

Plan d'échantillonnage

L'EAPA est une enquête postcensitaire, ce qui signifie que les répondants ont été sélectionnés d'après les réponses qu'ils ont fournies au Recensement de 2001. La majorité des Canadiens ont reçu le questionnaire abrégé du recensement (formulaire 2A), qui ne comprend que sept questions (nom, sexe, date de naissance, état matrimonial légal, situation vis-à-vis l'union libre, relation avec la Personne 1 et première langue apprise dans l'enfance). Un ménage sur cinq a reçu le questionnaire détaillé (formulaire 2B), qui comprend 55 questions, y compris des questions sur l'activité sur le marché du travail, le revenu, la scolarité, les limitations d'activités, la citoyenneté, le logement, l'origine ethnique, etc. Les communautés des Premières nations/réserves et les habitants des régions du Nord ont reçu le Questionnaire du Nord et des réserves (formulaire 2D), dont le contenu est identique à celui du formulaire 2B, sauf pour quelques exemples adaptés. Quatre questions du questionnaire détaillé (formulaires 2B et 2D) ont servi à déterminer la population cible de l'EAPA. La population cible de l'EAPA est constituée de toutes les personnes qui ont déclaré une ascendance autochtone à la question 17 OU qui ont répondu par oui à l'une ou l'autre des questions 18 (identité autochtone), 20 (membre d'une bande indienne ou première nation) ou 21 (indien des traités).

La collecte pour l'EAPA de 2001 a été effectuée en deux phases.

- La phase 1 s'est déroulée d'octobre à décembre 2001 et a mis l'accent sur la population « ayant une identité autochtone ». Cette population est constituée de toutes les personnes qui ont répondu oui à la question 18 (identité autochtone), à la question 20 (membre d'une bande/Première nation) ou à la question 21 (Indien des traités/inscrit).
- La phase II s'est déroulée d'avril à juin 2002 et a mis l'accent sur les personnes qui ont indiqué qu'elles étaient d'ascendance autochtone à la question 17, mais qui n'ont déclaré aucune identité autochtone à la question 18, ni appartenance à une bande/Première Nation à la question 20, ni statut d'Indien des traités ou d'Indien inscrit à la question 21. Cette population est appelée « population d'origine autochtone seulement ». Une faible proportion de la population ayant une identité autochtone était aussi comprise dans la phase II.

Deux plans d'échantillonnage ont été utilisés pour la phase I, un pour les Autochtones vivant « dans les communautés », et l'autre pour les Autochtones vivant « hors des communautés ». Le terme « dans les communautés » comprend les communautés participantes suivantes :

- 123 communautés des Premières nations (réserves);
- 53 communautés inuites dans les régions de l'Arctique;
- 38 communautés comptant au moins 250 Autochtones et une concentration de 40 % ou plus d'Autochtones, dont 28 (y compris 8 établissements métis en Alberta) comptent une forte concentration de Métis;
- 5 autres communautés comptant un nombre important d'Autochtones (Prince Albert, North Battleford, Wood Buffalo, Yellowknife et Whitehorse).

Le terme « hors des communautés » exclut les régions mentionnées précédemment et comprend toutes les autres régions rurales et urbaines à l'extérieur des réserves.

Deux plans d'échantillonnage similaires ont été utilisés pour la phase II, un pour la population d'origine autochtone seulement et l'autre pour la population ayant une identité autochtone¹.

Le tableau 1 présente les tailles d'échantillons et les taux de réponse pour chaque phase. Il convient de mentionner que les tailles d'échantillons et les taux de réponse ne comprennent pas les 21 réserves et 1 communauté inuite (5 455 personnes) qui ont été sélectionnées au départ, mais qui n'ont pas participé à l'enquête.

¹ Étant donné que la population autochtone hors des communautés est très petite au Nunavut, elle a aussi été exclue de l'enquête. Sont aussi exclues les personnes « d'origine autochtone seulement » vivant dans des communautés, à l'exception de cinq communautés importantes (voir **Plan d'échantillonnage – Phase 2**).

Tableau 1
Taille de l'échantillon et taux de réponse selon la phase

	Taille de l'échantillon	Nombre de répondants	Taux de réponse (%)
Phase I (dans les communautés)	57 560	50 594	87,9
Phase I (hors des communautés)	45 710	38 464	84,1
Phase II	13 971	9 591	68,6
Total	117 241	98 649	84,1 %

Plan d'échantillonnage (Phase I – Dans les communautés)

La base de sondage de la partie de la phase I portant sur les personnes vivant dans des communautés était constituée de listes de personnes qui ont été recensées au moyen d'un formulaire détaillé (2B ou 2D) dans le cadre du Recensement de 2001. Au cours des étapes de collecte du recensement, on a examiné les questionnaires provenant de secteurs de dénombrement (SD) déterminés au préalable², et on a dressé la liste de tous les Autochtones, afin de créer la base de sondage ayant servi à tirer l'échantillon. Une liste distincte a été établie pour chaque communauté autochtone.

Le plan d'échantillonnage pour toutes les communautés peut être décrit comme un plan stratifié à deux phases. L'échantillon des formulaires détaillés du Recensement de 2001 a servi à la première phase à identifier les Autochtones. Les personnes ont été réparties en groupes (« strates ») par suite d'un croisement des communautés selon les adultes et les enfants. À la deuxième phase, un échantillon a été sélectionné dans chaque strate, à partir d'un plan d'échantillonnage aléatoire systématique. (L'échantillonnage systématique comprend la sélection d'unités à partir d'une liste, selon un intervalle de sélection établi au préalable.)

Afin de déterminer la taille de l'échantillon de chaque communauté, un niveau d'erreur acceptable (c.-à-d. la fiabilité ou la précision souhaitée des estimations) et une proportion minimale pouvant être estimée à partir de

² Un SD correspond à la région géographique dont le dénombrement est assuré par un recenseur.

l'enquête ont dû être établis et approuvés. Statistique Canada utilise le coefficient de variation (CV.) comme mesure de la fiabilité. (Le CV est une mesure relative de précision qui établit un lien entre l'erreur-type d'une estimation et la taille de cette estimation.)

Pour la partie des personnes vivant dans les communautés de l'EAPA de 2001, une proportion minimale cible de 10 %, avec un CV maximum de 25 %, ont été fixés pour les adultes. La stratégie a alors consisté à estimer le nombre d'adultes nécessaires dans l'échantillon pour répondre à ces exigences et à appliquer la même fraction d'échantillonnage aux enfants. Il convient de souligner qu'étant donné que le nombre d'enfants est plus faible que le nombre d'adultes dans la plupart, sinon dans toutes les communautés, on obtiendra une précision moindre pour les enfants. Cela signifie que, pour les caractéristiques présentes pour 10 % des enfants, le CV sera supérieur à 25 % (moins précis). Par ailleurs pour un CV de 25 %, la proportion minimale d'enfants comportant une caractéristique particulière devra être supérieure à 10 %.

La taille de l'échantillon tient aussi compte des niveaux de réponse attendus. Les taux de réponse prévus allaient de 75 % pour les réserves des Premières nations à 90 % pour les communautés inuites. Ils ont été estimés à partir des résultats obtenus dans le cadre de l'EAPA de 1991. Étant donné que les taux de réponse observés étaient souvent plus élevés que prévu, cela a pu entraîner des CV plus faibles au moment de l'estimation dans certaines communautés.

Une autre approche a été utilisée lors de la détermination de la taille de l'échantillon dans le cas des communautés comportant des concentrations importantes de Métis. Étant donné que l'on a déterminé qu'il était nécessaire de fournir des estimations pour tous les Autochtones de la communauté, ainsi que pour les Métis seulement, on avait besoin d'une fraction d'échantillonnage plus grande. On a d'abord procédé à une estimation du nombre de Métis âgés de 15 ans et plus pour chacune des communautés métis, à partir des projections produites par la Division de la démographie de Statistique Canada et des chiffres du Recensement de 1996. On a par la suite calculé une taille d'échantillon pour la population de Métis adultes, afin d'atteindre les niveaux souhaités de fiabilité. La même fraction d'échantillonnage a été appliquée au reste de chaque communauté.

La taille de l'échantillon total pour cette partie de l'enquête était de 57 560. Le taux de réponse global obtenu a été de 87,9 %.

Plan d'échantillonnage (Phase I – Hors des communautés)

Le plan d'échantillonnage pour la phase I – personnes vivant hors des communautés peut être considéré comme un plan stratifié à deux degrés. Au premier degré, des SD ou groupes de SD (unités primaires d'échantillonnage (UPE)) ont été sélectionnés. Au deuxième degré, tous les répondants des formulaires détaillés (du groupe autochtone pour lequel l'UPE a été sélectionnée au premier degré) ont été sélectionnés. Une description plus détaillée du plan de sondage figure ci-après.

L'échantillon a été stratifié au moyen de ce que l'on appelle les « domaines d'estimation ». Ceux-ci correspondent aux régions géographiques pour lesquelles des estimations comportant un niveau de précision « acceptable » pour un groupe autochtone particulier (IAN, Métis, Inuit) sont requises. Une fois les strates/domaines définis, un échantillon d'UPE a été sélectionné à l'intérieur de chaque domaine. Il convient de souligner que les SD comprenant un petit nombre d'Autochtones ont été regroupés en UPE, tandis que les SD comportant des concentrations importantes ont constitué des UPE à part entière. Un certain nombre d'UPE pour chaque groupe autochtone a été sélectionné.

La méthode d'échantillonnage utilisée pour sélectionner les UPE est appelée « échantillon à probabilité proportionnelle à la taille » ou échantillonnage PPT. À partir de cette méthode, la probabilité de sélectionner une UPE est proportionnelle à sa taille (ou à sa taille prévue). Cela signifie qu'une UPE comptant 100 Autochtones est deux fois plus susceptible d'être sélectionnée qu'une UPE comptant 50 Autochtones. Cette méthode est efficace pour contrôler la taille de l'échantillon du point de vue du nombre d'Autochtones sélectionnés et a comme avantage qu'elle concentre l'échantillon dans les UPE plus grandes, ce qui réduit le coût de la collecte. Certaines UPE avaient une taille tellement importante qu'elles ont été échantillonnées avec certitude. Cela aura pour effet de réduire la variabilité d'échantillonnage.

À l'intérieur de chaque domaine, les UPE dont on savait qu'elles comporteraient une population autochtone très petite ou inexistante ont été exclues du processus de sélection de l'échantillon de la phase I. Étant donné que l'on avait besoin d'une couverture complète, il a été décidé de couvrir ce groupe de personnes à la phase II, au moyen d'un petit échantillon. (Cette décision a été prise parce que le processus de sélection de l'échantillon de la phase I a nécessité des opérations manuelles, tandis qu'à la phase II, la sélection de l'échantillon a été automatisée.)

Les tailles des UPE ont été estimées à partir des projections de la population canadienne de la Division de la démographie. Dans chaque

UPE, la proportion observée de population autochtone (pour chaque groupe) du Recensement de 1996 a été appliquée aux projections démographiques, en vue d'obtenir une projection de la taille de l'UPE.

Une fois qu'une UPE a été sélectionnée, tous les Autochtones ayant reçu un formulaire détaillé du recensement (pour le groupe donné pour lequel l'UPE a été sélectionnée) ont été inclus dans l'échantillon. Comme pour la composante des personnes vivant dans les communautés, la sélection de l'échantillon a été effectuée dans les Sous-sections de la collecte sur le terrain du recensement. Il convient de souligner que des SD pouvaient être sélectionnés pour plus d'un groupe autochtone.

Comme pour l'échantillon des personnes vivant dans les communautés, la taille de l'échantillon a été calculée à partir d'un niveau souhaité de fiabilité pour une proportion minimale donnée. Dans le cas des Inuits vivant hors des communautés (population très petite), on a ciblé uniquement des estimations au niveau du Canada. Les provinces comptant un nombre important d'Indiens de l'Amérique du Nord et de Métis ont été stratifiées selon les principales régions métropolitaines de recensement (Montréal, Ottawa-Gatineau, Toronto, Winnipeg, Regina, Saskatoon, Calgary, Edmonton et Vancouver), les « autres régions urbaines » et les « autres régions rurales » pour les deux groupes.

À l'intérieur de chaque domaine d'estimation (p. ex., les Indiens de l'Amérique du Nord vivant à Calgary), la proportion minimale cible était de 7,5 %, avec un CV variant selon la taille du domaine d'estimation (entre 20 % et 33 %). Plus le CV était petit, plus la précision était grande. On a ciblé des petits CV pour les grands domaines, et des grands CV pour les petits domaines.

La répartition de l'échantillon aux strates a été faite en utilisant une simulation faisant intervenir les données du Recensement de 1996 et les données de l'essai-pilote. L'essai-pilote a été utilisé pour estimer certains paramètres tels que les taux de réponse et les différences possibles dans les populations autochtones entre le Recensement et l'EAPA.

La taille de l'échantillon total pour cette portion de l'enquête était de 47 710, et le taux de réponse obtenu de 84,1 %. Il convient de souligner qu'en raison du faible nombre de personnes vivant hors des communautés au Nunavut, ces personnes ont été exclues de cette portion de l'enquête.

Plan d'échantillonnage (Phase II)

La phase II a été axée sur les personnes ayant déclaré une ascendance autochtone, mais pas une identité autochtone. Une faible proportion de la population déclarant une identité autochtone a aussi été couverte dans la phase II. (Les SD sans population autochtone prévue ou ceux dont la population autochtone constitue un faible pourcentage de la population totale n'ont pas été échantillonnés à la phase I, afin de réduire les opérations manuelles au moment de la sélection de l'échantillon. Cette population a été couverte à la phase II.)

Dans le cadre de la phase II, on a échantillonné uniquement les Autochtones vivant hors des communautés, à l'exception de cinq communautés autochtones (Whitehorse, Yellowknife, Prince Albert, Wood Buffalo et North Battleford) dont la population d'ascendance autochtone qui ne s'identifie pas comme autochtone est importante.

On disposait de la base non vérifiée du recensement pour sélectionner l'échantillon de la phase II. Cette partie de l'enquête a reposé sur un plan d'échantillonnage stratifié à deux phases. À la première phase, on a échantillonné environ un ménage sur cinq dans chaque SD, et à la deuxième phase, on a tiré un sous-échantillon d'Autochtones sélectionnés à la première phase. La stratification utilisée différait pour les échantillons d'ascendance autochtone et d'identité autochtone et comprenait des combinaisons de groupes autochtones, de régions infraprovinciales ou provinciales et d'adultes/enfants. Un échantillon systématique a été tiré de façon indépendante de chaque strate, après le tri des personnes selon des variables clés, comme la région infraprovinciale (lorsque la strate correspondait à la province), le groupe d'âge, le sexe, etc.

La taille d'échantillon était de 13 971 et le taux de réponse de 68,6 %.