

**Conseil économique et social**

Distr. générale
15 juillet 2015
Français
Original : anglais

Commission économique pour l'Europe**Conférence des statisticiens européens****Groupe d'experts des recensements de la population
et des habitations****Dix-septième session**

Genève, 30 septembre-2 octobre 2015

Point 2 b) de l'ordre du jour provisoire

**Innovations projetées pour le cycle de recensements de 2020
et résultats des essais****Plans pour le recensement canadien de 2016
et essais connexes****Note de Statistique Canada***Résumé*

Le Canada procède à un recensement de la population tous les cinq ans, le prochain devant avoir lieu en mai 2016. Cette fréquence présente de grandes possibilités car elle permet d'introduire plus rapidement des innovations concernant tous les aspects du programme, mais elle présente aussi des difficultés à ménager suffisamment de temps pour réaliser des essais élaborés. Une méthode utilisée consiste à concevoir des essais contrôlés directs pendant le recensement en prévision du cycle suivant.

Le présent document décrit succinctement la méthode appliquée au recensement de 2016, puis les essais menés en prévision de ce recensement et certaines des conclusions tirées. Il présente aussi les essais directs projetés pour 2016. Enfin, il décrit brièvement les travaux qu'il est prévu de mener pour que le programme de recensement s'appuie davantage sur les données des fichiers administratifs, ainsi que certaines activités de recherche et certains essais visant à vérifier les applications anticipées des sources susmentionnées.



I. Introduction

1. Statistique Canada est chargée, conformément à la loi sur la statistique, de procéder à un recensement de la population tous les cinq ans. Le programme comprend un recensement de la population s'appuyant sur un questionnaire comportant environ 10 questions se rapportant à la démographie de base et une enquête nationale auprès des ménages, qui recueille des renseignements socioéconomiques plus détaillés sur un échantillon de ménages. Une approche mixte est utilisée : d'une part, des données sont collectées à un moment donné auprès d'individus et au sujet d'unités d'habitation et, d'autre part, certaines informations sont obtenues à partir des registres administratifs (par exemple, des variables de revenu sont créées à partir des registres de l'impôt sur le revenu). Un recensement tous les cinq ans permet d'accélérer l'apprentissage de méthodes innovantes et d'apporter des modifications essentielles à tout élément du programme. Une certaine discipline est également indispensable pour que les essais nécessaires à la mise en œuvre de telles innovations soient effectués en temps voulu. Il est fondamentalement obligatoire, pendant le recensement, de planifier déjà des modifications plus importantes pour un cycle futur.

2. Dans le cadre du processus habituel de Statistique Canada, toutes les modifications majeures apportées aux méthodes et approches relatives à des programmes statistiques importants sont entièrement testées avant leur mise en application. Dans le programme de recensement de la population, des essais de contenu et des essais opérationnels sont généralement réalisés dans les deux à trois ans précédant la principale collecte elle-même. Pour le programme de 2016, un essai de contenu (BT2) a été mené en mai/juin 2014, et un essai visant les modifications opérationnelles (BT3) en octobre 2014. Des essais qualitatifs (BT1) concernant le contenu du questionnaire et les instruments utilisés ont été effectués avant l'essai BT2.

3. En particulier pour le programme de recensement, de nouvelles approches ont parfois également été testées dans des environnements contrôlés pendant les opérations de recensement en prévision du prochain cycle. De tels essais sont actuellement élaborés pour 2016.

4. La présente note décrit brièvement les essais menés ou projetés en rapport avec le recensement de 2016. On y trouvera les principales conclusions tirées pour les essais déjà réalisés. Des hypothèses de recherche sont exposées pour les essais à venir. De même, Statistique Canada mène actuellement des activités de recherche pour généraliser l'utilisation de données administratives lors du recensement. Le document décrit succinctement certains des projets de recherche et essais projetés liés au programme de 2016.

II. Essais de comportement pour le cycle de 2016

5. Au cours de ces dernières années, Statistique Canada a remanié son architecture opérationnelle globale, le but étant de rationaliser les investissements dans l'infrastructure et les processus d'entreprise en élaborant des approches globales par opposition à des approches locales applicables à des programmes individuels. Le programme de recensement fait partie de ces initiatives, en particulier du point de vue selon lequel les gros investissements requis par le programme peuvent être exploités pour appuyer la mise en place d'autres initiatives de l'institution. Ainsi, les systèmes opérationnels de terrain utilisés pour le recensement peuvent être développés pour répondre aux besoins de toutes les autres enquêtes, ce qui réduit les investissements institutionnels dans ces systèmes ainsi que le nombre de systèmes et de plates-formes à entretenir dans le temps. Pour que le programme de recensement puisse utiliser les

outils institutionnels en 2016, on a remanié la stratégie d'essai pour qu'elle vise principalement les modifications faites depuis le dernier recensement de 2011, au lieu de procéder à une « répétition générale » complète. Cette stratégie a été subdivisée en de petits essais de comportement (BT), l'objectif étant de vérifier le comportement des répondants et du personnel de terrain à l'égard des nouveaux systèmes et approches.

6. Le contenu du questionnaire et les instruments ont été testés dans l'essai de comportement 2 (BT2) en mai et juin 2014. En particulier, le nouveau questionnaire électronique a été utilisé en remplacement de l'application spécifique utilisée depuis le recensement de 2006. La capacité d'essai a permis de vérifier si le questionnaire électronique était applicable à diverses plates-formes technologiques, telles qu'ordinateurs de bureau, tablettes et smartphones. Avec un objectif de 65 % de réponses en ligne pour 2016 et compte tenu de l'évolution rapide des technologies, nous devons confirmer que le questionnaire électronique allait fonctionner correctement sur toutes les plates-formes susceptibles d'être utilisées par les répondants en 2016. L'essai a confirmé que l'objectif fixé pour l'utilisation d'Internet était adéquat et que le questionnaire fonctionnait comme prévu sur toutes les plates-formes. Un autre petit essai de plate-forme sera effectué à la fin de 2015 pour vérifier si le questionnaire électronique fonctionnera avec les dernières technologies.

7. L'essai de comportement 3 (BT3) a été réalisé à l'automne de 2014 dans le but de tester l'utilisation de nouveaux systèmes et processus de collecte par le personnel de terrain ainsi que le matériel de formation connexe. Nous avons en particulier testé le nouveau Système intégré de collecte et des opérations (SICO). Ce dernier remplace le système des opérations de terrain utilisé lors du recensement de 2011 et offre une interface Internet permettant au personnel de terrain de prendre connaissance des tâches qui lui sont assignées, d'enregistrer l'état d'avancement des travaux ainsi que les informations connexes en matière de rémunération et de dépenses. Le système comprend aussi un certain nombre de fonctions connexes telles que la messagerie, les rapports sur les informations de gestion, etc. L'essai BT3 a été différencié de l'essai BT2 et réalisé ultérieurement, contrairement aux pratiques passées afin de laisser plus de temps pour achever la mise au point des fonctions adéquates du SICO devant être utilisées lors du recensement.

8. L'essai portant sur le SICO a montré que certaines fonctions conçues pour être génériques (c'est-à-dire pour être utilisées par des enquêteurs ordinaires dans toutes les enquêtes et des agents recenseurs) étaient trop complexes dans le cadre d'un recensement. En effet, il est souhaitable en l'espèce de limiter le volume de formation requis/dispensé pour des raisons financières, et parce que les tâches assignées sont limitées à de courtes périodes, d'où la nécessité d'une courbe d'apprentissage assez courte. Sur la base des résultats de cet essai, certaines des fonctions génériques ont été simplifiées pour pouvoir être utilisées en 2016. Un exemple en est les multiples fonctions mises au point pour permettre aux superviseurs de terrain de répartir les tâches des agents recenseurs, telles que celles fondées sur le nombre moyen de tentatives par cas, la proximité des cas par rapport au domicile des agents recenseurs, etc. Ces fonctions se sont avérées trop complexes dans le cadre du recensement.

9. Plusieurs ajustements et modifications sont actuellement apportés aux systèmes d'après les résultats des essais de comportement. Pour confirmer que ces modifications résoudront efficacement les problèmes observés lors des essais, Statistique Canada effectuera un autre petit essai de validation des systèmes à l'automne de 2015. Des non-professionnels seront recrutés pour procéder à ces essais afin de vérifier que les systèmes seront intuitifs lors du recensement lorsque plus de 25 000 personnes seront recrutées et formées pour des missions de courte durée de six à dix semaines.

III. Essai concernant une utilisation accrue des données administratives

10. En 2014, Statistique Canada a lancé un programme de recherche pour étudier s'il était possible de généraliser l'utilisation des données administratives pour le recensement de la population. L'un des objectifs de ces recherches est de vérifier si les données provenant de sources administratives pourraient être utilisées pour remplacer les traditionnelles opérations de dénombrement des individus et de collecte des caractéristiques démographiques de base. Le Canada applique une méthode plus classique pour le recensement de la population puisqu'il ne tient pas de registre universel de la population ou des logements ni ne dispose d'identificateur personnel unique universel (Royce, 2013).

11. L'une des principales initiatives de Statistique Canada visant à appuyer les activités de recherche ayant trait aux sources administratives consiste à concevoir des prototypes de la Base statistique de données démographiques canadiennes (BSDDC) (Cyr *et al.*, 2013). Cette base de données est créée à partir des fichiers relatifs à l'impôt fédéral sur le revenu. S'y ajoutent les naissances et en sont soustraits les décès, compte tenu également de certains sous-groupes de population enregistrés dans d'autres fichiers tels que le Registre des Indiens et la Base de données longitudinales sur les immigrants. Les deux premiers prototypes de la BSDDC ont été conçus pour représenter la population à la date de référence de 2011, les écarts générés ne dépassant pas 1 % des estimations démographiques nationales pendant la même période. Les prototypes ont toutefois affiché de plus grands écarts à des niveaux géographiques inférieurs ou pour certains segments de la population définis selon l'âge et le sexe.

12. Statistique Canada n'a pas encore déterminé ce que les critères de succès ou les délais devraient être pour l'achèvement des travaux. Dans l'intervalle, les résultats éventuels de la BSDDC seront néanmoins évalués au cours des deux prochaines années et peuvent ne pas exiger le même niveau de couverture globale. Par exemple, si des renseignements démographiques de base étaient disponibles à temps pour un grand nombre de logements, ils pourraient éventuellement être utilisés pour remplacer les suivis coûteux effectués sur le terrain pendant le recensement. Il serait peut-être possible d'utiliser les informations pour valider l'état d'occupation des logements le jour du recensement, remplacer une partie du suivi des non-réponses sur le terrain, cibler les zones difficiles à recenser, appuyer les études de couverture, etc.

13. La création d'un prototype de profil statistique de la population (BSDDC) présente un certain nombre de défis, l'un étant le caractère d'actualité. Un fichier représentant une grande partie de la population peut-il être créé à temps pour coïncider avec les opérations de recensement? En l'espèce, la difficulté tient au temps nécessaire pour accéder aux fichiers de l'année de collecte et les mettre en correspondance, c'est-à-dire que les informations doivent être aussi à jour que possible. Une autre difficulté dans le contexte canadien est la capacité à placer les personnes figurant dans les fichiers administratifs dans le bon logement, à savoir leur résidence principale. La plupart des fichiers administratifs dont dispose Statistique Canada ne requièrent pas ce type d'information sur les adresses. Ces informations peuvent se rapporter à des résidences secondaires ou, dans le cas des données fiscales, elles pourraient même être l'adresse de celui qui a rempli la déclaration d'impôt.

14. Il est envisagé de recourir à deux ensembles d'essais afin d'évaluer si la BSDDC peut être utilisée pour faciliter le recensement. On peut vérifier les effets que pourrait avoir sur la qualité des données un remplacement du dénombrement direct par des données administratives et ce, en simulant l'utilisation de la BSDDC dans une configuration opérationnelle avec les informations relatives au recensement de 2016.

Des essais de contrôle seront mis au point pour savoir quelle aurait été l'incidence de l'utilisation de la BSDDC à la place du suivi des non-réponses concernant certains logements en 2016. Ainsi, nous pourrions instituer une règle selon laquelle il n'y aurait plus de suivi de certains logements après « x » tentatives, à condition que la BSDDC contienne de bons renseignements sur les occupants des logements retenus pour l'essai. La date utilisée pour la BSDDC devrait être aussi proche que possible de la date du recensement de 2016 (c'est-à-dire le 10 mai). On vérifierait si la collecte directe de données dans ces logements peut être remplacée par les renseignements figurant dans la BSDDC sans une incidence importante sur les données obtenues au niveau des petites zones. Comme la BSDDC ne contiendrait pas toutes les variables recueillies dans le questionnaire (des informations sur la langue seraient absentes par exemple), les renseignements manquants seraient estimés. On comparerait ensuite les résultats figurant dans le nouveau fichier aux résultats officiels du recensement de 2016 pour en vérifier l'exactitude.

15. L'un des principaux avantages que présente le remplacement du suivi sur le terrain par des données administratives serait que l'on obtiendrait une importante réduction des coûts, en diminuant le nombre de contacts avec les occupants de logements n'ayant pas répondu, cette opération étant l'une des plus onéreuses. Une partie des économies faites proviendrait d'une diminution des frais fixes en raison de la plus rapide réduction des effectifs de terrain, la collecte s'arrêtant plus tôt dans certaines zones.

IV. Essai concernant la collecte d'un identificateur personnel

16. Toutes les corrélations de registres effectuées dans le cadre des recherches sur les données administratives réalisées à Statistique Canada sont fondées sur des combinaisons du nom, de l'âge ou année de naissance, du sexe et de l'adresse des répondants. Bien qu'il n'y ait pas au Canada d'identificateur personnel unique universel qui puisse être utilisé pour faciliter ces corrélations, le numéro d'assurance sociale (NAS) pourrait être un substitut valable. La plupart des Canadiens ont un NAS, qui figure aussi dans certains des fichiers administratifs dont dispose Statistique Canada, par exemple les registres de l'impôt sur le revenu. Le problème est que Statistique Canada n'a pas actuellement un accès facile aux NAS.

17. La collecte des numéros d'assurance sociale directement auprès des répondants a été testée dans le cadre de l'essai BT2. Cette partie de l'essai visait à vérifier la propension des Canadiens à indiquer leur NAS dans le questionnaire, leur capacité à fournir des numéros exacts par procuration (une personne remplissant généralement le questionnaire pour l'ensemble du ménage) et les effets possibles sur les taux de réponse et de retour concernant le recensement et l'enquête nationale auprès des ménages pour des motifs de confidentialité et autres. Ces effets ont été mesurés pour les modes de réponse en ligne et sur papier, puis mesurés par rapport à des panels de contrôle dont le NAS n'a pas été collecté. Le volet recensement de l'essai a également été effectué selon les dispositions obligatoires de la loi sur la statistique afin de reproduire les conditions réelles d'un recensement.

18. L'essai a montré que la majorité des gens fourniraient leur NAS et que la qualité serait le plus souvent élevée. Un peu plus de 80 % des répondants ont fourni un NAS, et 1,6 % seulement un numéro non valable. L'analyse des résultats n'a pas porté sur les répondants qui n'ont pas de NAS, de sorte que la proportion réelle de ces répondants n'est pas connue. Certains effets opérationnels ont également été observés. La charge de travail pour les répondants a augmenté d'au moins 20 % pour les questionnaires demandant le NAS comme l'indique le temps supplémentaire

nécessaire pour remplir le questionnaire en ligne. Une baisse du taux de retour a aussi été observée – environ 2 % pour l'enquête nationale auprès des ménages et 2 à 3 % pour le recensement s'agissant des vagues 2 et 3 (lorsque des rappels sont envoyés aux non-répondants). Un taux de retour moindre ne donne pas forcément un faible taux de réponse à la fin de la collecte, mais signifierait une charge de travail plus importante au début du suivi des non-réponses, et probablement une augmentation du coût de la collecte sur le terrain.

19. Une analyse plus poussée a montré que le fait de connaître le NAS n'améliorait pas notablement les taux de corrélation avec les données fiscales (augmentation de moins de 1 %), ce qui confirme que la méthode actuellement appliquée par Statistique Canada est assez efficace. Cela étant, connaître le NAS accélérerait à l'évidence le processus de corrélation. Les résultats n'étant pas concluants, le NAS ne sera pas collecté lors du recensement de 2016.

V. Essais directs projetés pour le recensement de 2016

20. Statistique Canada procédera à quelques essais pendant la l'opération de collecte de données de 2016, afin de générer des connaissances qui serviront de base à la planification du prochain recensement de 2021. Ces essais porteront sur la collecte d'une adresse électronique comme information de contact, la collecte d'une adresse postale et d'une adresse de listage si elles sont différentes, l'envoi en ligne de l'invitation à répondre au recensement et la généralisation de l'envoi de questionnaires par la poste dans certaines régions.

21. Il sera demandé aux répondants leur adresse électronique dans le questionnaire de 2016. Ce renseignement sera ajouté à la base de sondage pour les enquêtes auprès des ménages, qui sert de source d'adresses pour le recensement. La collecte de l'adresse électronique fait partie des recherches ayant trait à d'autres modes de prise de contact avec les ménages canadiens, soit pour le recensement même, soit pour toute enquête utilisant comme base d'échantillonnage la base de sondage pour les enquêtes auprès des ménages. L'essai mesurera le taux de réponses positives (quelle proportion a donné une adresse électronique), la capacité à collecter ce renseignement avec exactitude et les éventuelles conséquences sur la saisie de données concernant les répondants utilisant les questionnaires sur papier. S'agissant plus particulièrement de ce dernier aspect, nous vérifierons dans quelle mesure la reconnaissance optique permet de lire directement et de reconnaître les adresses électroniques.

22. En parallèle à la collecte des adresses électroniques, Statistique Canada testera une méthode électronique permettant de prendre contact avec les ménages pendant les opérations de collecte de 2016. Il s'agira de prendre contact avec 50 000 à 100 000 ménages au moyen du service de boîte postale numérique offert par Postes Canada (« postel »). Cette méthode sera utilisée sur un panel d'abonnés au service « postel » et remplacera le mode de prise de contact par courrier normalement prévu. Nous choisirons aussi un panel de contrôle parmi les abonnés au service « postel » pour valider les résultats obtenus avec le panel initial. Le service « postel » exigera que les ménages acceptent d'abord Statistique Canada comme organisme autorisé à envoyer des données à leur boîte postale numérique. Ceux qui donneront leur consentement recevront une invitation à remplir le questionnaire de recensement dans leur boîte postale numérique. Le contenu du message numérique sera très semblable à celui de la lettre d'invitation adressée aux ménages par la poste traditionnelle. L'essai mesurera le taux de retour et de réponse du groupe numérique par rapport au groupe de contrôle. Les résultats serviront à planifier des stratégies de prise de contact avec les ménages pour 2021.

23. Statistique Canada a adopté plusieurs stratégies pour prendre contact avec les ménages. Selon l'hypothèse actuellement appliquée à 2016, les occupants d'environ 82 % des logements recevront initialement par courriel une lettre les invitant à répondre en ligne. Pour 17 % des logements, Statistique Canada enverra directement les questionnaires aux logements selon la méthode « listage-livraison », essentiellement en zone rurale, puisque le système de publipostage ne permet pas de distribuer du courrier à un logement précis sans disposer d'une adresse. Les résidents de ces zones recevront leur courrier dans les boîtes à lettres ordinaires puisque Statistique Canada ne peut adresser du courrier qu'aux « occupants ».

24. La délimitation des zones où les questionnaires sont envoyés par la poste et des zones de « listage-livraison » se fait en partie en fonction de critères opérationnels. Ainsi, pour rester dans les zones de « listage-livraison », certains logements pourraient recevoir des courriers, et ce afin d'éviter qu'il y ait des zones où une méthode mixte est appliquée, ce qui simplifie les procédures pour le personnel de terrain. Pour l'essai proposé, nous testerons une approche consistant à envoyer des questionnaires par la poste et à utiliser une méthode complémentaire. Pour certaines zones où Statistique Canada dispose d'une grande proportion d'adresses valables, nous enverrons par la poste une lettre d'invitation à ces adresses et nous fournirons au personnel de terrain, pour la remise des lettres, la liste des logements connus restants pour lesquels nous n'avons pas d'adresse postale valable. L'essai vérifiera si cette approche peut être mise en œuvre, si les coûts peuvent être réduits par rapport à une véritable approche « listage-livraison » et si la couverture des logements peut être maintenue. Si l'approche porte ses fruits, Statistique Canada pourrait étendre l'envoi par la poste à au moins 90 % des logements en 2021.

25. Enfin, Statistique Canada collectera aussi l'adresse de listage et l'adresse postale de chaque logement. Le registre des adresses contient des renseignements sur l'adresse de listage, qui parfois ne concorde pas avec l'adresse postale, notamment en zone rurale. L'obtention d'une concordance entre ces deux adresses peut permettre de généraliser le recours à la poste lors du prochain recensement de 2021. Cela appuiera aussi les recherches concernant les sources administratives en permettant de placer les personnes au bon endroit, élément essentiel pour la création d'un profil statistique de la population.

VI. Références

Boulet C., Mathieu P., Statistique Canada, 2015, Testing the addition of the Social Insurance Number to the Canadian Census of Population; document présenté aux réunions statistiques conjointes de l'American Statistical Association, Seattle, août 2015.

Royce D. 2011. Preliminary Report on Methodology Options for the 2016 Census. Document élaboré pour la Stratégie du recensement de 2016, Bureau de gestion du recensement.

Cyr A., Pignal J. Trépanier J., Statistique Canada, 2013 Canadian Statistical Demographic Database : A Research Project, document présenté au Symposium sur les questions de méthodologie organisé par Statistique Canada.