lorsque la réponse d'un répondant ne pouvait être facilement attribuée à une catégorie existante, de nombreuses questions ont aussi permis à l'intervieweur d'entrer une réponse en toutes lettres dans la catégorie « Autre – Précisez ».

Toutes les questions comportant des catégories « Autre – Précisez » ont fait l'objet d'un examen étroit pendant le traitement. Par suite d'un examen qualitatif des types de réponses en toutes lettres fournies, des lignes directrices de codage ont été élaborées pour chaque question. À partir de ces lignes directrices de codage, nombre des réponses en toutes lettres fournies ont été codées à nouveau dans l'une des catégories existantes énumérées au préalable. Les réponses qui étaient uniques et différentes des catégories existantes ont été conservées comme « Autre ». Pour certaines questions, une ou plusieurs nouvelles catégories ont été créées lorsqu'un nombre suffisant de réponses le justifiait. Dans le cas des questions où les réponses « Autre – Précisez » constituaient moins de 5 % des réponses globales à la question, on n'a pas procédé au codage et les réponses sont demeurées dans la catégorie « Autre ».

Environ 58 000 réponses pour 78 questions ont été consignées dans la catégorie « Autre – Précisez » et passées en revue pour le codage. L'annexe B résume les autres catégories ajoutées pour l'EAPA de 2012. Celles-ci seront prises en compte lorsque l'on précisera davantage les catégories de réponses pour les cycles futurs de l'enquête.

5.6.2 Questions ouvertes et classifications types

Quelques questions du questionnaire de l'EAPA de 2012 ont été consignées par les intervieweurs dans un format complètement ouvert. Cela comprenait des questions liées à la profession du répondant et à l'industrie dans laquelle il travaillait, ainsi qu'à son principal domaine d'études postsecondaires, le cas échéant. Ces réponses ont été codées à partir d'une combinaison de procédures de codage automatisées et interactives. Des systèmes de classification uniformisés ont servi à coder ces réponses. L'annexe C comporte des détails au sujet de ces classifications.

Une classification uniformisée a aussi servi à coder les langues autochtones que les répondants parlaient ou comprenaient, ainsi que la première langue apprise dans l'enfance. Dans le cas des langues, on a fourni aux intervieweurs un menu déroulant exhaustif de langues pour effectuer un choix, mais des réponses en toutes lettres ont aussi été saisies au besoin. Dans l'ensemble, 51 catégories de langues autochtones ont servi à coder les données sur la langue de l'EAPA. Pour plus de détails sur le système de classification utilisé pour les langues autochtones, voir l'annexe C.

Le codage de toutes les classifications a fait intervenir des procédures de codage et de contrôle de la qualité éprouvées, ainsi que d'autres procédures de vérification du traitement.

5.7 Contrôle et imputation

Après l'étape du codage dans le traitement, une série de contrôles personnalisés ont été effectués à l'égard des données. Ils comprenaient des vérifications de la validité des variables elles-mêmes et entre elles, afin de déterminer les lacunes, les incohérences, les valeurs aberrantes extrêmes et d'autres problèmes liés aux données. Pour résoudre les données problématiques déterminées dans le cadre des contrôles, des corrections ont été apportées selon des règles de contrôle logique. Dans certains cas, les données correspondantes ont été tirées des réponses des répondants de l'Enquête nationale auprès des ménages (ENM). On appelle cela l'imputation.

Un exemple de contrôle de validité pour une question est la variable du logement liée au nombre de pièces dans le logement, qui permettait à un intervieweur d'inscrire jusqu'à 95 pièces. Pour supprimer les réponses aberrantes dont on croyait qu'elles étaient invalides, un contrôle a été intégré pour faire en sorte que le nombre de pièces déclarées dans le logement ne dépasse pas 20. Par ailleurs, de nombreux contrôles de la cohérence entre les questions ont été exécutés en rapport avec les variables de l'éducation, afin d'éviter les contradictions dans les profils d'éducation. Par exemple, une personne qui n'avait pas déclaré avoir fréquenté un établissement d'enseignement postsecondaire particulier, comme une université, une école de métiers, un collège, un cégep ou un autre établissement non universitaire, puis avait déclaré par la suite tenter d'obtenir un certificat, un diplôme ou un grade d'un de ces établissements, a été considérée comme ayant fréquenté ce type d'établissement. La réponse à la question précédente a été modifiée de « non » à « oui » pour le type particulier d'établissement pour lequel le contrôle était requis.

Dans le cadre de l'EAPA de 2012, une série d'imputations importantes ont été effectuées en rapport avec les classifications de l'identité autochtone. Par exemple, les personnes pour lesquelles des données étaient manquantes pour les questions ID_Q02 sur le groupe d'identité autochtone, ID_Q03 sur le statut d'Indien inscrit, ou ID_Q05 sur l'appartenance à une Première Nation ou à une bande indienne ont reçu des valeurs imputées sur la base de leurs réponses à l'ENM. Dans le cas de celles qui se sont déclarées comme Autochtones à la question ID_Q01 de l'EAPA, mais qui n'ont pas déclaré de groupe autochtone particulier à la question ID_Q02, une imputation a aussi été effectuée sur la base de la réponse à l'ENM. En outre, une imputation a été effectuée pour une personne qui n'avait pas été identifiée comme appartenant à un groupe autochtone, mais qui s'était identifiée comme 1) un Indien avec statut, 2) un Indien inscrit en vertu de la loi C-31 ou de la loi C-3, ou 3) un membre d'une Première Nation ou d'une bande indienne – ces répondants ont été imputés à Premières Nations (Indiens de l'Amérique du Nord).

Enfin, même si toutes ces vérifications ont été effectuées systématiquement à l'égard de tous les sujets au moyen de contrôles programmés à l'ordinateur, certains cas de combinaisons très complexes d'information ont été passés en revue et corrigés manuellement.

5.8 Variables dérivées et couplage de l'ENM

Afin de faciliter une analyse plus approfondie du riche ensemble de données de l'EAPA, plus de 500 variables dérivées ont été créées en combinant des questions du questionnaire. Des variables dérivées (VD) ont été créées pour les principaux domaines de contenu. En outre, plus de 100 variables de l'ENM ont été couplées au fichier analytique final de l'EAPA pour 2012.

Nombre des variables dérivées étaient simples et comprenaient uniquement la combinaison de questions équivalentes, par exemple, celles concernant les cheminements en matière d'éducation. D'autres variables dérivées simples comprenaient le regroupement de catégories en catégories plus larges. Dans d'autres cas, deux variables ou plus ont été combinées pour créer une variable nouvelle ou plus complexe, utile pour les analystes des données. Certaines des variables dérivées étaient fondées sur des variables couplées de l'ENM, y compris les niveaux géographiques multiples de l'ENM et les régions inuites. L'ascendance autochtone a aussi été tirée de l'ENM, étant donné qu'elle n'est pas mesurée directement dans l'EAPA de 2012.

Au moment de l'élaboration des variables dérivées, une catégorie de réponse valide n'a généralement pas été attribuée à un répondant pour une variable dérivée donnée si l'une ou l'autre des parties de la question n'avait pas reçu de réponse (c'est-à-dire si l'une ou l'autre des questions utilisées pour la variable dérivée avait été codée « Ne sais pas », « Refus » ou « Non déclaré »). Dans ces cas, le code attribué à la variable dérivée a été étiqueté « Non déclaré ».

La plupart des noms des variables dérivées comportent un « D » comme premier caractère. Les VD géographiques font exception, étant donné qu'elles rendent compte des noms de variables correspondants de l'ENM. Pour toutes les variables couplées de l'ENM, le nom de la variable de l'ENM a été préservé dans la plus large mesure possible dans la base de données de l'EAPA. Certaines exceptions se sont appliquées, étant donné que les noms de variables de l'EAPA se limitent à huit caractères, tandis que les noms de variables de l'ENM dépassent parfois huit caractères.

Le dictionnaire de données de l'EAPA de 2012 indique de façon détaillée les variables dérivées et fournit des renseignements sur la façon dont les dérivations ont été effectuées. Les grandes lignes des VD figurent par thème dans l'annexe A, avec d'autres indicateurs de l'enquête. Une liste complète des variables couplées de l'ENM et de leurs notes connexes figure dans le dictionnaire de données de l'EAPA de 2012, qui accompagne le fichier analytique de l'EAPA.

5.9 Création de fichiers de données finaux et d'un dictionnaire de données

Quatre fichiers de données finaux ont été créés au moment du traitement des données :

- Fichier de traitement final
- Fichier analytique utilisé dans les centres de données de recherche
- Fichier de microdonnées à grande diffusion (FMGD)
- Fichiers de partage avec les Inuit, conformément à l'entente de partage de données avec les quatre régions inuites.

Le fichier de traitement final est un fichier interne qui comprend un certain nombre de variables temporaires utilisées exclusivement pour le traitement. Le fichier analytique, le FMGD et le fichier de partage avec les Inuit sont des fichiers de diffusion qui font l'objet d'un traitement additionnel aux fins de la diffusion. Les fichiers de diffusion doivent être distribués à divers moments suivant le jour de diffusion des données de l'EAPA, soit le 25 novembre 2013 (voir le chapitre 9 pour obtenir des descriptions plus détaillées et des renseignements sur la diffusion).

Le fichier analytique est distribué dans les centres de données de recherche au Canada, mais est accessible uniquement pour les chercheurs qui répondent à certaines exigences. Le fichier analytique est aussi utilisé à Statistique Canada pour produire des tableaux de données par suite de demandes des clients. Le FMGD est élaboré en vue d'une distribution publique plus large. Les fichiers de partage avec les Inuit sont produits en conformité avec les ententes de partage des données conclues avec les régions inuites : Nunatsiavut, Nunavik, Nunavut et région Inuvialuit. Pour tous ces fichiers de diffusion, de nombreuses mesures ont été prises pour assurer la confidentialité des répondants.

Afin de transformer le fichier de traitement final épuré en fichier analytique les chercheurs, un certain nombre d'étapes ont été suivies. Tout d'abord, une série de mesures ont été appliquées pour améliorer la protection de la confidentialité des répondants. En deuxième lieu, des poids ont été ajoutés au fichier pour les personnes. La pondération est décrite de façon plus détaillée dans la section 6. Enfin, toutes les variables temporaires ou variables utilisées exclusivement pour le traitement ont été supprimées du fichier de traitement final.

Le fichier analytique de l'EAPA de 2012 s'accompagne du cliché d'enregistrement, de la syntaxe SAS et SPSS (Statistical Package for the Social Sciences/ensemble des programmes statistiques relatif aux sciences sociales) servant à télécharger le fichier, ainsi que de métadonnées sous forme de dictionnaire de données, qui décrit chaque variable et fournit des fréquences pondérées et non pondérées.

Le FMGD fait l'objet d'un traitement plus exhaustif pour la protection de la confidentialité des répondants. Afin d'assurer la non-divulgation des données confidentielles, le niveau de détail du FMGD n'est pas aussi précis que celui des fichiers analytiques conservés par Statistique Canada. Des mesures sont prises pour empêcher que l'on reconnaisse des répondants comportant des combinaisons de caractéristiques potentiellement identifiables. Ces mesures de protection englobent la limite des régions géographiques comprises dans le fichier, des ajustements aux poids d'enquête, l'examen des chevauchements avec d'autres FMGD publiés, l'exclusion de variables, le regroupement de catégories pour certaines variables, le plafonnement de certaines valeurs numériques extrêmes, ainsi que la détermination des dossiers uniques à risque et des occurrences rares.

6. Pondération

Pour le calcul des poids initiaux, les « faux Autochtones » ont été traités comme s'ils n'avaient jamais existés, c'est-à-dire qu'ils ont été exclus de la base de sondage et de l'échantillon. Voir la sous-section 3.2.3 sur le plan de sondage et la répartition de l'échantillon.

Dans une enquête basée sur un échantillon, chaque personne sélectionnée est représentative non seulement d'elle-même, mais aussi des autres personnes qui n'ont pas été échantillonnées. Par conséquent, un poids est lié à chaque personne sélectionnée, afin d'indiquer le nombre de personnes que cette dernière représente. Ce poids doit être utilisé pour toutes les estimations. Par exemple, dans un échantillon aléatoire simple de 2 % de la population, chaque personne en représente 50. Le poids initial est par la suite ajusté afin de tenir compte d'éléments comme la non-réponse et les écarts entre les caractéristiques de l'échantillon et les totaux connus pour la population cible (poststratification ou stratification a posteriori). On a eu recours à sept étapes dans le processus de pondération.

6.1 Poids initiaux

Le poids initial d'une unité donnée dans une strate de l'Enquête auprès des peuples autochtones (EAPA) correspond au produit de deux composantes : l'inverse de la fraction de sondage dans la strate et le poids de l'Enquête nationale auprès des ménages (ENM) corrigé pour la non-réponse à l'ENM pour l'unité en question. La fraction de sondage dans la strate est calculée comme étant le nombre de personnes sélectionnées à l'EAPA dans chaque strate, divisé par le nombre total de répondants à l'ENM disponibles pour cette strate. Le poids de l'ENM utilisé est le poids de sondage de l'ENM corrigé pour la non-réponse, puis borné au 99e centile, tel que calculé par l'équipe de méthodologie travaillant sur l'estimation à l'ENM.

6.2 Ajustement pour les unités non envoyées à la collecte

Un nombre relativement faible d'unités échantillonnées n'ont pas été envoyées à la collecte, pour différentes raisons. Ces unités comprenaient entre autres:

- les cas où trois membres d'un même ménage avaient déjà été sélectionnés;
- les unités sans nom ni date de naissance;
- les « unités non admissibles de la vague 2 », c'est-à-dire les personnes sélectionnées à la vague 2 dans des ménages où au moins une personne avait refusé de participer à l'enquête à la vague 1.

Dans les deux premiers cas, un ajustement par ratio a été fait par région et par groupe autochtone de l'ENM. Dans le troisième cas, un ajustement par ratio a été fait par région, par groupe autochtone et par groupe de scolarité de l'ENM. À l'intérieur d'une région et d'un groupe autochtone (ou d'une région, d'un groupe autochtone et d'un groupe de scolarité dans le troisième cas), les poids des unités retirées ont été remis à zéro et les poids des unités restantes ont été accrus proportionnellement (ajustement par ratio).

6.3 Ajustement pour tenir compte de la non-réponse

Deux ajustements ont été effectués pour deux types différents de non-réponse : les personnes sélectionnées pour lesquelles aucun contact n'avait été établi avec la personne elle-même ou son parent ou tuteur (« non-contact » : 2 981 adultes et 770 enfants), et les personnes contactées qui n'ont pas fourni (ou qui n'ont pas pu fournir) de renseignements sur eux-mêmes ou leur enfant (« non-réponse avec contact » : 6 263 adultes et 1 763 enfants). Cette seconde forme de non-réponse est principalement associée aux refus ou aux « refus déguisés ». Un exemple de « refus déguisé » pourrait être une personne contactée à plusieurs reprises qui repousse continuellement l'entrevue à plus tard. Deux ajustements ont été effectués puisque les caractéristiques des gens qu'on ne réussit pas à contacter sont souvent différentes des caractéristiques des gens qui refusent de répondre, une fois contactés.

La distinction entre enfants et adultes est faite ici selon l'âge à l'ENM (et non l'âge au 1er février 2012, tel que mesuré à l'EAPA), soit les moins de 15 ans pour les enfants et les 15 ans et plus pour les adultes. Il s'agit d'une distinction importante puisque beaucoup moins de caractéristiques pouvant expliquer la non-réponse sont disponibles chez les enfants que chez les adultes à l'ENM. Par ailleurs, chez les enfants, ce ne sont pas les caractéristiques de l'enfant lui-même qui influent sur la réponse ou la non-réponse, mais plutôt les caractéristiques de la personne répondant pour l'enfant (parent ou tuteur). Il a donc fallu déterminer pour chaque enfant de moins de 15 ans à l'ENM qui était la personne la plus susceptible de répondre pour l'enfant selon la situation de l'enfant dans la famille de recensement, et ce, peu importe qu'une réponse ou non ait été obtenue à l'EAPA pour cet enfant. Dans les situations où les parents ou tuteurs de l'enfant vivaient en couple de sexes opposés, la préférence a été donnée à la personne de sexe féminin (mère, grand-mère, tante, par exemple).

Il convient de mentionner que la définition de non-contact a changé par rapport à ce qu'elle était lors de l'EAPA de 2006. Étant donné que les interviews en 2012 ont été menées par interview assistée par ordinateur (IAO) plutôt que par mode d'entrevue « papier-crayon » comme en 2006, un ensemble de variables de collecte appelées « paradonnées » sont disponibles pour toutes les unités de l'échantillon. En particulier, des informations sont recueillies pour chaque tentative de contact. Ainsi, une unité sera déclarée comme étant un « non-contact », si aucune des tentatives n'a permis d'entrer en contact avec la personne sélectionnée ou le parent ou tuteur de l'enfant sélectionné. En 2006, le non-contact était établi en fonction de la dernière tentative de contact seulement. En conséquence, en 2012, on a eu proportionnellement moins de « non-contact » et davantage de « non-réponse avec contact », comparativement à 2006.

Les poids ont d'abord été ajustés pour tenir compte des cas de « non-contact », puis des cas de « non-réponse avec contact », pour les adultes et les enfants séparément. Dans ce qui suit, le terme « non-réponse » sera utilisé pour les deux types de non-réponse. Le terme « répondant » désigne la personne qui fournit l'information sur la personne sélectionnée (habituellement elle-même dans le cas d'un adulte, ou le parent ou tuteur pour les enfants).

Chaque ajustement pour tenir compte de la non-réponse a été effectué en trois étapes. Dans la première étape, un modèle de régression logistique a été utilisé pour prédire la probabilité de répondre (probabilité d'obtenir une réponse) pour chaque unité (à la fois pour les unités répondantes et non répondantes), à partir d'une série de variables explicatives. Ces variables explicatives sont scindées en deux groupes. Dans le premier groupe, on retrouve les caractéristiques « personnes » ou « ménages » de l'ENM pour la personne sélectionnée ou du parent ou tuteur de l'enfant (p. ex. groupe autochtone de la personne sélectionnée ou du parent ou tuteur de l'enfant, nombre de personnes dans le ménage de la personne sélectionnée, etc.). Dans le second groupe de variables explicatives, on retrouve les variables de collecte appelées « paradonnées ». Le nombre de tentatives de contacts avec un sujet, le fait que des tentatives de dépistages aient été requises ou non sont des exemples de variables de paradonnées retenues par les modèles de régression logistique. Les paradonnées ont été considérées comme des variables particulièrement utiles pour prédire la réponse ou la non-réponse. En effet, nombre de ces variables mesurent les efforts déployés pour contacter une personne ou pour obtenir une réponse d'une personne contactée. Par exemple, les personnes pour lesquelles un nombre important de tentatives de contacts ont été nécessaires ont été considérées comme très similaires aux personnes non contactées (toutes les tentatives ont échouées).

Dans la deuxième étape, les répondants et non-répondants ayant des probabilités prédites de réponse similaires ont été regroupés en classes d'ajustement en se servant de l'analyse de classification. Une simulation a été effectuée pour déterminer de façon approximative le nombre optimal de classes ainsi que le nombre de répondants minimum par classe. Le taux de réponse a été calculé pour chacune des classes en fonction du nombre de répondants et de non-répondants dans la classe. Le taux de réponse calculé était pondéré par les poids obtenus à l'étape d'ajustement précédente.

Dans la troisième étape, l'inverse du taux de réponse pondéré dans une classe a servi de facteur d'ajustement pour cette classe, et les poids des unités répondantes à l'intérieur de la classe ont été ajustés en conséquence. Les poids des unités non-répondantes ont été mis à 0.

Il est important de noter qu'à cette étape, toutes les unités considérées comme étant hors du champ de l'enquête ont été classées comme étant des répondants. En effet, toutes les informations nécessaires ont été obtenues de ces personnes pour déterminer qu'elles étaient hors du champ de l'enquête. Les poids de ces unités ne faisant pas partie du champ de l'enquête ont ont été mis à 0 à la dernière étape de l'ajustement des poids et ces unités ont été retirées du fichier analytique. Le fait de les conserver jusqu'à la dernière étape permettra de produire à l'interne des estimations pondérées de différents groupes d'unités hors de la population cible. Ceci sera, entre autres, très utile pour l'estimation de certains paramètres lors d'une prochaine enquête.

6.4 Ajustement pour les répondants partiels

Les répondants partiels correspondent aux personnes ayant l'identité autochtone dans l'EAPA mais qui n'ont pas fourni suffisamment de renseignements pour correspondre à la définition de répondant tel que décrite au chapitre 5. Les répondants partiels sont au nombre de 157, ce qui devrait avoir peu d'impact sur les estimations.

L'ajustement a été effectué selon la région, le groupe autochtone et le groupe de scolarité mesurés dans l'ENM. Certains regroupements ont été effectués à partir du croisement de ces variables de façon à obtenir un nombre suffisant d'observations pour calculer le facteur d'ajustement. Sachant que ces répondants partiels ont déclaré l'identité autochtone, seuls les poids des répondants d'identité autochtone ont été augmentés pour tenir compte des répondants partiels (les poids des hors-champs, y compris les non-Autochtones à l'EAPA n'ont pas été modifiés). Les poids des répondants partiels ont alors été mis à 0.

6.5 Poststratification

La poststratification (aussi appelée ajustement par stratification a posteriori) fait en sorte que la somme des poids ajustés des unités répondantes correspond aux estimations de l'ENM, selon différents groupes appelés poststrates.

Dans le cas de l'EAPA, deux poststratifications distinctes ont été effectuées. La première poststratification visait à ajuster les poids à la population d'identité ou d'ascendance autochtone de l'ENM par poststrate en utilisant les variables d'identité et d'ascendance qui ont été prises de la BDR (voir la sous-section 3.1.3) lors de la sélection de l'échantillon (et non les variables mesurées à l'EAPA qui font plutôt l'objet de la 2e poststratification). Les poststrates sont définies à partir de certaines combinaisons de la région, du type d'Autochtone (identité autochtone ou ascendance autochtone seulement), du groupe autochtone (Première Nation avec statut, Première Nation sans statut, Métis, Inuit, autre) et du groupe d'âge (6 à 14 ans, 15 à 44 ans et 45 ans et plus). La distinction entre Première Nation avec et sans statut a été faite uniquement pour les provinces entre l'Ontario et la Colombie Britannique. Il est important de mentionner que les estimations de l'ENM sur lesquelles les poids sont ajustés correspondent exactement à la couverture de l'EAPA. Il s'agit donc de la population d'identité ou d'ascendance autochtone âgée de 6 ans et plus au 1er février 2012, excluant celle vivant dans les réserves et certaines communautés de Premières Nations dans les territoires.

Les poids ont été ajustés selon le ratio de l'estimation pondérée de l'ENM et de l'estimation pondérée de l'échantillon pour chaque poststrate. Ceci fait en sorte que l'échantillon ne sous-représentait ou ne surreprésentait pas certaines combinaisons de groupes autochtones, régions et groupes d'âge de l'ENM.

Étant donné que les réponses aux questions définissant la population d'identité autochtone (présentées à la sous-section 3.1.1) peuvent différer entre l'EAPA et l'ENM, une seconde poststratification a été effectuée. Il est à noter que les questions de l'EAPA définissant la population d'identité autochtone sont légèrement différentes de celles posées dans l'ENM (voir le tableau 1 à la section 2 et la sous-section 3.1.1). La seconde poststratification faisait en sorte que la population d'identité autochtone estimée à partir des questions de l'EAPA, correspondait à celle définie dans l'ENM à l'intérieur de chaque poststrate. À noter que, contrairement à la première poststratification, la deuxième n'était pas une poststratification « classique » où les poids étaient réajustés pour tenir compte de la sous-représentation ou surreprésentation de certains groupes dans l'échantillon. En effet, les réponses aux questions sur l'identité autochtone à l'EAPA pouvaient être différentes de celles obtenues par l'ENM pour diverses raisons (section 8.1). Il s'agissait plutôt d'une poststratification « pratique » qui faisait en sorte que les comptes de population d'identité autochtone selon l'EAPA étaient les mêmes que ceux obtenus à l'ENM. Après cette étape, seuls les répondants ayant l'identité autochtone selon l'EAPA avaient des poids positifs.

Il est important de noter que le système de traitement et d'imputation de l'EAPA de 2012 faisait en sorte qu'une catégorie d'identité autochtone n'existait plus. Il s'agissait du groupe « Indien avec statut ou membre d'une Première Nation/bande indienne seulement » (voir la sous-section 3.2.1). Les cas appartenant à ce groupe ont été imputés comme étant des Premières Nations dans l'enquête de 2012. Lors de la seconde poststratification, les personnes de ce groupe à l'ENM ont aussi été combinées avec les Premières Nations. Comme il était impossible de préserver les comptes d'identité multiples entre l'EAPA et l'ENM (comptes trop petits et écarts trop grands), les personnes déclarant une identité de Première Nation et Métis, Première Nation et Inuit ou Première Nation, Métis et Inuit ont été combinées à celles déclarant une identité de Première Nation lors de la seconde poststratification. Les personnes déclarant une identité de Métis et d'Inuit ont été combinées aux Métis. Les secondes poststrates sont formées de certaines combinaisons spécifiques de la région, du groupe d'identité autochtone (Première Nation avec statut, Première Nation sans statut, Métis, Inuit) ainsi que du groupe d'âge (6 à 14 ans, 15 à 44 ans et 45 ans et plus).

6.6 Ajustement des poids extrêmes – méthode de l'écart-sigma

Une fois les ajustements de poids précédents effectués, certains poids pouvaient avoir de très grandes valeurs comparés aux autres. Ceci pouvait créer des problèmes lors de l'estimation si, en plus, les observations ayant de grands poids avaient des caractéristiques très distinctes des observations avec de plus petits poids. Une méthode appelée « écart-sigma » a été utilisée pour détecter ces poids extrêmes à l'intérieur de chaque poststrate, les poststrates étant étroitement reliées aux domaines d'estimation de l'enquête (voir la sous-section 3.2.1). Un exemple d'application de la méthode de l'écart-sigma est décrit dans Bernier et Nobrega (1998)[3]. La méthode de l'écart-sigma, telle qu'utilisée ici, visait à détecter les « valeurs aberrantes » (poids excessivement grands) en calculant la différence entre deux poids successifs, après que ceux-ci aient été triés en ordre décroissant. On a comparé cette différence à *n*l'écart type* des poids à l'intérieur de chaque poststrate. Si la différence dépassait *n*l'écart type* des poids, le poids le plus grand était déterminé comme aberrant. Lorsqu'un poids était déterminé aberrant, tous les autres qui lui étaient supérieurs, dans sa poststrate, étaient automatiquement considérés comme aberrants. Ces poids étaient alors tronqués à la valeur du premier poids non aberrant. Ensuite, la masse des poids tronqués était redistribuée à l'intérieur des poststrates par un ajustement ratio. Après avoir étudié plusieurs scénarios, une valeur de 2 a finalement été choisie pour *n*. Cette valeur particulière de *n* permettait de trouver les poids que nous aurions intuitivement considérés comme étant aberrants.

3. Bernier, J. et Nobrega, K. 1998. Outlier detection in asymmetric samples: A comparison of an inter-quartile range method and a variation of a sigma-gap method. Congrès annuel de la société Statistique du Canada, juin 1998..

7. Qualité des données

7.1 Aperçu de l'évaluation de la qualité des données

L'objectif de l'Enquête auprès des peuples autochtones (EAPA) de 2012 est de produire des estimations de qualité dans les domaines de l'éducation, de l'emploi, de la santé et d'autres indicateurs de base pour les membres des Premières Nations vivant à l'extérieur des réserves, les Métis et les Inuit âgés de 6 ans et plus.

Les sections 7.2 et 7.3 ci-après expliquent les deux types d'erreurs qui se produisent dans les enquêtes : erreurs d'échantillonnage et erreurs non dues à l'échantillonnage. Chaque type d'erreur est évalué dans le contexte de l'EAPA de 2012. L'erreur d'échantillonnage est la différence entre les données obtenues à partir de l'échantillon de l'enquête et les données qui auraient découlé d'un recensement complet de l'ensemble de la population mené dans des conditions similaires. Ainsi, l'erreur d'échantillonnage peut être décrite comme les différences découlant de la variabilité d'un échantillon à l'autre. Les erreurs non dues à l'échantillonnage sont toutes les autres erreurs qui ne sont pas liées à l'échantillonnage. Les erreurs non dues à l'échantillonnage peuvent se produire à n'importe quelle étape du processus d'enquête et comprennent la non-réponse à l'enquête, ainsi que les erreurs qui se sont produites pendant la collecte des données ou le traitement informatique. Il se peut que des erreurs figurent dans les réponses des répondants qui tentent de se rappeler des faits du passé, par exemple, ou se produisent lorsque la réponse est obtenue par personne interposée. Il se peut qu'une réponse soit saisie de façon incorrecte en raison de la fatigue de l'intervieweur ou d'un problème informatique. Une erreur peut avoir été faite au moment de la programmation, lorsque les données ont été traitées ou totalisées. Ce sont tous là des exemples d'erreurs non dues à l'échantillonnage.

Le présent chapitre décrit les diverses mesures adoptées pour prévenir les erreurs dans la mesure du possible et pour corriger les erreurs qui se produisent aux différentes étapes de l'EAPA. Les domaines où il convient de faire preuve de prudence pour l'interprétation des données de l'EAPA sont aussi notés.

7.2 Erreurs d'échantillonnage et méthode bootstrap

Les estimations qui peuvent être calculées à partir de l'Enquête auprès des peuples autochtones (EAPA) de 2012 sont fondées sur un échantillon de personnes. Des estimations quelque peu différentes auraient pu être obtenues si l'on avait procédé à un recensement complet au moyen des mêmes questionnaires, intervieweurs, superviseurs, méthodes de traitement, etc. que ceux réellement utilisés. La différence entre une estimation obtenue à partir de l'échantillon et celle découlant d'un dénombrement complet effectué dans des conditions similaires est appelée « erreur d'échantillonnage » de l'estimation.

Afin de fournir des estimations de l'erreur d'échantillonnage pour les statistiques produites dans le cadre de l'EAPA, on a eu recours à une méthode bootstrap particulière (le bootstrap étant elle-même une méthode de ré-échantillonnage sparticulière). Plusieurs méthodes bootstrap existent dans la littérature, mais aucune d'elles n'était appropriée pour le plan d'échantillonnage de l'EAPA. Les particularités du plan d'échantillonnage qui rendaient difficile l'estimation des erreurs d'échantillonnage étaient les suivantes :

- Plan d'échantillonnage à trois phases dans lequel des ménages (ou logements) étaient sélectionnés aux deux premières phases, et des personnes l'étaient à la troisième (sous-section 3.2.3).
- Fraction de sondage de l'échantillon de première phase (échantillon de l'Enquête nationale auprès des ménages, ENM) non négligeable (environ 1/3 dans les régions N1) et fraction de sondage de l'EAPA généralement assez élevée dans la plupart des strates.
- Strates de l'EAPA (combinaisons de domaines d'estimation, régions N1 ou N2, répondant initial vs répondant au suivi des cas de non-réponse (SCNR), identité vs ascendance seulement) non imbriquées dans les strates de l'ENM (unités de collecte ou groupes d'unités de collecte).
- Méthode utilisée devait être suffisamment souple pour produire des statistiques standards, comme des proportions, des totaux, des moyennes et des ratios, mais aussi des statistiques plus poussées, y compris des centiles, des coefficients de régression logistique, etc.

Certaines de ces particularités avaient été rencontrées lors de l'EAPA de 2006. Cependant, en 2006, la base de sondage était construite à partir du formulaire long du Recensement. La fraction de sondage était alors d'environ un ménage sur cinq et le taux de réponse était élevé en raison du caractère obligatoire du Recensement. En 2011, l'ENM avait une fraction de sondage encore plus grande d'environ un ménage sur trois. L'ENM étant une enquête volontaire, la non-réponse était passablement plus élevée qu'au questionnaire long de 2006, et un sous-échantillon de non-répondants a été sélectionné pour un suivi de la non-réponse, ce qui a complexifié le plan d'échantillonnage. Une description plus détaillée du plan d'échantillonnage de l'ENM se trouve au chapitre 3 du *Guide de l'utilisateur de l'Enquête nationale auprès des ménages*.

Comme la non-réponse à l'ENM était relativement considérable, pour **les besoins du calcul de la variance seulement** (la variance étant une mesure particulière de l'erreur d'échantillonnage), les répondants au SCNR ont été considérés comme formant un échantillon de 3e phase, où la probabilité d'inclusion d'un ménage était égale à une probabilité de répondre qui lui était propre et qui ne dépendait pas des probabilités de répondre des autres ménages.

Plusieurs méthodes bootstrap existent dans la littérature pour l'échantillonnage à une phase ainsi que l'échantillonnage à plusieurs degrés. La plus courante est appelée *bootstrap avec remise* (en anglais, « with-replacement bootstrap ») et consiste à sélectionner *M* sous-échantillons avec remise à partir de l'échantillon principal et à produire des estimations pour chaque sous-échantillon. L'estimation de la variance bootstrap est par la suite calculée comme fonction du carré des différences entre les estimations provenant de chacun des *M* sous-échantillons bootstrap et l'estimation provenant de l'échantillon de l'enquête.

L'utilisation de poids bootstrap simplifie grandement le calcul de la variance. Pour chaque sous-échantillon (répétition bootstrap), on doit ajuster les poids d'échantillonnage initiaux en fonction du sous-échantillonnage bootstrap, ce qui produit ce que l'on appelle les « poids bootstrap initiaux ». Étant donné que chaque échantillon bootstrap est le résultat de la sélection d'unités avec remise, une unité peut se retrouver plusieurs fois dans un échantillon bootstrap particulier. On peut démontrer que les poids bootstrap sont une fonction du poids d'échantillonnage initial de l'observation multiplié par ce que l'on appelle « la multiplicité » de l'unité dans l'échantillon bootstrap, c'est-à-dire le nombre de fois où l'unité est sélectionnée dans l'échantillon bootstrap. La multiplicité d'une unité dans l'échantillon bootstrap est une variable aléatoire qui suit ce que l'on appelle une « distribution multinomiale ». Ainsi, les poids bootstrap peuvent être considérés comme le produit des poids d'échantillonnage initiaux par un facteur d'ajustement aléatoire (qui est, dans ce cas-ci, une fonction de la multiplicité de l'unité). Une fois les poids bootstrap initiaux calculés, tous les ajustements de poids appliqués aux poids d'échantillonnage initiaux sont appliqués aux poids bootstrap initiaux en vue d'obtenir les poids bootstrap finaux, qui rendront compte de la variance liée non seulement au plan d'échantillonnage particulier, mais aussi à tous les ajustements de poids appliqués à l'ensemble de l'échantillon pour calculer les poids finaux.

Pour l'EAPA de 2006, une méthode bootstrap généralisée pour plan d'échantillonnage à deux phases[4] a été élaborée.

Tel que mentionné précédemment, les poids bootstrap peuvent être considérés comme le produit des poids d'échantillonnage initiaux et d'un facteur d'ajustement aléatoire. Il s'agit du concept qui sous-tend la méthode bootstrap généralisée. Dans le cas d'un échantillon à deux phases, la variance peut être décomposée en deux composantes, chacune étant associée à une phase de l'échantillonnage. La méthode bootstrap généralisée à deux phases produit un facteur d'ajustement aléatoire pour chaque phase d'échantillonnage. Dans ce cas, le poids bootstrap initial d'une unité est le produit du poids d'échantillonnage initial par ces deux facteurs d'ajustement aléatoires.

Or en 2011, **pour l'estimation de la variance seulement**, on considère que l'ENM avait deux phases supplémentaires, soit celle correspondant au sous-échantillon du SCNR ainsi que celle associée à la non-réponse au sous-échantillon du SCNR. Dans le cadre de l'EAPA de 2012, nous avons combiné les trois phases de l'ENM en une seule pour ensuite utiliser la méthode bootstrap généralisée à deux phases (une phase à l'ENM et une phase à l'EAPA). Dans la méthode bootstrap généralisée à deux phases, les facteurs d'ajustement aléatoires sont fonctions des probabilités d'inclusion simples et doubles associées à chacune des phases. Afin de regrouper les trois phases en une seule, on a combiné les probabilités d'inclusion

4. Langlet, É., Beaumont, J.-F., et Lavallée, P. 2008. Bootstrap Methods for Two-Phase Sampling Applicable to Postcensal Surveys. Article présenté au Comité consultatif des méthodes statistiques de Statistique Canada, mai 2008, Ottawa.

simples et doubles pour les trois phases de l'ENM. Les probabilités d'inclusion simples et doubles des trois phases combinées sont données par le produit des probabilités d'inclusion simples et doubles de chacune des trois phases. Les détails de la méthodologie utilisée se trouvent dans Haddou (2013)[5].

Une fois les trois phases de l'ENM combinées en une seule, on a utilisé la méthode bootstrap généralisée pour plan d'échantillonnage à deux phases, qui impliquait le calcul de deux ensembles de facteurs d'ajustement aléatoires pour chacune des phases.

La présence de ces deux ensembles de facteurs d'ajustement aléatoires comportait un avantage majeur. Le premier ensemble pouvait être utilisé pour les estimations fondées sur la première phase seulement, c'est-à-dire, les estimations fondées sur l'échantillon de l'ENM. Ces estimations ont été utilisées lorsque les poids ont été ajustés en fonction des totaux de l'ENM au moment de la poststratification (section 6.5). Cette méthode a produit des totaux de l'ENM variables pour chaque échantillon bootstrap. Cela rendait compte du fait que les totaux de l'ENM utilisés étaient fondés sur un échantillon et non sur des totaux fixes connus.

Dans le cas de l'EAPA, 1 000 ensembles de poids bootstrap ont été produits au moyen de la méthode décrite précédemment. La méthode utilisée est légèrement biaisée à la hausse, en ce sens qu'elle surestime légèrement la variance. Cependant, l'ampleur de la surestimation est considérée comme négligeable pour l'EAPA. La méthode peut aussi produire des poids bootstrap négatifs. Pour remédier à ce problème, on a procédé à une transformation des poids bootstrap en vue de réduire leur variabilité. Par conséquent, la variance calculée à partir de ces poids bootstrap transformés doit être multipliée par un facteur qui est fonction d'un certain paramètre appelé phi. La valeur de ce paramètre correspond au plus petit nombre entier faisant en sorte que tous les poids bootstrap soient positifs. Pour l'EAPA, ce facteur est de 4. Les variances calculées à partir des poids bootstrap transformés doivent donc être multipliées par $4^2 = 16$. De même, les coefficients de variation obtenus (racine carrée de la variance divisée par l'estimation proprement dite) doivent être multipliés par 4. Toutefois, la plupart des logiciels qui produisent des estimations de l'erreur d'échantillonnage à partir de poids bootstrap comportent une option permettant de préciser ce facteur d'ajustement, afin que la bonne estimation de la variance soit obtenue sans la nécessité d'une étape additionnelle de multiplication par la constante.

Il est extrêmement important d'utiliser le facteur multiplicatif approprié pour n'importe quelle mesure de l'erreur d'échantillonnage, telle la variance, l'erreur type et le coefficient de variation (c.v.). L'omission de ce facteur multiplicatif produira des résultats et des conclusions erronés. Ce facteur est souvent spécifié comme étant « l'ajustement de Fay » dans les logiciels qui produisent des estimations de l'erreur d'échantillonnage à partir de poids bootstrap.

À noter que si C est le facteur multiplicatif de la variance, certains logiciels utilisent plutôt le paramètre k tel que $k = 1 - 1/\sqrt{C}$. Dans notre cas, puisque C=16, alors k=0,75. Pour des exemples de procédures utilisant l'ajustement de Fay, voir « Enquête auprès de peuples autochtones de 2012 : Guide d'utilisation du fichier analytique».

La mesure de l'erreur d'échantillonnage utilisée pour l'EAPA est le coefficient de variation (c.v.) de l'estimation, c'est-à-dire l'erreur type de l'estimation divisée par l'estimation proprement dite. Dans le cas de cette enquête, lorsque le c.v. d'une estimation est inférieur ou égal à 16,6 %, l'estimation peut être utilisée sans restriction. Si le c.v. est supérieur à 16,6 %, mais inférieur ou égal à 33,3 %, l'estimation est accompagnée de la lettre « E », ce qui indique que les données doivent être utilisées avec prudence. Lorsque le c.v. d'une estimation est supérieur à 33,3 %, l'estimation de la cellule est remplacée par la lettre « F », ce qui indique que l'estimation a été supprimée pour des raisons de fiabilité. La lettre « X » est utilisée pour indiquer qu'une estimation a été supprimée parce qu'elle était confidentielle en vertu des dispositions de la *Loi sur la statistique*. Le tableau 5 résume ces règles.

5. Haddou, M. 2013. Bootstrap Variance Estimation Specifications - Aboriginal Peoples Survey. Document interne, janvier 2013.

Tableau 5 Lignes directrices relatives à la variabilité d'échantillonnage

Type d'estimation	c.v. (en %)	Lignes directrices	Symbole
Acceptable	c.v. ≤ 16,6	On peut envisager une diffusion générale non restreinte des estimations. Aucune annotation particulière n'est nécessaire.	
Marginale	16,6 < c.v. ≤33,3	On peut envisager une diffusion générale non restreinte des estimations, en y joignant une mise en garde aux utilisateurs quant à la variabilité d'échantillonnage élevée liée aux estimations. Les estimations de ce genre doivent être identifiées par la lettre E (ou d'une autre manière similaire).	E – utiliser avec prudence
Inacceptable	c.v. > 33,3	Statistique Canada recommande de ne pas publier des estimations dont la qualité est inacceptable. Toutefois, si l'utilisateur choisit de le faire, il doit alors adjoindre la lettre F (ou un autre identificateur semblable) et les diffuser avec l'avertissement suivant : « Nous avisons l'utilisateur que ... (précisez les données)... ne répondent pas aux normes de qualité de Statistique Canada pour ce programme statistique. Les conclusions tirées de ces données ne sauraient être fiables et seront fort probablement erronées. Ces données et toute conclusion qu'on pourrait en tirer ne devraient pas être publiées. Si l'utilisateur choisit de les publier, il est alors tenu de publier également le présent avertissement. »	F – pas assez fiable pour être publié

7.3 Erreurs non dues à l'échantillonnage

Outre l'échantillonnage, un certain nombre de facteurs à presque toutes les étapes d'une enquête peuvent entraîner des erreurs dans les résultats. Les erreurs non dues à l'échantillonnage sont attribuables principalement aux causes suivantes : la non-réponse, la couverture, la mesure et le traitement. Pour chacun de ces domaines, les sous-sections qui suivent examinent les diverses mesures utilisées pour réduire et corriger les erreurs. Par exemple, les erreurs de mesure peuvent être dues à une mauvaise compréhension par les répondants des questions et aux réponses inexactes; par ailleurs, il se peut que les réponses soient entrées de façon incorrecte pendant la saisie des données et que des erreurs se produisent au moment du traitement et de la totalisation des données. Le passage du questionnaire papier-crayon, en 2006, à l'interview assistée par ordinateur (IAO) en 2012 a réduit considérablement le nombre d'erreurs non dues à l'échantillonnage, parce que l'IAO permet la saisie directe des réponses, des enchaînements automatisés entre les questions, ainsi que des contrôles intégrés qui éliminent les incohérences et les valeurs aberrantes, etc. (pour plus de renseignements sur l'IAO, voir la section 2.1).

En présence d'un grand nombre d'observations, les erreurs aléatoires auront peu d'effet sur les estimations de l'enquête. Toutefois, les erreurs qui se produisent systématiquement contribueront au biais dans les estimations de l'enquête. Ainsi, beaucoup de temps et d'énergie a été consacré à la réduction du nombre d'erreurs non dues à l'échantillonnage dans l'enquête, comme décrit aux sous-sections suivantes.

7.3.1 Erreurs de non-réponse

Les erreurs de non-réponse surviennent lorsqu'il est impossible d'obtenir des données complètes pour toutes les unités de l'échantillon sélectionné. La non-réponse produit des erreurs dans les estimations d'enquête de deux façons. Tout d'abord, les non-répondants ont souvent des caractéristiques différentes des répondants, ce qui peut donner lieu à des estimations d'enquête biaisées si la non-réponse n'est pas corrigée de façon appropriée. Plus le taux de non-réponse est élevé, plus le risque de biais possible le sera. En deuxième lieu, le fait d'avoir un plus grand nombre de non-répondants réduit la taille réelle de l'échantillon. Par conséquent, la précision des estimations diminue (l'erreur d'échantillonnage des estimations augmente). Ce deuxième aspect peut être évité grâce à la sélection d'un échantillon plus important au départ. Toutefois, cela ne réduira pas le biais possible dans les estimations.

Il existe de nombreux types de non-réponse. L'une des formes de non-réponse est la non-réponse partielle (parfois appelée « non-réponse d'item »), c'est-à-dire un répondant qui ne répond pas à une ou plusieurs questions, mais qui répond à la majeure partie du questionnaire. La non-réponse partielle peut être attribuable à la difficulté à comprendre une question particulière.

En général, l'étendue de la non-réponse partielle a été relativement faible dans l'EAPA. Des examens qualitatifs exhaustifs et une mise à l'essai des questionnaires ont eu lieu avant l'enquête, ce qui a réduit l'étendue de la non-réponse partielle. Une réponse à des questions clés déterminées au préalable était requise pour classer un cas comme « répondant », conformément à la description de la sous-section 5.3.1. On retrouvait un nombre important de cas, toutefois, dans lesquels il manquait une proportion importante

des réponses à des questions clés. Ces cas ont été éliminés de la base de données des répondants (ne répondaient pas à la définition de répondant) et ont été traités pendant la pondération comme des cas spéciaux de non-réponse totale (section 6.4). Enfin, il y a non-réponse totale lorsque la personne sélectionnée pour participer à l'enquête n'a pu être contactée ou n'a pas participé une fois qu'elle a été contactée. Les poids des répondants ont été augmentés, afin de compenser pour ceux qui n'ont pas répondu, conformément à la description de la section 6.3.

Afin d'atténuer le nombre de cas de non-réponse, de nombreuses initiatives ont été entreprises. Au cours des mois ayant mené à l'enquête, une stratégie de communication exhaustive décrite au chapitre 4 a été mise en œuvre pour encourager la participation. En outre, les intervieweurs ont reçu une formation approfondie, colligée dans des manuels à leur intention et offerte par des formateurs expérimentés qui ont également supervisé les activités sur le terrain. Des efforts pour joindre les non-répondants, sous forme de rappels et de suivis, ont aussi été faits par des intervieweurs principaux pour encourager les répondants à participer à l'enquête. Des intervieweurs employant l'IPAO ont aussi assuré un suivi sur le terrain dans de nombreuses régions particulières.

Un tableau détaillé des taux de réponse finaux obtenus pour l'EAPA de 2012 figure à la section 3.3 du présent guide (tableau 4).

7.3.2 Erreurs de couverture

Tel que mentionné à la section 3.1, la population cible de l'EAPA 2012 incluait la population d'identité autochtone du Canada, âgée de 6 ans et plus au 1er février 2012, vivant dans les logements privés à l'extérieur des établissements indiens ou des réserves, de même qu'à l'extérieur de certaines communautés des Premières Nations au Yukon et aux Territoires du Nord-Ouest. La population échantillonnée ou couverte par l'enquête correspondait aux répondants de l'ENM ayant déclaré soit l'ascendance ou l'identité autochtone (voir la sous-section 3.1.1) avec les mêmes restrictions que celles concernant la population cible en ce qui a trait à l'âge et à la géographie. Pour des données sur les Premières Nations vivant dans les réserves, on conseille aux chercheurs d'utiliser les données de l'ENM de 2011. Sinon, de l'information sur cette population sera disponible par l'entremise de l'Enquête régionale sur l'éducation, l'emploi et la petite enfance chez les Premières Nations menée par le Centre de gouvernance de l'information des Premières Nations (voir la section 2.2).

Des erreurs de couverture se produisent lorsqu'il y a des différences entre la population cible et la population échantillonnée (population couverte par la base de sondage). Le surdénombrement n'est généralement pas problématique, étant donné que les unités de l'échantillon hors du champ de l'enquête sont habituellement déterminées pendant la collecte des données et peuvent être estimées pour l'ensemble de la base de sondage. Il pourrait toutefois y avoir sous-dénombrement. Comme l'échantillon de l'EAPA a été sélectionné à partir des personnes qui avaient participé à l'ENM, les personnes qui n'ont pas participé à l'ENM n'ont pas pu être échantillonnées pour l'EAPA (le taux de réponse non pondéré à l'ENM s'établit à 68,6 %). En conséquence, un biais de non-réponse à l'ENM peut se traduire par un biais de couverture dans l'EAPA (même si, techniquement, cela pourrait être considéré comme un biais de non-réponse pour l'EAPA). Statistique Canada a mené plusieurs études, avant et après la collecte des données de l'ENM, pour évaluer le risque et la portée du biais de couverture possible. Un certain nombre de mesures ont été prises pour atténuer ces effets, nommément des procédures de suivi particulières de la non-réponse, qui ont été utilisées pour réduire le biais possible pour les populations à risque, comme la population autochtone. Des stratégies de pondération particulières ont aussi été utilisées pour réduire le biais. Pour un examen complet de la qualité des données de l'ENM, voir le *Guide de l'utilisateur de l'Enquête nationale auprès des ménages*.

7.3.3 Erreurs de mesure

Des erreurs de mesure se produisent lorsqu'une réponse fournie diffère de la valeur réelle. Ces erreurs peuvent être attribuables aux répondants, à l'intervieweur, au questionnaire, à la méthode de collecte ou au système de conservation des dossiers du répondant. Des efforts exhaustifs ont été déployés afin d'élaborer des questions pour l'EAPA de 2012 qui seraient compréhensibles, pertinentes et adaptées sur le plan culturel.

Par suite de la diffusion des données de l'EAPA de 2006, un examen exhaustif du contenu a été mené pour les questions existantes de l'EAPA. Cet examen a réuni l'expertise d'un groupe diversifié de chercheurs et de spécialistes du domaine provenant de Statistique Canada et de l'extérieur. Une analyse a été effectuée au sujet des questions qui fonctionnaient le mieux et qui étaient les plus efficaces pour produire des indicateurs valides. Ce processus s'est aussi prolongé par une recherche exhaustive de questions pertinentes à partir d'autres questions d'enquête uniformisées à Statistique Canada. Les questions sélectionnées pour être incluses dans le questionnaire de 2012 ont par la suite fait l'objet de plusieurs rondes d'essais qualitatifs au moyen d'interviews individuelles avec les répondants dans huit communautés différentes de diverses régions du Canada, y compris Iqaluit dans le Nord. La mise à l'essai a touché les Premières Nations, les Métis et les Inuit. Un essai qualitatif du questionnaire d'enquête a été mené par le Centre de ressources en conception de questionnaires (CRCQ) de Statistique Canada. Pour réduire l'erreur de mesure, des ajustements ont été apportés au libellé des questions et aux enchaînements, à partir de ces résultats.

De nombreux autres moyens ont aussi été pris pour réduire de façon particulière l'erreur de mesure, y compris le recours à des intervieweurs qualifiés la formation exhaustive des intervieweurs en ce qui a trait aux procédures et au contenu de l'enquête, ainsi que l'observation et le contrôle des intervieweurs, afin de déceler les problèmes de conception du questionnaire ou de mauvaise compréhension des instructions.

7.3.4 Erreurs de traitement

Des erreurs de traitement peuvent se produire à diverses étapes du processus d'enquête, y compris la saisie des données, le codage et le contrôle. Des procédures de contrôle de la qualité ont été appliquées à chaque étape du traitement des données de l'EAPA, afin de réduire ce type d'erreur. Comparativement à l'EAPA de 2006, les erreurs de traitement en 2012 ont été considérablement réduites, du fait que la méthode de collecte papier-crayon de 2006 a été remplacée par l'interview assistée par ordinateur.

Les questionnaires de l'EAPA ont d'abord été passés en revue sur le terrain par le superviseur des intervieweurs. À l'étape du traitement des données, un ensemble détaillé de procédures et de règles de contrôle ont servi à déterminer et à corriger les incohérences entre les réponses fournies. Un ensemble de procédures rigoureuses et systématisées a été élaboré pour évaluer la qualité de chaque variable et pour apporter des corrections à chaque erreur décelée. Un instantané des fichiers de sortie a été pris à chaque étape et une vérification a été effectuée en comparant les fichiers de l'étape en cours et de l'étape précédente. La programmation de toutes les règles de contrôle a fait l'objet d'un essai exhaustif avant d'être appliquée aux données. Parmi les exemples de vérifications du traitement des données figurent l'examen de tous les enchaînements de questions, y compris les séquences très complexes, afin de faire en sorte que les valeurs de sauts de questions soient attribuées avec exactitude et distinguées des différents types de valeurs manquantes; le double codage de contrôle de la qualité des réponses « Autre – Précisez »; la supervision experte du codage des classifications uniformisées; et l'examen de toutes les variables dérivées par rapport aux variables qui les composent, afin d'assurer une programmation appropriée de la logique de calcul, y compris les calculs très complexes. Voir le chapitre sur le traitement des données du présent guide pour plus de détails (chapitre 5).

8. Différences entre l'Enquête auprès des peuples autochtones et d'autres sources de données

En raison de certaines différences entre la méthodologie utilisée pour l'Enquête auprès des peuples autochtones (EAPA) de 2012, pour les cycles précédents de l'EAPA et pour d'autres enquêtes de Statistique Canada, il faut faire preuve de prudence lorsque l'on compare les données de diverses sources. Les sections qui suivent donnent un aperçu détaillé des éléments touchant la comparabilité des données, et fournissent aux utilisateurs de données des renseignements importants sur les facteurs à prendre en compte au moment d'effectuer des analyses à l'aide des données de l'EAPA.

8.1 Différences entre l'Enquête auprès des peuples autochtones et l'Enquête nationale auprès des ménages

L'échantillon de l'EAPA est sélectionné à partir des répondants ayant fourni des réponses spécifiques à l'Enquête nationale auprès des ménages (ENM). Des renseignements plus détaillés sur la façon dont les réponses de l'ENM ont été utilisées pour déterminer la population d'intérêt pour l'EAPA sont fournis au chapitre 3 (conception de l'enquête).

L'ENM et l'EAPA sont deux sources importantes d'information sur les Autochtones, qui se complètent l'une l'autre. L'EAPA repose sur des concepts qui sont abordés dans l'ENM et comporte des questions plus approfondies, en vue de produire des renseignements plus détaillés. Par exemple, l'ENM fournit de l'information sur les titres scolaires obtenus (certificats, diplômes et grades, soit les questions 27 à 30). L'ajout d'information provenant de l'EAPA permet d'en apprendre davantage sur les niveaux de scolarité inférieurs aux études secondaires, l'identité autochtone des professeurs, l'obtention d'aide financière pour les études postsecondaires et les raisons évoquées pour ne pas poursuivre des études.

L'EAPA porte aussi sur des sujets ou des thèmes complets qui ne sont pas couverts par l'ENM. Par exemple, l'EAPA fournit des renseignements détaillés sur l'emploi et la santé des peuples autochtones.

Bien que les deux enquêtes englobent conceptuellement la « population d'identité », contrairement à l'EAPA de 2006, l'EAPA de 2012 ne couvrait pas la « population d'ascendance seulement », décrite au chapitre 3. (Les répondants de l'ENM ayant déclaré l'ascendance autochtone seulement faisaient partie de l'échantillon de l'EAPA parce qu'ils avaient une probabilité non négligeable de déclarer l'identité à l'EAPA; ces répondants n'ont été conservés dans l'ensemble de données de l'EAPA que s'ils avaient effectivement déclaré une identité autochtone à l'EAPA).

Les comptes de population d'identité autochtone provenant de l'EAPA de 2012 pour certaines sous-populations peuvent différer de ceux obtenus à partir de l'ENM, et ce, même si nous limitons l'univers de la population de l'ENM à celui de l'EAPA. La seconde poststratification décrite à la section 6.5 faisait en sorte que le nombre de personnes ayant déclaré avoir une identité autochtone était le même dans l'ENM et l'EAPA, mais cela concernait seulement certaines combinaisons du groupe autochtone, de la région et du groupe d'âge. Les comptes de population d'identité autochtone peuvent différer pour d'autres sous-populations, qui n'ont pas été contrôlées lors de la poststratification. Par ailleurs, pour une même personne, l'identité autochtone déclarée peut, dans certains cas, différer entre l'ENM et l'EAPA. Il existe de nombreuses raisons pour lesquelles l'identité autochtone diffère entre les deux enquêtes.

8.1.1 Méthodes d'interview différentes et effet de la déclaration par personne interposée

Dans la plupart des régions, les données de l'ENM de 2011 ont été recueillies par autodénombrement. Les questionnaires ont été remplis sur Internet ou retournés par la poste. Pour les réserves indiennes et les régions éloignées, y compris toutes les communautés inuites, on a eu recours à des intervieweurs du recensement. Souvent, un membre du ménage a rempli le questionnaire de l'ENM pour tous les membres du ménage. On parle alors de déclaration par personne interposée.

Tel qu'il est décrit au chapitre 4, les données de l'EAPA ont été recueillies au moyen d'une interview assistée par ordinateur (IAO) auprès de la personne sélectionnée dans le cas d'un adulte et auprès du parent ou du tuteur dans le cas d'un enfant (il était également permis d'interviewer directement les jeunes de 15 à 17 ans avec le consentement du parent ou du tuteur). Puisque la personne contactée dans le cadre de l'EAPA n'est pas nécessairement celle qui a rempli le questionnaire du recensement, les réponses à des questions similaires peuvent varier.

8.1.2 Questionnaires différents

Parmi les autres sources d'écart entre l'ENM et l'EAPA, on trouve la question sur l'autodéclaration autochtone. Pour l'EAPA de 2012, cette question a été scindée en deux, comme il est décrit à la sous-section 8.2.3. De plus, contrairement à l'EAPA qui ne posait plus la question sur l'ascendance autochtone en 2012, la question 18 de l'ENM sur l'autodéclaration de l'identité autochtone est précédée de la question 17, qui porte sur les origines ethniques et culturelles. Plusieurs origines autochtones y sont données à titre d'exemple. Cela pourrait avoir une incidence sur les réponses fournies à la question 18.

Le libellé de la première question sur l'autodéclaration autochtone de l'EAPA est légèrement différent de celui de la question 18 de l'ENM. Entre autres, le terme « Indien de l'Amérique du Nord » qui figure entre parenthèses après le terme « Première Nation » n'est pas mentionné dans le libellé de la question correspondante à l'EAPA. Il est cependant précisé dans une note à l'intervieweur que les termes « Première Nation » et « Indien(ne) de l'Amérique du Nord » peuvent être remplacés l'un par l'autre.

Ces changements dans le questionnaire peuvent avoir entraîné des différences quant à la façon de répondre.

8.1.3 Contextes différents

Dans le questionnaire de l'ENM, on pose un nombre limité de questions générales à l'ensemble de la population du Canada, alors que celui de l'EAPA est conçu précisément pour les Autochtones. Par conséquent, compte tenu du contexte plus précis de l'EAPA, il se peut que le concept d'identité autochtone ait été mieux compris. Ainsi, il est possible que des personnes qui avaient déclaré être Autochtones à l'ENM ont pu ne plus se déclarer Autochtones à l'EAPA. À l'inverse, des personnes qui avaient déclaré être d'ascendance autochtone seulement à l'ENM peuvent avoir déclaré par la suite une identité autochtone à l'EAPA.

8.1.4 Effet du temps

Le concept d'identité autochtone n'est pas nécessairement une caractéristique statique dans le temps. Les événements qui affectent les droits de certains groupes autochtones ou le changement de perception de la population face aux Autochtones, par exemple, peuvent influer sur la façon dont les personnes répondent aux questions sur l'identité autochtone. Les personnes qui se considèrent d'ascendance autochtone seulement à un certain moment peuvent se déclarer comme étant Autochtone à un autre moment. Par ailleurs, une même personne peut se considérer comme étant de Première Nation à un certain moment, puis Métis à un autre.

8.1.5 Traitement

Il est très important de préciser que, dans le cadre de l'EAPA de 2012, toutes les personnes qui ont déclaré être un Indien avec statut ou un membre d'une Première Nation ou d'une bande indienne sans se considérer également comme un Autochtone ont été imputées comme étant de Première Nation dans l'EAPA. À ce titre, ce groupe relativement petit dans l'ENM n'existe pas dans le fichier de données finales de l'EAPA. Ainsi, lorsque l'on effectue des comparaisons entre les comptes d'identité autochtone en fonction du groupe autochtone entre l'ENM et l'EAPA de 2012, les répondants de l'ENM qui ont déclaré être un Indien avec statut ou être un membre d'une Première Nation ou d'une bande indienne et qui ne se sont pas déclarer comme un Autochtone devraient être inclus dans les estimations des Premières Nations.

8.1.6 Différences dans les estimations

Les tableaux qui suivent comparent les estimations tirées de l'ENM à celles de l'EAPA pour différents groupes autochtones et régions géographiques. Les quatre régions inuites sont séparées du reste du Canada.

Le tableau 6 compare les estimations de l'ENM et celles de l'EAPA pour la population d'identité autochtone sans double comptabilisation. Il s'agit de la population d'identité autochtone âgée de 6 ans et plus au 1er février 2012, vivant dans les logements privés à l'extérieur des réserves indiennes ou des établissements indiens, de même qu'à l'extérieur de certaines communautés des Premières Nations au Yukon et dans les Territoires du Nord-Ouest; cela correspond à la couverture de l'EAPA. Il convient de noter que pour l'ENM, il existe une catégorie intitulée « Indien avec statut ou membre d'une Première Nation/bande indienne seulement », et que celle-ci n'existe pas pour l'EAPA après imputation (ces cas ont été imputés comme étant Première Nation dans l'EAPA). Dans l'ENM, une colonne a été créée pour le total des catégories « Première Nation » et « Indien avec statut ou membre d'une Première Nation/bande indienne seulement ».

Le tableau 7 est semblable, mais il tient compte de la double comptabilisation. Ainsi, une personne ayant une identité multiple de Première Nation et de Métis sera prise en compte à la fois dans les catégories Première Nation et Métis.

Tous les comptes des tableaux suivants ont été arrondis à la dizaine près. Comme les totaux sont arrondis de façon indépendante des cellules individuelles, la somme des cellules peut ne pas correspondre exactement aux totaux.

Tableau 6 Nombre de personnes âgées de 6 ans et plus[1] vivant hors réserve[2] et déclarant une identité autochtone aux deux enquêtes, sans double comptabilisation

	Premières Nations				Métis	
Région	**ENM PN**	**ENM IB**	**ENM tot**	**EAPA**	**ENM**	**EAPA**
Nunatsiavut	10	0	10	20	20	10
Nunavik	50	0	50	50	40	10
Inuvialuit	790	10	800	730	180	210
Nunavut	110	10	130	100	130	90
Nunangat (total)	**960**	**20**	**980**	**890**	**360**	**320**
Atlantique, excluant le Nunatsiavut	36 650	4 050	40 700	40 760	21 540	21 330
Québec excluant le Nunavik	40 850	4 150	45 010	45 160	37 830	37 840
Ontario	140 250	7 220	147 460	149 030	80 060	79 980
Manitoba	45 510	900	46 420	46 910	71 450	71 470
Saskatchewan	40 680	900	41 580	41 640	46 420	46 360
Alberta	61 130	2 900	64 030	64 410	87 460	87 290
Colombie-Britannique	93 620	3 300	96 920	97 630	63 550	63 610
Yukon	4 050	40	4 090	4 070	710	710
Territoires du Nord-Ouest excluant l'Inuvialuit	3 250	120	3 370	3 340	1 880	1 870
Reste du Canada (total)	**465 990**	**23 580**	**489 570**	**492 960**	**410 900**	**410 460**
Canada (total)	**466 950**	**23 610**	**490 560**	**493 850**	**411 260**	**410 780**

Tableau 6 Nombre de personnes âgées de 6 ans et plus[1] vivant hors réserve[2] et déclarant une identité autochtone aux deux enquêtes, sans double comptabilisation (suite)

	Inuit		Multiple		Identité (total)	
Région	**ENM**	**EAPA**	**ENM**	**EAPA**	**ENM**	**EAPA**
Nunatsiavut	2 090	2 090	0	10	2 120	2 120
Nunavik	9 270	9 270	20	40	9 380	9 380
Inuvialuit	2 920	2 920	10	50	3 900	3 900
Nunavut	23 190	23 190	20	80	23 460	23 460
Nunangat (total)	**37 480**	**37 480**	**40**	**180**	**38 860**	**38 860**
Atlantique, excluant le Nunatsiavut	4 740	5 220	600	740	67 570	68 050
Québec excluant le Nunavik	1 520	1 130	1 360	1 200	85 720	85 340
Ontario	2 910	3 480	2 520	1 050	232 960	233 530
Manitoba	480	590	950	430	119 290	119 410
Saskatchewan	270	180	500	500	88 770	88 680
Alberta	1 780	1 340	1 590	1 370	154 860	154 420
Colombie-Britannique	1 460	1 320	2 100	1 330	164 030	163 890
Yukon	150	130	10	30	4 970	4 950
Territoires du Nord-Ouest excluant l'Inuvialuit	800	710	30	60	6 070	5 990
Reste du Canada (total)	**14 110**	**14 110**	**9 670**	**6 720**	**924 250**	**924 250**
Canada (total)	**51 580**	**51 580**	**9 710**	**6 900**	**963 110**	**963 110**

ENM Enquête nationale auprès des ménages
EAPA Enquête auprès des peuples autochtones
PN Première Nation
IB Indien avec statut ou membre d'une Première Nation/bande indienne seulement
Tot Total pour Première Nation ou membre d'une Première Nation/bande indienne seulement
1. Au 1er février, 2012.
2. Et hors certaines communautés de Premières Nations au Yukon et T.N.-O.

Tableau 7 Nombre de personnes âgées de 6 ans et plus[1] vivant hors réserve[2] et déclarant une identité autochtone aux deux enquêtes, avec double comptabilisation

	Premières Nations				Métis		Inuit	
Région	NHS FN	NHS IB	NHS Tot	APS	NHS	APS	NHS	APS
Nunatsiavut	10	0	10	20	20	10	2 090	2 100
Nunavik	60	0	70	80	40	20	9 290	9 320
Inuvialuit	800	10	800	760	180	220	2 930	2 960
Nunavut	120	10	130	170	140	100	23 210	23 270
Nunangat (total)	**990**	**20**	**1 010**	**1 030**	**380**	**360**	**37 520**	**37 650**
Atlantique, excluant le Nunatsiavut	37 090	4 050	41 140	41 140	22 050	22 030	5 000	5 630
Québec excluant le Nunavik	42 180	4 150	46 330	46 330	39 140	39 020	1 730	1 190
Ontario	142 670	7 220	149 890	149 890	82 440	80 940	3 230	3 750
Manitoba	46 430	900	47 330	47 330	72 390	71 900	520	610
Saskatchewan	41 170	900	42 070	42 070	46 920	46 810	270	290
Alberta	62 590	2 900	65 490	65 490	88 960	88 640	2 010	1 660
Colombie-Britannique	95 650	3 300	98 950	98 950	65 600	64 940	1 600	1 330
Yukon	4 060	40	4 100	4 100	730	740	160	140
Territoires du Nord-Ouest excluant l'Inuvialuit	3 270	120	3 390	3 390	1 900	1 920	810	740
Reste du Canada (total)	**475 110**	**23 580**	**498 700**	**498 700**	**420 120**	**416 940**	**15 330**	**15 330**
Canada (total)	**476 100**	**23 610**	**499 710**	**499 730**	**420 500**	**417 300**	**52 850**	**52 980**

ENM Enquête nationale auprès des ménages
EAPA Enquête auprès des peuples autochtones
PN Première Nation
IB Indien avec statut ou membre d'une Première Nation/bande indienne seulement
Tot Total pour Première Nation ou membre d'une Première Nation/bande indienne seulement
1. Au 1er février, 2012.
2. Et hors certaines communautés de Premières Nations au Yukon et T.N.-O.

8.2 Différences entre l'EAPA de 2006 et de celle de 2012

Plusieurs changements importants sont survenus entre l'EAPA de 2006 et celle de 2012, non seulement sur le plan du contenu mais aussi pour ce qui est de la méthodologie. En raison de ces changements, les estimations de population tirées de l'EAPA de 2006 et de 2012 devraient être comparées avec prudence. En fait, la stratégie de diffusion de l'EAPA n'a jamais été axée sur la production de comptes de la population autochtone; elle visait plutôt à présenter les caractéristiques de cette population pour certains sous-groupes sous la forme de proportions.

8.2.1 Différences sur le plan méthodologique

Sur le plan méthodologique, la différence la plus importante entre l'EAPA de 2012 et l'EAPA de 2006 provient du fait que la population d'ascendance autochtone seulement ne faisait plus partie de la population cible de l'enquête en 2012 comme cela a été le cas en 2006. En fait, l'EAPA de 2012 ne comportait aucune question sur l'ascendance autochtone. Il faut savoir que les personnes ayant déclaré une ascendance autochtone seulement à l'ENM ont été échantillonnées puisqu'elles avaient une probabilité non négligeable de déclarer l'identité à l'EAPA. Comme la population cible de 2012 n'incluait plus la population d'ascendance autochtone seulement, il a été possible lors de la deuxième poststratification d'ajuster les poids de sorte que les comptes d'identité correspondent à ceux de l'ENM pour certaines combinaisons de régions, de groupes autochtones et de groupes d'âge (section 6.5). En 2006, par contre, il n'était pas possible d'ajuster à la fois les comptes d'identité et d'ascendance seulement. La deuxième poststratification avait fait en sorte que la population autochtone totale (identité ou ascendance), estimée à partir des questions de sélection de l'EAPA, correspondait à celle des questions de sélection du recensement. Cette poststratification avait été effectuée selon la région géographique et le groupe d'âge, en fonction du nombre total d'Autochtones au Recensement de 2006, et non pas selon chaque groupe autochtone.

Les deux stratégies très distinctes de poststratification entre 2006 et 2012 font en sorte que la croissance de la population autochtone ne peut être mesurée à partir des données de l'EAPA de 2006 et de 2012. En 2006, en raison du moins grand nombre de transitions d'identité autochtone à ascendance autochtone seulement ou d'identité autochtone à non autochtone (ni identité ni ascendance) que de transitions entre ascendance autochtone et identité autochtone entre le recensement et l'EAPA, les comptes d'identité autochtone tirés de l'EAPA de 2006 étaient plus élevés que ceux provenant du Recensement de 2006, et les comptes d'ascendance autochtone seulement étaient plus faibles dans l'EAPA de 2006 que dans le Recensement de 2006.

Une autre différence méthodologique importante est le fait que l'échantillon de l'EAPA de 2006 a été tiré à partir des répondants du Recensement de 2006, alors que l'échantillon de l'EAPA de 2012 a été tiré à partir des répondants de l'ENM de 2011. Les caractéristiques des répondants de l'ENM peuvent être différentes des caractéristiques des répondants d'un recensement. Le fait que les non-répondants aient des caractéristiques différentes de celles des répondants crée ce qu'on appelle le biais de non-réponse. Bien que lors de l'ENM ce biais ait été réduit par les stratégies de suivi ainsi que par les stratégies d'ajustement de non-réponse à la pondération, il est possible qu'un certain biais de non-réponse persiste. Voir la sous-section 7.3.2 pour obtenir plus de renseignements.

8.2.2 Différences en ce qui a trait au contenu

Pour ce qui est du contenu, il a été souligné à la sous-section 8.1.2 qu'une différence majeure entre les deux cycles les plus récents de l'EAPA était le fait que la question sur l'autodéclaration autochtone a été scindée en deux dans le questionnaire de l'EAPA de 2012; les détails sont présentés à la sous-section suivante. Cela peut évidemment créer des différences dans la façon dont les personnes répondent à l'autodéclaration autochtone. Une autre différence très importante est le fait que, en 2012, la question sur l'autodéclaration autochtone n'était plus précédée, comme en 2006, des trois questions portant sur l'ascendance autochtone (trois questions en une). Ces trois questions sur l'ascendance autochtone posées en 2006 ont probablement fait en sorte que les personnes étaient plus susceptibles de déclarer être un Autochtone à la question suivante et étaient plus susceptibles de déclarer des identités multiples. En effet, la proportion de personnes ayant déclaré une identité autochtone multiple était beaucoup plus élevée en 2006 qu'elle ne l'était en 2012.

8.2.3 Modifications des questions sur l'identité autochtone

Plusieurs modifications ont été apportées aux questions de l'EAPA de 2012 visant à mesurer l'identité autochtone par rapport aux questions correspondantes de l'EAPA de 2006. Certaines de ces modifications rendent compte de celles adoptées pour l'ENM de 2011 par rapport au Recensement de 2006. Les changements ont été apportés afin d'améliorer l'exactitude de la mesure, de tenir compte de la terminologie courante et de rendre compte des modifications législatives récentes.

Groupe autochtone autodéclaré

Pour l'EAPA de 2012, la mesure du groupe autochtone autodéclaré a été scindée en deux, comme l'illustre le tableau ci-après. On visait ainsi à améliorer la mesure en permettant aux répondants de s'identifier comme Autochtones à la question ID_Q01, même s'ils ne s'identifiaient pas à un groupe particulier désigné à la question ID_Q02. Le terme « Première Nation » a été adopté pour la première fois en 2012. En 2006, on utilisait uniquement le terme « Indien de l'Amérique du Nord ». En outre, la question de 2012 comportait des précisions à l'intention des répondants, à savoir que le terme « Première Nation » comprenait à la fois les Indiens avec statut et les Indiens sans statut. Enfin, en 2012, on a employé les termes « Inuk » et « Inuit », alors qu'en 2006, on a utilisé uniquement « Inuk ».

Tableau 8 Groupe autochtone autodéclaré, EAPA de 2012, EAPA de 2006, ENM de 2011 et Recensement de 2006

EAPA de 2012	EAPA de 2006
ID_Q01 — Êtes-vous un(e) Autochtone, c'est-à-dire, Première Nation, Métis(se) ou Inuk (Inuit)? Première Nation comprend les Indiens(nes) avec statut et les Indiens(nes) sans statut. ID_Q02 — Êtes-vous Première Nation, Métis(se) ou Inuk (Inuit)? Première Nation (Indien de l'Amérique du Nord) Métis Inuk (Inuit)	Êtes-vous un(e) Autochtone, c'est-à-dire un(e) Indien(ne) de l'Amérique du Nord, un(e) Métis ou un(e) Inuk?
ENM de 2011	**Recensement 2006**
Cette personne est-elle un Autochtone, c'est-à-dire Première Nation (Indien de l'Amérique du Nord), Métis ou Inuk (Inuit)?	Cette personne est-elle un Autochtone, c'est-à-dire un Indien de l'Amérique du Nord, un Métis ou un Inuit (Esquimau)?

Note : Première Nation (Indien de l'Amérique du Nord) comprend les Indiens avec statut et les Indiens sans statut

Indien avec statut (Indien inscrit ou des traités)

Dans le cadre de l'EAPA de 2012, la principale question visant à mesurer le statut d'Indien (ID_Q03) a été reformulée et le terme « statut » a été utilisé pour la première fois. Les questions subséquentes ont été mises à jour pour rendre compte d'une modification législative apportée en 2011 : une modification à la *Loi sur les Indiens*, appelée projet de loi C-3, *Loi sur l'équité entre les sexes relativement à l'inscription au registre des Indiens*.

Tableau 9 Indien avec statut, EAPA de 2012, EAPA de 2006, ENM de 2011 et Recensement de 2006

EAPA de 2012	EAPA de 2006
ID_Q03 — Êtes-vous un(e) Indien(ne) avec statut, soit un(e) Indien(ne) inscrit(e) ou des traités, aux termes de la *Loi sur les Indiens* du Canada? ID_Q04A — Avez-vous déjà présenté une demande au ministère des Affaires autochtones et Développement du Nord Canada (précédemment nommé AINC) afin d'être inscrit(e) comme Indien(ne) de plein droit en vertu de la Loi C-31 ou de la Loi C-3? ID_Q04B — Êtes-vous inscrit(e) comme Indien(ne) de plein droit en vertu de la Loi C-31 ou de la Loi C-3?	Êtes-vous un(e) Indien(ne) visé(e) par un traité ou un(e) Indien(ne) inscrit(e) aux termes de la Loi sur les Indiens du Canada? Avez-vous déjà posé votre candidature au ministère des Affaires indiennes et du Nord canadien afin d'être inscrit(e) comme Indien(ne) de plein droit en vertu de la Loi C-31? Êtes-vous inscrit(e) comme Indien(ne) de plein droit en vertu de la Loi C-31?
ENM de 2011	**Recensement de 2006**
Cette personne est-elle un Indien avec statut (Indien inscrit ou des traités aux termes de la *Loi sur les Indiens* du Canada)? (Question posée pour la première fois avant la question concernant l'appartenance à une Première Nation ou à une bande indienne)	Cette personne est-elle un Indien des traités ou un Indien inscrit aux termes de la *Loi sur les Indiens* du Canada?

Membre d'une Première Nation ou d'une bande indienne

Dans le cadre de l'EAPA de 2012, le terme « Première Nation » précédait le terme « bande indienne »; en 2006, le terme « bande indienne » apparaissait en premier.

Tableau 10 Membre d'une Première Nation ou d'une bande indienne, EAPA de 2012, EAPA de 2006, ENM de 2011 et Recensement de 2006

EAPA de 2012	EAPA de 2006
ID_Q05 — Êtes-vous membre d'une Première Nation ou d'une bande indienne?	Appartenez-vous à une bande indienne ou à une Première Nation?
ENM de 2011	**Recensement de 2006**
Cette personne est-elle membre d'une Première Nation/bande indienne?	Cette personne appartient-elle à une bande indienne ou à une Première Nation?

9. Diffusion des données

9.1 Aperçu de la diffusion des données de l'Enquête auprès des peuples autochtones de 2012

Les données de l'Enquête auprès des peuples autochtones (EAPA) de 2012 ont été diffusées le 25 novembre 2013. La diffusion comprenait un article analytique et une courte vidéo, tous deux sont offerts gratuitement au public dans le site Web de Statistique Canada. Un ensemble de tableaux régionaux accompagnant l'article analytique est offert sur demande par courriel. D'autres tableaux de données personnalisés sont aussi disponibles moyennant des frais.

Les diffusions ultérieures des données de l'EAPA, qui auront lieu principalement en 2014, comprendront des tableaux de données, d'autres documents d'analyse, un fichier de microdonnées à grande diffusion et un fichier analytique de l'EAPA. Chacun de ces produits est abordé de façon plus détaillée ci-après.

Statistique Canada donne aussi des présentations spéciales de l'EAPA aux organismes autochtones nationaux, aux chercheurs des centres de données de recherche, à d'autres intervenants clés et dans le cadre de diverses conférences. Par ailleurs, les conseillers du Programme de liaison avec les Autochtones de Statistique Canada distribueront des bulletins portant sur l'EAPA et répondront aux besoins en données de l'EAPA dans leur région.

9.2 Produits de données et services

Un ensemble de tableaux de données sur les thèmes de l'éducation et de l'emploi sera offert peu après la première diffusion des données de l'EAPA de 2012, soit le 25 novembre 2013. Ces tableaux viennent compléter l'article analytique (voir la section 9.3 pour obtenir la description de l'article). Ils fournissent des données aux échelons national et provinciaux pour les Premières Nations vivant hors réserve, les Inuit et les Métis, et comparent les finissants et les décrocheurs de l'école secondaire âgés de 18 à 44 ans.

En 2014, un ensemble plus exhaustif d'indicateurs de l'EAPA sera offert dans CANSIM, le Système canadien d'information socioéconomique de Statistique Canada. L'EAPA de 2012 est la première EAPA pour laquelle des tableaux CANSIM seront disponibles. CANSIM permet un accès rapide et facile à une vaste gamme des statistiques les plus récentes disponibles au Canada. Les tableaux CANSIM sont offerts gratuitement en ligne.

Le fichier analytique de l'EAPA de 2012 pourra être consulté dans les centres de données de recherche (CDR) de Statistique Canada partout au pays. Afin d'accéder au fichier, les chercheurs doivent subir un examen d'approbation par le comité de recherche et d'éthique. Leur utilisation des données doit se faire en conformité avec les politiques, lignes directrices et normes de Statistique Canada (par exemple, seules des estimations statistiques agrégées qui sont conformes aux dispositions en matière de confidentialité de la Loi sur la statistique peuvent être diffusées à l'extérieur de Statistique Canada). Pour obtenir plus de renseignements sur le Programme des CDR, voir la Foire aux questions.

Le fichier analytique est accompagné du cliché d'enregistrement, de la syntaxe SAS (Statistical Analysis System/système d'analyse statistique) et SPSS (Statistical Package for the Social Sciences/ensemble des programmes statistiques relatif aux sciences sociales) servant à télécharger le fichier, ainsi que de métadonnées sous forme d'un dictionnaire de données qui décrit chaque variable et fournit des fréquences pondérées et non pondérées. En outre, un guide de l'utilisateur détaillé fournit aux chercheurs les lignes directrices dont ils ont besoin pour effectuer une analyse statistique des données de l'EAPA de 2012.

Un fichier de microdonnées à grande diffusion (FMGD) sera aussi diffusé en 2014, et il permettra une utilisation plus large et plus directe des données par les chercheurs, y compris les étudiants universitaires d'un bout à l'autre du pays. Afin de donner une plus grande assurance en ce qui a trait à la non-divulgation des données confidentielles, le niveau de détail du FMGD n'est pas aussi grand que celui du fichier analytique conservé par Statistique Canada.

Dans le cadre de l'EAPA de 2012, quatre ententes de partage de données sont en place avec chacune des régions inuites au Canada : le Nunatsiavut (côte nord du Labrador), le Nunavik (Nord du Québec), le territoire du Nunavut et la région Inuvialuit des Territoires du Nord-Ouest. Ces ententes sont une excellente méthode pour augmenter l'utilisation des données et rendre ces dernières plus accessibles. Seuls les répondants qui sont des Inuit et qui ont approuvé le partage de leurs données avec leur région respective feront partie du fichier de partage des données. Les fichiers de partage des données seront diffusés en conformité avec les ententes de partage des données en 2014.

L'outil d'accès à distance en temps réel (ADTR) de Statistique Canada constitue un autre moyen d'accéder au fichier de données. Il s'agit d'un service par abonnemment payant offert aux clients rattachés à un établissement universitaire, à un ministère ou à un organisme sans but lucratif. L'ADTR est un outil d'accès à distance en ligne qui permet aux utilisateurs d'exécuter en temps réel les programmes SAS sur des ensembles de microdonnées situés dans un endroit centralisé et sécurisé. Les chercheurs qui utilisent l'ADTR n'ont pas directement accès aux microdonnées et ne peuvent pas consulter le contenu des fichiers. Les utilisateurs se servent plutôt des programmes SAS pour extraire les résultats sous forme de tableaux de fréquences agrégés. L'ADTR complète les modes d'accès existants aux microdonnées confidentielles, y compris les centres de données de recherche. Grâce à un nom d'utilisateur et à un mot de passe, l'ADTR offre un accès jour et nuit aux résultats de l'enquête; il suffit d'avoir à sa disposition un ordinateur doté d'un accès à Internet. Des règles de confidentialité et des lignes directrices en matière fiabilité sont appliquées de façon automatisée à toutes les demandes ce qui élimine la nécessité d'une intervention manuelle et permet d'accéder rapidement aux résultats. On peut joindre un agent de l'ADTR à l'adresse rtra_adtr@statcan.gc.ca; la marche à suivre pour obtenir un compte d'utilisateur est expliquée à l'adresse www.statcan.gc.ca/rtra.

Outre ces produits de données et services, les clients peuvent obtenir des tableaux de données personnalisés auprès des Services à la clientèle de Statistique Canada, par courriel, à sasd-dssea@statcan.gc.ca, ou par téléphone, au 1-800-263-1136. Toutes les demandes personnalisées font l'objet d'un examen de confidentialité et les données agrégées sont arrondies avant d'être diffusées aux clients.

9.3 Produits analytiques

Au cœur de la première diffusion des données de l'EAPA de 2012 se trouve un article analytique intitulé « Expériences au chapitre de l'éducation et de l'emploi des Premières Nations vivant hors réserve, des Inuit et des Métis : certains résultats de l'Enquête auprès des peuples autochtones de 2012 ». Cet article comprend une analyse à la fois approfondie des expériences en matière d'éducation et d'emploi, et distincte pour les membres des Premières Nations vivant hors réserve, les Inuit et les Métis âgés de 18 à 44 ans. Pour chaque groupe, on compare les finissants et les décrocheurs de l'école secondaire pour des indicateurs clés des expériences des élèves, de l'environnement scolaire et de la participation de la famille. On y examine aussi les expériences en matière d'études postsecondaires et d'emploi chez les finissants et les décrocheurs de l'école secondaire. Un sommaire des résultats présentés dans l'article est aussi diffusé en format vidéo.

D'autres documents d'analyse, qui présentent les faits saillants des données de l'EAPA de 2012 sur des sujets spécialisés, seront diffusés en 2014. Ils porteront sur les thèmes suivants :

- la santé des Inuit;
- l'emploi des Métis;
- l'éducation des jeunes des Premières Nations vivant hors réserve.

9.4 Produits de référence

Il est possible d'obtenir des renseignements sur l'EAPA dans le site Web de Statistique Canada. L'organisme offre une Base de métadonnées intégrée (BMDI) en ligne pour toutes les enquêtes qu'il mène, y compris l'Enquête auprès des peuples autochtones de 2012. L'objectif de la BMDI consiste à offrir au public des renseignements qui l'aideront à interpréter les données que Statistique Canada publie. Les renseignements (aussi appelés métadonnées) servent à mieux faire comprendre les concepts de base qui définissent les données, y compris les variables et les classifications, les méthodes et enquêtes statistiques sous-jacentes, et des aspects clés de la qualité des données. Un accès direct au questionnaire de l'EAPA de 2012 est aussi fourni.

Par ailleurs, le présent Guide des concepts et méthodes est offert en ligne pour un examen détaillé du contenu de l'enquête, du plan d'échantillonnage, de la collecte et du traitement des données, de la pondération et de la qualité des données, des différences entre l'EAPA de 2012 et l'Enquête nationale auprès des ménages (ENM) de 2011, ainsi que des produits de diffusion de l'EAPA.

Dans le cas des chercheurs qui utilisent le fichier analytique des centres de données de recherche (CDR) de Statistique Canada, un guide de l'utilisateur des CDR peut être consulté. Il énonce en détail les concepts et les méthodes de l'enquête et fournit la marche à suivre pour utiliser le fichier de données. Le guide de l'utilisateur des CDR décrit la structure du fichier de données en détail, y compris toutes les variables de base, les variables dérivées et les couplages de données avec l'ENM. Un dictionnaire de données détaillé renferme toutes les variables disponibles. Le guide de l'utilisateur des CDR fournit également des lignes directrices détaillées pour la totalisation et l'analyse statistique, la façon d'appliquer les poids nécessaires aux données, des renseignements sur les logiciels disponibles et des lignes directrices pour la diffusion des données, notamment les règles d'arrondissement. Le processus d'estimation de la fiabilité des estimations, tant quantitatif que qualitatif, est abordé plus longuement.

Enfin, un guide distinct, le *Guide de l'utilisateur du FMGD*, a été conçu pour les utilisateurs du FMGD de l'EAPA. Il comprend des questions similaires liées aux variables disponibles, un dictionnaire de données, le processus d'estimation, l'utilisation de poids de même que des lignes directrices pour la totalisation, l'analyse statistique et la diffusion des données.

9.5 Contrôle de la divulgation

La loi interdit à Statistique Canada de divulguer toute information recueillie qui pourrait dévoiler l'identité d'une personne, d'une entreprise ou d'un organisme sans leur permission ou sans en être autorisé par la *Loi sur la statistique*. Diverses règles de confidentialité s'appliquent à toutes les données diffusées ou publiées afin d'empêcher la publication ou la divulgation de toute information jugée confidentielle. Au besoin, des données sont supprimées pour empêcher la divulgation directe ou par recoupement de données reconnaissables.

Annexe A Indicateurs de l'Enquête auprès des peuples autochtones de 2012 selon le module du questionnaire

La liste ci-après comprend des indicateurs du questionnaire, ainsi que des variables dérivées (VD), qui combinent différentes questions du questionnaire.

Il convient de souligner qu'il ne s'agit pas d'une liste complète de toutes les variables disponibles à partir de l'Enquête auprès des peuples autochtones (EAPA) de 2012. Pour obtenir la liste complète des variables, voir le dictionnaire de données de l'EAPA de 2012. Pour de plus amples renseignements sur la façon d'obtenir ce dictionnaire, veuillez communiquer avec les Services à la clientèle de Statistique Canada, par courriel à l'adresse sasd-dssea@statcan.gc.ca, ou en composant le 1-800-263-1136.

Introduction

- Objectif de l'enquête
- Nature volontaire de l'enquête
- Explication du couplage de données avec l'Enquête nationale auprès des ménages (ENM)

Identification autochtone

- Groupe autochtone — Premières Nations, Métis ou Inuit
- Statut — Indien inscrit ou des traités en vertu de la Loi sur les Indiens du Canada
- Demande ou inscription en tant qu'Indien inscrit en vertu du projet de loi C-31 ou C3
- Membre d'une Première Nation ou d'une bande indienne

Les VD de l'identité autochtone comprennent des groupes d'identité unique et multiple et des reprises des définitions historiques pour assurer la comparabilité avec les enquêtes précédentes. L'ascendance autochtone est tirée de l'ENM, parce qu'elle n'est pas mesurée directement au moyen de l'EAPA de 2012.

Composition du ménage et état matrimonial

- Âge, sexe, état matrimonial de la « personne la mieux renseignée » répondant pour l'enfant et son lien avec l'enfant
- Nombre de membres du ménage et nombre de personnes de moins de 18 ans
- Membres de la famille vivant dans le ménage

Les VD de la composition du ménage comprennent les groupes d'âge, le nombre de personnes dans le ménage, le nombre d'enfants, la présence d'un partenaire et le nombre de générations dans le ménage.

Mobilité

- A vécu dans la communauté actuelle toute sa vie
- Raisons du déménagement dans la communauté actuelle
- Nombre d'années depuis le déménagement dans la communauté actuelle
- Déjà vécu dans une réserve ou une communauté inuite
- Déjà vécu à l'extérieur d'une communauté inuite

- Raisons du déménagement de la communauté (réserve, communauté inuite seulement)
- Raisons de demeurer dans la communauté
- Fréquence des déménagements

Les VD de la mobilité comprennent la résidence dans une communauté inuite ou des Premières Nations, le nombre d'années depuis le départ d'une communauté inuite ou des Premières Nations, ou le retour.

Éducation

Situation par rapport aux études

- Inscription dans une école
- Niveau de scolarité actuel
- Plus haut niveau de scolarité (déterminé en fonction des renseignements fournis sur les études postsecondaires)
- Diplôme d'études secondaires (et équivalence d'études secondaires)
- Fréquentation d'une école secondaire (et équivalence d'études secondaires)
- Lieu du programme d'équivalence d'études secondaires

Cheminements scolaires

Éducation 1, 2 et 3 : de la 1re à la 6e année, de la 7e à la 12e année, finissants et décrocheurs âgés de moins de 45 ans

(À noter que la plupart de ces indicateurs, mais pas la totalité, ont été mesurés pour les trois groupes.)

- Fréquentation d'un établissement préscolaire (y compris un programme de développement de la petite enfance)
- Fréquentation d'un établissement préscolaire s'adressant aux enfants autochtones (y compris un programme de développement de la petite enfance)
- Nombre d'écoles fréquentées
- Raison du changement d'école
- École située dans une communauté des Premières Nations (dans une réserve)
- Mobilité pour fréquenter une école
- Exposition aux langues autochtones : enseignement d'une langue autochtone et enseignement de matières dans une langue autochtone (autres que les cours de langue)
- Rendement scolaire
- Soutien et tutorat : besoin, raisons, fréquence, source
- Communication avec le parent ou le tuteur
- Perceptions de l'environnement scolaire
- Participation de la famille à l'école
- Soutien de l'apprentissage à la maison
- Fréquence de la lecture par l'élève
- Accès à Internet
- Absences de l'école, cours manqués et retards à l'école : fréquence et raisons

- Attentes et aspirations des parents concernant les études
- Planification et épargne des parents pour les études
- Antécédents scolaires des membres de la famille : frères et sœurs, mère, père

Éducation 2 et 3 : de la 7e à la 12e année, finissants et décrocheurs âgés de moins de 45 ans — Indicateurs additionnels

- Besoin de soutien social et source du soutien
- Sports et activités physiques : participation, fréquence; activités offertes à l'école, à l'extérieur de l'école ou les deux
- Art, théâtre et musique : participation, fréquence; activités offertes à l'école, à l'extérieur de l'école ou les deux
- Activités d'un groupe ou d'un club : participation, fréquence; activités offertes à l'école, à l'extérieur de l'école ou les deux
- Activités liées à la culture des Premières Nations, des Métis ou des Inuit : participation, fréquence
- Temps passé avec des aînés
- Bénévolat et travail non rémunéré dans la communauté
- Travail dans la communauté (gardien[ne] d'enfants, commis de magasin, tuteur)
- Attitudes et comportements des pairs (au secondaire)
- Décrochage scolaire : fréquence et raisons

Éducation 3 et 4 : Finissants et décrocheurs de l'école secondaire âgés de moins de 45 ans et de 45 ans et plus

(À noter que la plupart, mais pas la totalité, de ces indicateurs sont mesurés pour les deux groupes.)

- Raisons du retour aux études
- Âge au moment de l'obtention du diplôme d'études secondaires
- Âge au moment des dernières études primaires ou secondaires
- Études en vue d'obtenir un certificat d'études postsecondaires
- Type d'établissement d'enseignement actuel
- Certificat(s), diplôme(s) ou grade(s) obtenu(s)
- Principal domaine d'études
- Fréquentation actuelle d'une école (études postsecondaires)
- Obtention d'un diplôme d'études postsecondaires
- Année de l'abandon ou de la fin des études
- Raisons de l'abandon des études postsecondaires
- Études à temps plein ou à temps partiel
- Domaine actuel d'études
- Programme de préparation des élèves
- Déménagement pour poursuivre des études postsecondaires
- Formation à distance
- Soutien personnel
- Prêt étudiant du gouvernement

- Méthode de financement
- Obstacles financiers aux études postsecondaires
- Raisons de ne pas avoir poursuivi des études postsecondaires
- Futurs projets d'études
- Antécédents scolaires des membres de la famille : frères et sœurs, mère, père

Les VD de l'éducation comprennent plus de 300 VD combinant des questions équivalentes pour différents cheminements scolaires et 50 autres combinaisons de questions, comme le plus haut niveau de scolarité, le principal domaine d'études, la participation de la famille aux activités de l'école ainsi que le besoin de soutien et le soutien reçu.

Langues autochtones

- Langue(s) autochtone(s) parlée(s) et comprise(s)
- Langue autochtone principale
- Autoévaluation de la capacité de parler et de comprendre une langue autochtone
- Importance de parler et de comprendre une langue autochtone
- Exposition à une langue autochtone
- Première langue apprise pendant l'enfance

Les VD des langues autochtones comprennent plus de 50 VD qui rendent compte de la complexité de la mesure de la langue unique ou des langues multiples parlées et comprises.

Pensionnat

- Fréquentation d'un pensionnat
- Fréquentation d'un pensionnat par les membres de la famille

Les VD du pensionnat comprennent la fréquentation d'un pensionnat par le répondant et les membres de la famille.

Emploi (répondants âgés de 15 ans et plus)

Activité sur le marché du travail

- Taux de chômage, d'emploi et d'activité
- Raison(s) de l'absence du travail

Situation vis-à-vis de l'activité

- Recherche d'un emploi
- Raisons pour ne pas avoir cherché de travail
- Nombre moyen d'heures de recherche
- Capable de travailler
- Raison de la non-disponibilité pour travailler

Recherche d'emploi

- Méthodes de recherche d'emploi
- Durée du chômage
- Obstacles à la recherche d'emploi

Travailleurs découragés

- Recherche d'emploi
- Raisons pour ne pas avoir cherché de travail
- Capable de travailler
- Raison de la non-disponibilité pour le travail

Autres modules de l'emploi

- Lien à l'emploi précédent
- Emplois multiples
- Catégorie de travailleurs
- Entreprise constituée en société
- Durée d'occupation de l'emploi
- Industrie
- Professions
- Nombre habituel d'heures travaillées

Emploi à temps partiel

- Principale raison du travail à temps partiel

Emploi permanent

- Emploi permanent

Mobilité de la main-d'œuvre

- Mobilité de la main-d'œuvre
- Raisons de ne pas avoir déménagé pour trouver un emploi ailleurs

Les VD de l'emploi comprennent la situation vis-à-vis de l'activité, l'emploi à temps plein ou à temps partiel, l'emploi permanent et non permanent, et les emplois multiples.

Activités traditionnelles

- Fabrication de vêtements ou de chaussures
- Fabrication d'œuvres artistiques ou artisanales (p. ex. sculptures, dessins, bijoux)
- Chasse, pêche ou piégeage
- Cueillette de plantes sauvages (p. ex. baies, riz, foin d'odeur)
- Participation à ces activités traditionnelles et fréquence

- Raison de la participation à des activités traditionnelles
- Obstacles à la participation
- Intérêt à participer

Les VD des activités traditionnelles comprennent la participation à des activités traditionnelles au cours de l'année précédente, la participation contre rémunération et l'intérêt à l'égard des activités traditionnelles.

Revenu

- Sources de revenu et principale source de revenu personnel
- Revenu personnel total
- Revenu d'emploi

Les VD du revenu comprennent la source du revenu personnel en 2011, le total des groupes de revenu personnel et les groupes de revenu d'emploi.

Santé

Santé générale 1 — État de santé

- État de santé autodéclaré

Grossesse et accouchement

- Grossesse
- Nombre de naissances
- Âge au premier accouchement

Taille et poids

- Taille et poids

Santé générale 2 — Consultation de différents professionnels de la santé

- Consultation d'un médecin régulier
- Obstacles à la consultation d'un médecin régulier
- Consultation de différents professionnels de la santé (p. ex. dentiste, psychologue)
- Obstacles à la consultation de différents professionnels de la santé
- Type de soins nécessaires

Problèmes de santé chroniques

- Problèmes de santé chroniques chez les personnes de moins de 12 ans
- Limitations physiques (enfants ayant des problèmes de santé chroniques)
- Problèmes de santé chroniques chez les personnes de 12 ans et plus

Blessures

- Blessures
- Fréquence des blessures
- Type de blessure
- Lieu où la blessure est survenue
- Cause de la blessure
- Blessure(s) causée(s) par une chute et raison

Santé mentale

- État de la santé mentale autoévalué

Détresse

- Échelle de détresse

Suicide

- Idées suicidaires
- Tentatives de suicide

Usage du tabac

- Fréquence de l'usage du tabac
- Âge auquel l'usage du tabac a commencé
- Exposition à la fumée secondaire à la maison

Consommation d'alcool

- Fréquence de la consommation d'alcool
- Fréquence de la consommation de cinq verres d'alcool et plus en une même occasion

Consommation de drogues

- Usage de drogues de la rue
- Consommation de médicaments prescrits

Sécurité alimentaire

- Échelle de sécurité alimentaire

Soutien de la communauté

- Soutien de la communauté

Les VD de la santé comprennent la taille, le poids, l'indice de masse corporelle, l'échelle de détresse, le suicide et le soutien.

Logement

- Logement occupé en propriété ou en location
- Logement nécessitant des réparations
- Surpeuplement, nombre de pièces
- Logement subventionné

Les VD du logement comprennent le nombre de personnes par pièce et la durée d'attente d'un logement subventionné.

Géographie

Les VD de la géographie comprennent les régions géographiques multiples de l'ENM et les régions inuites.

Annexe B Catégories de réponses supplémentaires créées pour les questions « Autre — Précisez »

Le tableau ci-après résume les catégories additionnelles qui ont été ajoutées à certaines questions de l'enquête pendant le codage des données (questions comportant une catégorie « Autre — Précisez »).

Thème	Indicateur	Questions	Catégories additionnelles
Pensionnat	Fréquentation d'un pensionnat par les membres de la famille	RS_S02F	RS_Q02G — Tantes ou oncles RS_Q02H — Cousins
Santé	Obstacles à la consultation d'un médecin régulier	GH2_S02	07 — Médecins qui changent toujours dans la région; aucun médecin permanent disponible
Santé	Type de soins nécessaires	GH2_S08	06 — Soins dentaires
Emploi	Raisons de ne pas déménager dans une autre ville pour améliorer les perspectives d'emploi	LM_S02	15 — Problème de santé ou incapacité 16 — Heureux/à l'aise là
Mobilité	Raisons de demeurer dans cette communauté	MOB_S13	14 — Culture/mode de vie traditionnel
Activités traditionnelles	Raisons de la participation à des activités traditionnelles : fabrication de vêtements et d'œuvres artisanales	TA_S01C TA_S02C	05 — Pour un projet scolaire 06 — Pour des cadeaux, œuvres de bienfaisance ou collectes de fonds
Activités traditionnelles	Obstacles à la participation à la cueillette	TA_S04E TA_S04G	07 — Incapacité physique, santé ou raisons médicales
Éducation	Lieu du programme d'équivalence d'études secondaires	EDS1_S06 EDS2_S05 ED4_S04	5 — Dans un collège ou un institut technique
Éducation	Raisons du retour aux études	ED3F_S39	09 — Pour mes enfants/pour servir de modèle
Éducation	Absences de l'école, cours manqués et retards à l'école Raisons	ED1_S28B ED2F_S37B	16 — Grasse matinée/fatigue 17 — Participer à des sports ou à des activités culturelles
Éducation	Raisons de ne pas recevoir d'aide ou de tutorat	ED2B_S17 ED3B_S16	07 — Élève non intéressé, manque de volonté 08 — Envisage de le faire, demande en cours 09 — Tutorat non demandé ou poursuivi 10 — Abandon des études/problèmes d'assiduité

Annexe C Classifications types utilisées pour l'Enquête auprès des peuples autochtones de 2012

Principal domaine d'études : Classification des programmes d'enseignement

Les répondants de l'Enquête auprès des peuples autochtones (EAPA) de 2012 ont fourni des renseignements sur le principal domaine d'études correspondant au certificat, diplôme ou grade le plus élevé qu'ils ont obtenu. Les réponses ont été codées selon la Classification des programmes d'enseignement (CPE), Canada, 2011.

La CPE a été élaborée au départ en 1980 par le National Center for Educational Statistics (NCES) aux États-Unis. Le NCES a diffusé des mises à jour en 1985, 1990, 2000 et 2010. La CPE Canada 2011 est la deuxième version canadienne de cette classification; la première version est la CPE Canada 2000.

La CPE est conçue pour classer les programmes d'enseignement. La CPE est organisée selon le domaine d'études. À Statistique Canada, un domaine d'études est défini comme « la discipline ou le domaine d'apprentissage ou de formation ».

Industrie et profession : Système de classification des industries de l'Amérique du Nord et Classification nationale des professions

Les répondants de l'EAPA ont fourni des renseignements sur leur profession et sur l'industrie dans laquelle ils ont travaillé pendant la semaine de référence de l'enquête. Les réponses ont été codées au moyen du Système de classification des industries de l'Amérique du Nord (SCIAN) Canada et de la Classification nationale des professions (CNP). Les données ont été codées au niveau de détail à quatre chiffres et selon les versions courantes/nouvelles et historiques de ces systèmes de classification, de la façon suivante :

SCIAN Canada 2012 — Système de classification des industries de l'Amérique du Nord (SCIAN) Canada

SCIAN Canada 2007 — Système de classification des industries de l'Amérique du Nord (SCIAN) Canada

CNP — 2011 : Classification nationale des professions (CNP)

CNP-S — 2006 : Classification nationale des professions pour statistiques (CNP-S)

Le SCIAN est un système de classification des industries qui a été conçu par les organismes statistiques du Canada, du Mexique et des États-Unis. Il a été créé dans le contexte de l'Accord de libre-échange nord-américain. La révision du SCIAN pour 2012 avait pour principal objectif de modifier ou de créer des industries qui reflètent les activités et les technologies nouvelles, émergentes ou en évolution. Le SCIAN repose sur des principes orientés sur l'offre ou la production. Les critères utilisés pour regrouper les établissements en industries dans le SCIAN sont basés sur la similitude des structures des facteurs de production, des compétences de la main-d'œuvre et des processus de production.

La CNP 2011 est la taxonomie acceptée à l'échelle nationale et le cadre organisationnel des professions sur le marché du travail au Canada. Elle représente une mise à jour de la Classification nationale des professions de Ressources humaines et Développement des compétences Canada et de la Classification nationale des professions pour statistiques (CNP-S) de Statistique Canada. Comme par le passé, cette révision est fondée sur une recherche exhaustive sur les professions, une analyse et des consultations menées à l'échelle du pays. Les professions sont déterminées et regroupées principalement selon le genre de travail effectué, qui est déterminé d'après les tâches, les fonctions et les responsabilités du travailleur.

Langues autochtones

Pour le codage des langues autochtones pour lesquelles des données ont été recueillies dans le cadre de l'EAPA de 2012, on a utilisé 51 catégories de langues autochtones. Le codage a été effectué en conformité avec la classification du Recensement de 2011 et de l'Enquête nationale auprès des ménages de 2011. Cela diffère des classifications utilisées en 2006 et en 2001. Par exemple, la famille de langues mitchif comportait une nouvelle catégorie distincte en 2011, et les langues inuites comportent maintenant un niveau de détails plus grand. Pour de plus amples renseignements sur la classification complète et les concordances avec les années précédentes, voir l'annexe D du *Dictionnaire du Recensement de 2011*.

Annexe D Glossaire des termes de l'enquête

B

Bande indienne

Groupe d'**Indiens de l'Amérique du Nor**d dont les terres ont été mises de côté et dont l'argent est détenu par l'État. Chaque bande a son propre conseil de bande, qui est habituellement constitué d'un ou de plusieurs chefs et de plusieurs conseillers. Les membres de la communauté choisissent les chefs et les conseillers par élection, ou parfois selon la coutume traditionnelle. Les membres d'une bande partagent généralement des valeurs, des traditions et des pratiques communes qui sont enracinées dans leur patrimoine ancestral. Aujourd'hui, de nombreuses bandes préfèrent l'appellation **Premières Nations**.

Base de données, ensemble de données

Liste organisée et triée de faits ou de données concernant un ensemble de personnes, de ménages, d'entreprises ou d'autres unités pertinentes. Un ensemble de données de Statistique Canada est généralement produit à partir de données d'enquête ou de données administratives, conservé dans un ordinateur et organisé pour pouvoir être accessible facilement dans le cadre d'une vaste gamme d'applications statistiques.

Base de sondage

Liste, carte ou spécification conceptuelle des unités constituant la population d'enquête, à partir de laquelle des personnes peuvent être sélectionnées. Il s'agit, par exemple, d'un répertoire téléphonique ou du répertoire d'une ville, ou encore d'une liste de membres d'une association ou d'un groupe particulier.

C

Centres de données de recherche (CDR)

Le Programme des centres de données de recherche (CDR) fournit aux chercheurs l'accès, dans un environnement sécuritaire géré par Statistique Canada, aux microdonnées des enquêtes sur la population et les ménages. Le programme des CDR fait partie d'une initiative de Statistique Canada, du Conseil de recherches en sciences humaines (CRSH) et d'un consortium d'universités, en vue de renforcer les capacités dans le domaine de la recherche sociale au Canada et d'appuyer la communauté des chercheurs stratégiques. Le programme est aussi soutenu par la Fondation canadienne pour l'innovation (FCI) et les Instituts canadiens de recherche en santé (ICRS).

Centre de population

Le terme « centre de population » remplace le terme « région urbaine » (tel qu'utilisé au Recensement de la population jusqu'en 2006). Un centre de population contient une concentration démographique d'au moins 1 000 habitants et une densité de population de 400 habitants ou plus au kilomètre carré. Les centres de population sont classés en trois groupes, selon la taille de leur population :

- centres de population de petite taille : de 1 000 à 29 000 habitants
- centres de population de moyenne taille : de 30 000 à 99 999 habitants
- centres de population de grande taille : de 100 000 habitants et plus

Cohorte

Groupe de personnes dans le domaine de la démographie ayant une caractéristique commune, par exemple, toutes les personnes d'une population donnée qui sont nées en 1940, ou toutes les personnes qui ont un problème de santé particulier.

Coefficient de variation (c.v.)

Lors d'une enquête sur échantillon, les résultats de l'échantillon sont utilisés pour estimer ce que les résultats auraient été si l'ensemble de la population avait été mesuré. Dans ce processus d'estimation, un certain niveau d'erreur est inévitable. Le coefficient de variation (c.v.) est une façon d'exprimer l'**erreur d'échantillonnage** liée à une **estimation**. Tout d'abord, on calcule une **erreur type** ou erreur « moyenne » pour l'estimation. Le c.v. est obtenu en divisant l'erreur type de l'estimation par l'estimation proprement dite, la fraction qui en résulte étant exprimée en pourcentage. Plus le c.v. est faible, plus la qualité des données est élevée (voir **Marge d'erreur**).

Confidentialité

Résultat d'un rapport de confiance implicite entre la personne qui fournit les données et la personne ou l'organisation qui les recueille. Ce rapport repose sur l'assurance que les données ne seront pas divulguées sans la permission de la personne. En vertu de la ***Loi sur la statistique***, les données qui permettent d'identifier une personne, une entreprise ou une institution ne peuvent pas être divulguées sans que ces personnes, entreprises ou institutions le sachent ou y consentent.

Contrôle

Processus qui permet de veiller à ce que les données d'enquête soient exactes, complètes et uniformes. Un ensemble de règles ou de conditions de contrôle s'applique à un **ensemble de données**. Les données qui ne respectent pas les conditions sont examinées et corrigées au besoin.

Couverture

Il s'agit de la mesure dans laquelle chaque personne ou unité devant être incluse dans une enquête ou un recensement est dénombrée dans les faits et dénombrée seulement une fois. Des erreurs de couverture se produisent lorsque des personnes ou des unités de l'enquête ou du recensement sont laissées de côté (sous-dénombrement) ou dénombrées plus d'une fois (surdénombrement). Des études sont souvent menées par Statistique Canada pour produire des estimations du sous-dénombrement et du surdénombrement d'une enquête ou d'un recensement donné, ou pour examiner des questions connexes. Par exemple, Statistique Canada a étudié et analysé la mesure dans laquelle l'utilisation du téléphone cellulaire a des répercussions sur la couverture dans le cas des enquêtes par téléphone.

D

Diffusion

Processus visant à fournir des produits et des services statistiques au grand public et à des utilisateurs de données particuliers. Statistique Canada diffuse les données et les analyses sous forme de résultats d'enquête, de rapports de recherche, de documents techniques, de périodiques, de produits du recensement et de recueils de recherches. Les produits en ligne vont de 1996 à aujourd'hui. Des données historiques peuvent être retrouvées au moyen du catalogue de la bibliothèque. Les données de Statistique Canada sont aussi distribuées par l'entremise d'un réseau approuvé de bibliothèques de dépôt. L'objectif des activités de diffusion est de produire des données pertinentes au moment opportun, selon des formats utiles et des modes accessibles. Les activités en place à l'appui de la diffusion des produits comprennent les services de consultation offerts aux clients, le marketing, les promotions, la formation des utilisateurs et d'autres services à la clientèle.

Distribution de fréquences

Graphique ou tableau montrant combien de fois chaque valeur ou fourchette de valeurs d'une **variable** se retrouve dans un **ensemble de données**. On appelle parfois cela un tableau de fréquences unidimensionnel, afin d'indiquer que la distribution comprend des chiffres pour une variable seulement.

Données

Observations et mesures recueillies dans le cadre d'une enquête, d'un recensement ou d'une autre étude. Faits ou chiffres à partir desquels on peut tirer des conclusions.

Données confidentielles

Il s'agit d'un terme utilisé par Statistique Canada pour décrire les données qui sont assujetties aux dispositions relatives au secret de la *Loi sur la statistique*. Les données sont considérées comme confidentielles parce qu'elles permettent d'identifier directement une **unité répondante**, par exemple, à partir de son nom, ou parce qu'elles pourraient permettre l'identification d'unités répondantes particulières, même après la suppression des identificateurs, en raison des détails qu'elles comportent, de leur structure géographique ou de leur format.

E

Écart type

Mesure de l'écart ou de la dispersion d'un ensemble de données par rapport à la moyenne. Il s'agit de la mesure de dispersion la plus largement utilisée. Au niveau mathématique, l'écart type correspond à la racine carrée de la **variance**.

Échantillonnage

Processus de sélection d'une partie de la population à observer, afin d'estimer un élément d'intérêt concernant l'ensemble de la population. Parmi les exemples des différentes méthodes d'échantillonnage figurent : l'échantillonnage aléatoire simple, l'échantillonnage aléatoire stratifié, l'échantillonnage en grappes, l'échantillonnage à plusieurs phases, et l'échantillonnage à plusieurs degrés.

Échantillonnage stratifié, stratification

Procédure d'échantillonnage dans laquelle la population est divisée en sous-groupes ou strates homogènes, la sélection des échantillons se faisant de façon indépendante dans chaque strate.

Enquête nationale auprès des ménages (ENM)

Cette enquête, à participation volontaire, a été instaurée en 2011, en remplacement du questionnaire complet du recensement, plus largement connu comme la formule « 2B/2D » du recensement. L'ENM vise à recueillir des données sociales et économiques sur la population canadienne. L'objectif de l'ENM est de fournir des données à de petits niveaux géographiques et pour de petits groupes de population. Pour de plus amples renseignements sur l'ENM, suivre ce lien : http://www12.statcan.gc.ca/nhs-enm/index-fra.cfm?HPA

Enquête postcensitaire

Enquête dans laquelle toutes les **unités enquêtées** sont sélectionnées à partir de leurs réponses au **Recensement de la population** ou à l'**Enquête nationale auprès des ménages (ENM)**. Ces enquêtes sont généralement menées peu après le recensement ou l'ENM.

Enregistrement

Données pour **une unité répondante** individuelle, dans un fichier comprenant des données pour toutes les unités répondantes d'une enquête.

Ensemble de données, base de données

Liste organisée et triée de faits ou de données concernant un ensemble de personnes, de ménages, d'entreprises ou d'autres unités pertinentes. Un ensemble de données de Statistique Canada est généralement produit à partir de données d'enquête ou de données administratives, conservé dans un ordinateur et organisé pour pouvoir être accessible facilement dans le cadre d'une vaste gamme d'applications statistiques.

Erreur type

Dans une enquête sur échantillon, les résultats de l'échantillon sont utilisés pour estimer ce que les résultats auraient été si l'ensemble de la population avait été mesuré. L'**erreur d'échantillonnage** désigne le fait que les résultats de l'échantillon pondéré diffèrent dans une certaine mesure des résultats qui auraient été obtenus à partir de la population totale. La différence est désignée comme l'erreur d'échantillonnage. L'erreur d'échantillonnage réelle est évidemment inconnue, mais il est possible de calculer une valeur « moyenne », appelée « erreur type ».

Erreurs

Dans une enquête sur échantillon, les résultats de l'échantillon servent à estimer ce que les résultats auraient été si l'ensemble de la population avait été mesuré. L'exactitude d'une telle **estimation** représente une mesure de la façon dont l'estimation diffère du chiffre correct ou « véritable ». Les écarts par rapport aux chiffres véritables sont appelés « erreurs ». Elles peuvent provenir de nombreuses sources, mais peuvent être regroupées en quelques grandes catégories : erreurs de couverture, erreurs dues à la non-réponse, erreurs de réponse, erreurs de traitement et erreurs d'échantillonnage.

Erreur d'échantillonnage

L'erreur d'échantillonnage rend compte du fait que les résultats de l'échantillon pondéré diffèrent dans une certaine mesure des résultats qui auraient été obtenus auprès de la population totale. La différence est appelée erreur d'échantillonnage. L'erreur d'échantillonnage réelle est évidemment inconnue, mais il est possible de calculer une valeur « moyenne », appelée « **erreur type** ».

Erreurs de couverture

Les erreurs de couverture ont trait aux personnes ou aux unités laissées de côté par l'enquête (sous-dénombrement) ou dénombrées plus d'une fois (surdénombrement).

Erreurs de réponse

Les erreurs de réponse montrent qu'une réponse peut ne pas être entièrement précise. Le répondant peut avoir mal interprété la question ou peut ne pas connaître la réponse, par exemple, si elle concerne un membre du ménage absent.

Erreurs de traitement

Les erreurs de traitement comprennent les erreurs commises pendant l'entrée des données, le codage, la totalisation ou d'autres formes de traitement des données.

Erreurs dues à la non-réponse

Les erreurs dues à la non-réponse se produisent lorsqu'il est impossible d'obtenir des réponses complètes d'une personne, d'un ménage ou d'une organisation. Même si certains ajustements pour tenir compte des données manquantes peuvent être effectués pendant le traitement, la non-réponse signifie inévitablement une perte d'exactitude.

Estimation

Utilisation des résultats de l'échantillon pondéré pour estimer les caractéristiques de l'ensemble de la population.

F

Fichier analytique

Ensemble de **microdonnées** de Statistique Canada pour une enquête donnée, qui peut être utilisé dans les centres de données de recherche (CDR) au Canada. Les CDR permettent aux chercheurs d'accéder aux microdonnées d'enquêtes sur les ménages et sur la population. Les centres comptent des employés de Statistique Canada. Ils sont opérés en vertu des dispositions de la ***Loi sur la statistique*** et sont administrés conformément à toutes les règles de confidentialité. Ils ne sont accessibles qu'aux chercheurs dont les projets ont été approuvés et qui ont été assermentés comme « personnes réputées être employées » en vertu de la ***Loi sur la statistique***.

Fichier de microdonnées à grande diffusion (FMGD)

Fichiers qui donnent accès aux **unités répondantes**, ce qui fait que les utilisateurs peuvent mener leur propre recherche ou analyse. Ils comprennent un ensemble de données non identifiables incluant les caractéristiques relatives aux unités de l'enquête (par exemple, personnes, ménages ou entreprises). La diffusion au public de ces ensembles de données a été autorisée par le Comité de la diffusion des microdonnées de Statistique Canada. Les ensembles de données ne comprennent aucun renseignement confidentiel, étant donné que les identificateurs individuels ont été supprimés, et que toute combinaison de données ou de niveaux géographiques, qui pourrait potentiellement révéler l'identité d'une unité répondante, a été modifiée.

Fréquence

Nombre de fois qu'un événement ou qu'un élément se retrouve dans un **ensemble de données**.

G - H - I

Guides de l'utilisateur

Ces guides accompagnent les ensembles de données d'enquête de Statistique Canada, comme les **fichiers analytiques** et les **fichiers de microdonnées à grande diffusion (FMGD)**, et fournissent des renseignements techniques détaillés nécessaires pour utiliser les données de façon appropriée. Le guide comprend généralement des renseignements importants à connaître avant l'analyse des données : variables de pondération utilisées, procédures relatives à l'estimation de la variance et précautions à prendre pour la diffusion des données.

Imputation

Remplacement de données manquantes ou invalides par des données valides. Elle est généralement exécutée selon des règles déterminées au préalable ou à partir de données d'un « voisin statistique », c'est-à-dire, une autre **unité répondante** qui comporte des caractéristiques similaires. L'imputation est souvent combinée au **contrôle**.

Indien de l'Amérique du Nord

Terme qui décrit tous les Autochtones au Canada qui ne sont pas des **Inuit** ou des **Métis**. Les Indiens de l'Amérique du Nord sont l'un des trois groupes de personnes reconnues comme Autochtones dans la Loi constitutionnelle de 1982. Ce terme désigne aussi les **Premières Nations**, comprenant les Indiens avec statut et les Indiens sans statut.

Information

Données qui ont été enregistrées, classées, organisées, mises en correspondance ou interprétées à l'intérieur d'un cadre, afin d'en obtenir une signification.

Intelligibilité

Facilité avec laquelle l'utilisateur peut comprendre, utiliser de façon appropriée et analyser les données ou l'information. Le degré d'intelligibilité est déterminé dans une large mesure par : la pertinence des définitions des concepts, des populations cibles et des variables; la terminologie qui sous-tend les données, et les renseignements concernant les limites des données.

Inuit

« Inuit » signifie « peuple » en Inuktitut, une des langues des Inuit. La plupart des Inuit vivent dans les Territoires du Nord-Ouest, au Nunavut, dans le Nord du Québec et au Labrador.

Inuit Nunangat

L'Inuit Nunangat est la patrie des Inuit du Canada. Elle comprend les communautés incluses dans les quatre régions inuites : Nunatsiavut (côte nord du Labrador), Nunavik (Nord du Québec), le territoire du Nunavut et la région inuvialuite des Territoires du Nord-Ouest. Ces régions englobent le territoire traditionnellement habité par les Inuit au Canada.

Inuk

Singulier d'**Inuit** (c'est-à-dire « une personne »).

J - K - L

Loi sur la statistique

Loi concernant la statistique au Canada. Comprend la définition du mandat de Statistique Canada : « Est maintenu, sous l'autorité du ministre, un bureau de la statistique appelé Statistique Canada, dont les fonctions sont les suivantes :

- recueillir, compiler, analyser, dépouiller et publier des renseignements statistiques sur les activités commerciales, industrielles, financières, sociales, économiques et générales de la population et sur l'état de celle-ci;
- collaborer avec les ministères à la collecte, à la compilation et à la publication de renseignements statistiques, y compris les statistiques qui découlent des activités de ces ministères;
- recenser la population du Canada et faire le recensement agricole du Canada de la manière prévue à la présente loi;
- veiller à prévenir le double emploi dans la collecte des renseignements par les ministères;
- en général, favoriser et mettre au point des statistiques sociales et économiques intégrées concernant l'ensemble du Canada et chacune des provinces, et coordonner des projets pour l'intégration de telles statistiques. »

Loi sur les Indiens

Dispositions législatives fédérales adoptées pour la première fois en 1876, qui énonce certaines obligations du gouvernement fédéral et réglemente la gestion des terres des réserves indiennes. La loi a été modifiée plusieurs fois, dont la dernière en 1985.

M

Marge d'erreur

Dans une enquête sur échantillon, les résultats de l'échantillon sont utilisés pour estimer ce que les résultats auraient été si l'ensemble de la population avait été mesuré. Dans ce processus d'**estimation**, un certain niveau d'erreur est inévitable. La marge d'erreur, utilisée pour former des intervalles de confiance, sert d'indicateur brut de la précision d'une estimation. Par exemple, les sondeurs disent souvent qu'un certain pourcentage de la population, plus ou moins la marge d'erreur (exprimée en points de pourcentage), devrait voter pour un certain candidat, 19 fois sur 20. Pour calculer la marge d'erreur, qui dans cet exemple correspond à un intervalle de confiance de 95 %, le sondeur utiliserait l'équivalent de plus ou moins deux erreurs types de l'estimation (voir **Erreur type**).

Méthode bootstrap

Méthode servant à estimer l'erreur d'un ensemble de données associé à l'**échantillonnage**. L'échantillonnage entraîne des erreurs, parce que l'on ne recueille pas des données pour l'ensemble de la population, mais uniquement pour un sous-ensemble, appelée échantillon, qui sert ensuite à produire des **estimations** pour l'ensemble de la population. Il existe plusieurs méthodes pour estimer le niveau d'**erreur d'échantillonnage**. Dans le cadre de la méthode bootstrap, on sélectionne habituellement un certain nombre de sous-échantillons de l'échantillon principal et on produit des estimations pour chaque sous-échantillon. L'erreur d'échantillonnage est estimée comme fonction des différences observées entre les estimations des différents sous-échantillons et de l'échantillon complet.

Méthodologie

Ensemble de méthodes et de techniques de recherche appliquées à un domaine particulier d'étude. À Statistique Canada, par méthodologie, on entend la méthodologie d'enquête.

Métis

Bien qu'il n'existe pas de définition unique de Métis, ce terme désigne généralement toute personne ayant une ascendance mixte d'**Indien de l'Amérique du Nord** et européenne qui s'identifie comme Métis.

Microdonnées

Fichier d'**enregistrements** se rapportant à des unités **répondantes** individuelles.

N – O

Observation

Données recueillies pour une variable donnée concernant une **unité répondante** particulière. Parmi les exemples figurent les valeurs particulières d'une unité répondante pour des caractéristiques comme l'âge, le sexe ou l'état matrimonial : l'observation pourrait être « 77 », « femme » et « veuve ».

P

Peuples autochtones

Les peuples autochtones du Canada comprennent les **Premières Nations**, les **Métis** et les **Inuit**.

Plan d'échantillonnage

Ensemble de spécifications qui décrivent les éléments d'échantillonnage d'une enquête de façon détaillée. Ces éléments comprennent la **population**, la **base de sondage**, les **unités** enquêtées par l'enquête, la taille de l'échantillon, la sélection de l'échantillon et la méthode d'**estimation**.

Poids

Nombre moyen d'unités de la population qu'une unité représente dans l'enquête. Parmi les exemples d'unités figurent une personne ou un ménage. Les poids sont appliqués aux **unités répondantes** dans un ensemble de données d'échantillon, afin de veiller à ce que, lorsque l'on fait des inférences à partir des données d'enquête en fonction de paramètres de population, les estimations des caractéristiques correspondent à la population totale.

Population, population cible

Groupe complet d'unités auxquelles les résultats de l'enquête s'appliquent. Ces unités peuvent être des personnes, des ménages, des entreprises, des institutions, etc. Il s'agit de la population sur laquelle on veut obtenir des informations.

Première Nation

Terme qui est devenu d'usage courant dans les années 1970 pour remplacer le mot « Indien », que nombre de personnes trouvaient offensant. Même si le terme Première Nation est largement utilisé, il n'existe pas de définition juridique de ce terme. Le terme « Premières Nations » se rapporte notamment aux **Indiens de l'Amérique du Nord** au Canada avec ou sans statut. Nombre de personnes ont aussi adopté le terme « Première Nation » pour remplacer le mot « bande » dans le nom de leur communauté.

Produit d'information

Organisation des résultats des activités de Statistique Canada, y compris les fichiers de données, les **bases de données**, les tableaux, les graphiques, les cartes et le texte. Cette organisation peut être définie au préalable (produit d'information standard) ou faire suite à des demandes spéciales (produit d'information personnalisé). Les produits d'information sont disponibles sous forme imprimée ou électronique.

Proportion

Relative au nombre de réponses qui se retrouvent dans une catégorie de réponse donnée, en rapport avec les réponses totales. Elle est calculée en divisant la fréquence de la catégorie de réponse par le nombre total de réponses à la question.

Q – R

Qualité des données

Degré ou niveau de confiance quant à l'adaptation des données et des renseignements statistiques à leur utilisation. Les questions particulières de qualité ou d'adaptation à l'utilisation qui doivent être prises en compte par Statistique Canada sont la pertinence, l'exactitude, l'actualité, l'accessibilité, l'**intelligibilité** et la cohérence.

Recensement de la population

Collecte de données concernant toutes les unités d'une population, et est aussi parfois appelé enquête auprès d'un échantillon de 100 %. En vertu de la ***Loi sur la statistique*** de 1971, il est obligatoire de mener un recensement dans l'ensemble du pays tous les cinq ans. Le Recensement de la population fournit des données qui servent à planifier des services communautaires comme des écoles, garderies, services de police et de protection contre les incendies; à prévoir la demande des consommateurs; et à procéder à des études de marché.

Région métropolitaine de recensement (RMR) et agglomération de recensement (AR)

Territoire formé d'une ou de plusieurs municipalités voisines les unes des autres qui sont situées autour d'un grand noyau urbain. Une région métropolitaine de recensement doit avoir une population d'au moins 100 000 habitants et le noyau urbain doit compter au moins 50 000 habitants. L'agglomération de recensement doit avoir un noyau urbain d'au moins 10 000 habitants.

Régression

Méthode statistique servant à prédire la valeur d'une caractéristique, grâce à l'étude de son rapport avec une ou plusieurs autres caractéristiques. Ce rapport est exprimé grâce à une équation de régression.

Régression logistique

Il s'agit d'une forme d'analyse de **régression** utilisée lorsque la variable de réponse est une **variable** binaire (une variable comportant deux valeurs possibles).

Répondant, unité répondante

Personne qui fournit les données pour l'**unité enquêtée**, qu'il s'agisse d'une personne, d'un ménage, d'une entreprise ou d'une institution. Dans le cas de l'EAPA, le répondant est l'adulte sélectionné de 18 ans et plus ou le parent ou le tuteur de chaque enfant ou jeune sélectionné de 6 à 14 ans Pour les jeunes de 15 à 17 ans, l'approbation du parent ou tuteur doit être obtenue pour mener l'entrevue directement avec le jeune. Chez les 15 à 17 ans, le répondant est donc le jeune lui-même ou son parent ou tuteur.

S

Saut valide

Indique le saut d'une question d'enquête parce qu'elle ne correspondait pas à la situation du répondant d'après les réponses valides aux questions précédentes. En pareils cas, le répondant est considéré comme ne faisant pas partie de la population cible ou de l'univers pour cette question.

Subdivision de recensement (SDR)

Il s'agit du terme général pour désigner les municipalités (déterminées en vertu des dispositions législatives provinciales / territoriales) ou les régions traitées comme des équivalents de municipalité aux fins statistiques (p. ex. les réserves indiennes, les établissements indiens et les territoires non organisés).

Suppression

Processus grâce auquel des données particulières ne sont pas diffusées, sur la base de critères conçus pour protéger la confidentialité. La suppression de « cellules » se rapporte aux procédures utilisées pour protéger les données de tableaux de nature délicate de la divulgation, une cellule étant une entrée dans un tableau. Dans le cas de l'EAPA, les données ont aussi été supprimées pour des raisons de qualité (c.v. plus élevé que 33,3 %).

T - U – V

Taux d'échantillonnage ou fraction de sondage

Taille d'échantillon divisé par la taille de la population.

Taux de réponse

Proportion d'un échantillon pour laquelle une réponse au questionnaire est obtenue, habituellement exprimée en pourcentage. La non-réponse englobe ceux qui refusent de participer, ainsi que les personnes que l'on n'a pas pu contacter dans le cadre de l'enquête.

Unité d'échantillonnage ou unité échantillonnée

Unité sélectionnée au moyen du **plan d'échantillonnage** et au sujet de laquelle des mesures sont recueillies dans le cadre d'une enquête. Parmi les exemples figurent les personnes, les ménages, les familles ou les entreprises. Dans le cas de l'EAPA, l'unité d'échantillonnage est la personne.

Unité enquêtée

Unité sélectionnée pour laquelle des mesures sont recueillies pour une enquête sur échantillon ou un recensement. Parmi les exemples figurent les personnes, les ménages, les familles ou les entreprises. Dans le cas de l'EAPA, l'unité enquêtée (qui est aussi l'unité échantillonnée, étant donné que l'EAPA est une enquête sur échantillon) comprend les enfants ou les jeunes de 6 à 14 ans et les adultes de 15 ans et plus.

Variable

Caractéristique qui peut prendre plus d'une valeur, à laquelle une mesure numérique est attribuée (p. ex. revenu, âge et poids).

Variable dérivée

Nouvelle **variable** élaborée grâce à l'application d'opérations logiques ou mathématiques à une ou plusieurs variables existantes, afin de répondre à un besoin de données particulier. Par exemple, une variable de l'âge peut être calculée à partir des données sur la date de naissance. Par ailleurs, une variable dérivée appelée « présence d'un problème de santé chronique » peut être obtenue selon qu'un répondant a répondu « oui » au moins une fois à une série de questions concernant des problèmes de santé chronique particuliers, comme l'asthme, le diabète, la maladie cardiaque, etc.

Variance

Mesure de dispersion d'une caractéristique ou d'une variable donnée dans un ensemble de données. Elle montre dans quelle mesure une variabilité existe pour cette caractéristique. Au niveau technique, elle est calculée comme étant la moyenne des écarts au carré entre chaque observation et leur moyenne sur l'ensemblede données pour une variable particulière.

Bibliographie

BERNIER, Julie, et Karla NOBREGA. 1998. *Outlier detection in asymmetric samples: A comparison of an inter-quartile range method and a variation of a sigma gap method*, congrès annuel de la Société statistique du Canada, juin 1998.

CONSEIL CANADIEN SUR L'APPRENTISSAGE (CCA). (2009). « État de l'apprentissage chez les Autochtones au Canada : Une approche holistique de l'évaluation de la réussite ». Ottawa.

HADDOU, Mohammed. 2013. *Bootstrap Variance Estimation Specifications—Aboriginal Peoples Survey*. Document interne, janvier 2013.

LANGLET, Éric, Jean-François BEAUMONT et Pierre LAVALLÉE. 2008. *Bootstrap Methods for Two-Phase Sampling Applicable to Postcensal Surveys*. Article présenté au Comité consultatif des méthodes statistiques de Statistique Canada, mai 2008, Ottawa.

VERRET, François. 2013. *The estimation methodology of the 2011 National Household Survey*. Article présenté au Joint Statistical Meeting de 2013.