
Rapport technique sur la couverture

Recensement de la population, 2021



Date de diffusion : le 23 octobre 2024



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 suivants :

- | | |
|---|----------------|
| • Service de renseignements statistiques | 1-800-263-1136 |
| • Service national d'appareils de télécommunications pour les malentendants | 1-800-363-7629 |
| • Télécopieur | 1-514-283-9350 |

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 « Contactez-nous » > « [Normes de service à la clientèle](#) ».

Note de reconnaissance

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

Publication autorisée par le ministre responsable de Statistique Canada

© Sa Majesté le Roi du chef du Canada, représenté par le ministre de l'Industrie, 2024

L'utilisation de la présente publication est assujettie aux modalités de l'[entente de licence ouverte](#) de Statistique Canada.

Une [version HTML](#) est aussi disponible.

This publication is also available in English.

Date de diffusion : le 23 octobre 2024

N° 98-303-X au catalogue, numéro 2021001

ISBN 978-0-660-47834-0

Table des matières

Introduction	1
1. Estimation des erreurs de couverture de la population	3
1.1 Introduction	3
1.2 Sous-dénombrement net	3
1.3 Sous-dénombrement	9
1.4 Surdénombrement	13
2. Univers du recensement	15
2.1 Introduction	15
2.2 Univers de la population	15
2.3 Univers des logements	15
2.4 Lieu habituel de résidence	17
3. Erreur de couverture de la population	18
3.1 Sources	18
3.2 Contrôle	19
3.3 Définitions	20
3.4 Évaluation	22
4. Collecte des données du recensement	23
5. Traitement des données du recensement	25
6. Enquête sur la classification des logements	26
6.1 Introduction	26
6.2 Méthodologie	26
6.2.1 Stratification et sélection de l'échantillon	26
6.2.2 Interviews sur le terrain	28
6.2.3 Traitement et estimation	28
6.2.4 Imputation des ménages entiers du recensement	28
6.3 Estimations	29
6.3.1 Logements inoccupés	30
6.3.2 Logements non répondants	34
7. Étude sur le sous-dénombrement du recensement	39
7.1 Échantillonnage	39
7.2 Traitement et classification	45
7.2.1 Traitement	45
7.2.2 Classification	46

7.3 Collecte.....	49
7.3.1 Aperçu	49
7.3.2 Opérations	50
7.3.3 Dépistage.....	52
7.3.4 Statistiques de collecte.....	52
7.4 Estimation	55
7.4.1 Calcul du poids initial	55
7.4.2 Ajustements aux poids initiaux	55
7.4.3 Rajustement de la non-réponse	55
7.4.4 Ajustements finaux aux poids des personnes classées	56
7.4.5 Distribution pondérée selon la classification	58
7.4.6 Calcul du sous-dénombrement du recensement	59
8. Étude sur le surdénombrement du recensement.....	61
8.1 Aperçu.....	61
8.2 Étapes du couplage.....	61
8.2.1 Données utilisées pour les couplages	61
8.2.2 Couplage déterministe.....	62
8.2.3 Couplage probabiliste	62
8.3 Création de la base de sondage de l'Étude sur le surdénombrement du recensement	63
8.4 Plan d'échantillonnage.....	65
8.4.1 Strate du couplage déterministe seul.....	66
8.4.2 Strate du couplage probabiliste seul	67
8.4.3 Strate du couplage probabiliste-couplage déterministe	71
8.4.4 Tailles d'échantillon finales (par paire).....	72
8.5 Opération de vérification manuelle	73
8.6 Pondération et estimation.....	75
8.6.1 Pondération	75
8.6.2 Estimation	76
8.7 Résultats	77
8.7.1 Surdénombrement par composante.....	77
8.7.2 Surdénombrement par scénario	79
9. Estimation.....	82
10. Évaluation des études sur la couverture	84
10.1 Étude sur le sous-dénombrement du recensement.....	84
10.1.1 Introduction.....	84
10.1.2 Comparaisons avec les chiffres du recensement	84
10.1.3 Comparaison avec les estimations démographiques.....	86

Rapport technique sur la couverture

10.2 Étude sur le surdénombrement du recensement	89
10.3 Estimations démographiques	90
10.3.1 Erreur en fin de période	90
10.3.2 Exactitude des estimations postcensitaires	91
11. Estimations historiques des erreurs de couverture de la population	93
11.1 Estimations	93
11.2 Modifications apportées à la conception des études sur la couverture de la population	105
12. Sujets spéciaux	109
12.1 Sous-dénombrement de collecte	109
12.2 Participation des réserves et des établissements	112
12.2.1 Introduction	112
12.2.2 Réserves et établissements partiellement dénombrés	112
Annexe A — Qui inscrire dans le questionnaire du recensement	114
Bibliographie	115

Introduction

Le Recensement de 2021 a nécessité la participation de l'ensemble de la population canadienne, soit environ 37 millions de personnes réparties sur un territoire couvrant 9 millions de kilomètres carrés. Même si des normes de qualité rigoureuses régissent la collecte et le traitement des données, il est impossible d'éliminer toutes les erreurs. Pour utiliser les données du recensement correctement et de façon appropriée, il est important de comprendre le cadre conceptuel et les définitions utilisées pour effectuer le recensement, ainsi que les méthodes de collecte et de traitement des données. Les utilisateurs doivent également connaître les principales sources d'erreur et, dans la mesure du possible, l'ampleur des erreurs et toute circonstance inhabituelle qui pourrait limiter l'utilité ou l'interprétation des données du recensement. Grâce à cette information, les utilisateurs peuvent évaluer les risques associés à l'utilisation des données du recensement pour tirer des conclusions ou prendre des décisions.

Le présent rapport technique porte sur les erreurs de couverture du Recensement de 2021. Il existe deux types d'erreurs de couverture. Le premier, appelé **sous-dénombrement de la population**, désigne l'exclusion de personnes qui auraient dû être dénombrées. Le second, appelé **surdénombrement de la population**, désigne quant à lui l'inclusion de personnes qui n'auraient pas dû être dénombrées ou qui l'ont été plus d'une fois. Dans le premier cas de surdénombrement, l'erreur globale est jugée peu importante. En général, le sous-dénombrement est plus fréquent que le surdénombrement. L'incidence nette du sous-dénombrement et du surdénombrement sur la taille d'une population étudiée est appelée le **sous-dénombrement net de la population**. Le sous-dénombrement net représente le nombre de personnes exclues qui auraient dû être dénombrées (sous-dénombrement), moins le nombre de dénombrements en trop de personnes qui ont été dénombrées plus d'une fois (surdénombrement). Les erreurs de couverture figurent parmi les erreurs les plus importantes puisque du point de vue de l'exactitude, elles ont une incidence non seulement sur les chiffres des différents univers du recensement, mais également sur l'ensemble des données du recensement relatives aux caractéristiques de ces univers.

Les erreurs de couverture commises lors du recensement sont mesurées à l'aide de trois études. La première, l'Enquête sur la classification des logements (ECL), porte sur les erreurs de couverture attribuables aux erreurs de classification des logements. Les données du recensement sont corrigées pour tenir compte de ce type d'erreur de couverture. La deuxième, l'Étude sur le sous-dénombrement du recensement (ESoR) auparavant appelée Contre-vérification des dossiers (CVD), permet d'évaluer le sous-dénombrement de la population. La troisième, l'Étude sur le surdénombrement du recensement (ESuR), permet quant à elle d'évaluer le surdénombrement de la population. Les données du recensement ne sont pas corrigées pour tenir compte des erreurs de couverture décelées par l'ESoR ou l'ESuR. Statistique Canada utilise plutôt des estimations du sous-dénombrement net pour établir des estimations démographiques. Les études menées dans le cadre du Recensement de 2021 sont assez semblables à celles de 2016, bien que certains changements et certaines améliorations aient été apportés.

Les utilisateurs des données du recensement doivent savoir que la présence d'erreurs de couverture commises lors du Recensement de 2021 signifie que les produits du recensement font l'objet d'un dénombrement incomplet ou en double. Le taux de sous-dénombrement, par exemple, est plus élevé chez les jeunes hommes adultes. Pour obtenir les estimations des erreurs de couverture commises lors du Recensement de 2021 pour divers niveaux et regroupements démographiques et géographiques, voir la [section 1](#).

La [section 2](#) porte sur le cadre conceptuel du Recensement de 2021 et définit l'univers de la population, l'univers des logements et le lieu habituel de résidence, que le recensement vise à mesurer. La [section 3](#) décrit les erreurs de couverture, leur cause, les méthodes de recensement qui permettent de les réduire au minimum et le cadre conceptuel utilisé pour les évaluer. Elle présente également les études sur la couverture du recensement. La [section 4](#) et la [section 5](#) décrivent quant à elles la méthodologie adoptée lors du Recensement de 2021 en matière de collecte et de traitement des données.

La [section 6](#) présente la méthodologie utilisée lors de l'ECL de 2021 et les résultats obtenus. Les estimations des taux d'erreur de couverture ne sont produites que pour l'univers de la population. La [section 7](#) et la [section 8](#) décrivent la méthodologie utilisée et les résultats obtenus dans le cadre de l'ESoR de 2021 et de l'ESuR de 2021, respectivement. La [section 9](#) indique pour sa part la façon dont les résultats de l'ESoR et de l'ESuR ont été combinés aux données du recensement pour produire les estimations des erreurs de couverture de la population et les erreurs-types connexes.

La [section 10](#) porte sur les résultats des évaluations effectuées pour l'ESoR et l'ESuR, ainsi que sur l'évaluation de l'erreur en fin de période. L'« erreur en fin de période » fait référence à la différence entre les estimations démographiques ayant comme population de base les chiffres du Recensement de 2016, corrigés pour tenir compte du sous-dénombrement net, et les données du Recensement de 2021, également corrigées pour tenir compte du sous-dénombrement net.

Statistique Canada mène des études sur la couverture de la population du recensement depuis la première ESoR, qui remonte au Recensement de 1961¹. Dans une perspective chronologique, la [section 11](#) retrace les erreurs de couverture des recensements de 1971 à 2021.

La [section 12](#) aborde d'autres sujets, notamment le concept de personnes non dénombrées et la participation des réserves et des établissements au recensement.

Le présent rapport a été préparé par : Melanie Abeysondera, Karen Bruce, Emily Campling, Monique Daweside, Heather Farr, Vincent Martin, Michel Parenteau, Martin Provost, Hélène St-Jean, Martin St-Pierre et Zhao Zhao de la Division des méthodes d'intégration statistique, et Julien Bérard-Chagnon du Centre de démographie.

De nombreuses personnes ont fait des commentaires utiles en ce qui a trait aux versions préliminaires, ce qui a permis d'améliorer la teneur et la lisibilité du rapport définitif.

Le [Dictionnaire, Recensement de la population, 2021](#), produit n° 98-301-X au catalogue de Statistique Canada, renferme de plus amples renseignements à propos des concepts, des variables et de la géographie du recensement. Les [produits de référence, Recensement de 2021](#) fournissent des renseignements supplémentaires à propos du processus de recensement.

1. La première ESoR a été réalisée en 1961, mais il n'y avait pas de base de sondage pour les personnes omises lors du recensement précédent. Dans l'ESoR de 1966, on a utilisé les résultats de l'ESoR de 1961 en vue de l'établissement de la base de sondage des personnes omises lors du Recensement de 1961.

1. Estimation des erreurs de couverture de la population

1.1 Introduction

Le recensement définit la population devant être dénombrée, ainsi que les règles applicables aux fins du dénombrement (voir la [section 3](#)). Des erreurs de couverture surviennent lorsque des erreurs sont commises relativement à ces règles et définitions. Les principales erreurs de couverture sont attribuables à l'omission d'un logement (et, par conséquent, de ses résidents) et à des erreurs des répondants (omission de personnes devant être incluses ou inclusion de personnes ne devant pas être incluses). La présente section fournit des estimations du sous-dénombrement net, du sous-dénombrement et du surdénombrement de la population survenus lors du Recensement de 2021. Un sous-dénombrement et un surdénombrement peuvent entraîner un biais dans les chiffres officiels et les estimations, car les caractéristiques des personnes non incluses peuvent différer de celles des personnes incluses, et les caractéristiques de personnes incluses plusieurs fois (doublons) peuvent différer de celles des personnes incluses une seule fois. Le sous-dénombrement net indique la mesure dans laquelle le nombre de dénombrements inclus dans les données du recensement est supérieur ou inférieur au dénombrement complet.

1.2 Sous-dénombrement net

Le taux de sous-dénombrement net de la population lors du Recensement de 2021 a été évalué à 3,00 %². Autrement dit, la différence entre le nombre de personnes non incluses dans le recensement, mais faisant partie de la population cible du recensement, et le nombre de dénombrements inclus en trop à cause des doublons est estimée à 3,00 % de la population cible du recensement. Le taux de sous-dénombrement de la population est évalué à 4,98 % (1 897 876 personnes), tandis que le taux de surdénombrement de la population est évalué à 1,98 % (755 635 personnes). Un taux de sous-dénombrement de 4,98 % signifie que les personnes non incluses faisant partie de la population cible représentent 4,98 % de la population cible du recensement. Un taux de surdénombrement de 1,98 % signifie que les doublons représentent 1,98 % de la population visée par le recensement.

Le taux de sous-dénombrement estimé a augmenté par rapport au Recensement de 2016, alors que le taux de surdénombrement est demeuré relativement stable. Par conséquent, le taux de sous-dénombrement net, qui est la différence entre le sous-dénombrement et le surdénombrement, a augmenté par rapport au Recensement de 2016.

La méthodologie des études sur la couverture du Recensement de 2021, telle que décrite aux sections [7](#) et [8](#), a été améliorée puisque le but premier de ces études est toujours de produire les meilleures estimations de couverture possible pour le plus récent recensement. Ces améliorations limitent toutefois la comparabilité de ces estimations avec celles du sous-dénombrement net de 2016 et expliquent en partie la différence observée. De plus, comme les estimations du sous-dénombrement net sont mesurées à partir d'échantillons, la marge d'erreur associée aux estimations de 2016 et de 2021 peut aussi expliquer en partie la différence.

2. Les taux de sous-dénombrement net dans le présent rapport peuvent légèrement différer des taux diffusés le 27 septembre 2023, car des réserves et des établissements partiellement dénombrés sont inclus. Toutes les estimations d'erreurs de couverture comprises dans le présent document font abstraction des erreurs de couverture visant ce groupe.

Tableau 1.2a

Estimations du taux d'erreur de couverture de la population et de l'erreur-type pour le Canada, recensements de 2016 et de 2021

Erreur de couverture	Recensement de 2016		Recensement de 2021	
	Taux estimé	Erreur-type	Taux estimé	Erreur-type
	pourcentage			
Sous-dénombrement	4,32	0,11	4,98	0,09
Surdénombrement	1,96	0,04	1,98	0,02
Sous-dénombrement net	2,36	0,12	3,00	0,10

Sources : Statistique Canada, études sur la couverture des recensements de 2016 et de 2021.

La présente section fait état des estimations du sous-dénombrement net établies en fonction des variables géographiques et démographiques suivantes :

- [province ou territoire](#) de résidence au moment du recensement;
- [âge](#) et [genre](#);
- [état matrimonial](#) et [genre](#);
- [langue maternelle](#);
- [région métropolitaine de recensement \(RMR\)](#) du [lieu habituel de résidence](#) le jour du recensement.

Le tableau 1.2b présente une estimation du sous-dénombrement net, des erreurs-types liées à l'estimation ainsi que du taux de sous-dénombrement net correspondant et son erreur-type selon diverses caractéristiques. Des estimations négatives de sous-dénombrement net indiquent un taux de surdénombrement supérieur au taux de sous-dénombrement. La [section 9](#) explique de quelle façon une telle situation peut se produire.

Tableau 1.2b

Estimations du sous-dénombrement net de la population et de l'erreur-type pour diverses caractéristiques, Recensement de 2021

Caractéristique	Sous-dénombrement net de la population		Taux de sous-dénombrement net de la population	
	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)
Canada	1 142 241	38 241	3,00	0,10
Province ou territoire				
Terre-Neuve-et-Labrador	16 235	1 829	3,08	0,34
Île-du-Prince-Édouard	6 901	995	4,28	0,59
Nouvelle-Écosse	27 852	3 452	2,79	0,34
Nouveau-Brunswick	13 624	2 731	1,73	0,34
Québec	42 868	16 829	0,50	0,20
Ontario	585 370	28 636	3,95	0,19
Manitoba	39 284	5 145	2,84	0,36
Saskatchewan	34 912	4 464	2,99	0,37
Alberta	153 166	12 527	3,47	0,27
Colombie-Britannique	212 846	16 069	4,08	0,30
Yukon	2 467	174	5,78	0,38
Territoires du Nord-Ouest	3 589	228	8,04	0,47
Nunavut	3 128	250	7,82	0,58

Rapport technique sur la couverture

Tableau 1.2b

Estimations du sous-dénombrement net de la population et de l'erreur-type pour diverses caractéristiques, Recensement de 2021

Caractéristique	Sous-dénombrement net de la population		Taux de sous-dénombrement net de la population	
	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)
Genre et groupe d'âge				
Total, genre	1 142 241	38 241	3,00	0,10
0 à 4 ans	65 492	6 028	3,45	0,31
5 à 14 ans	5 808	13 172	0,14	0,31
15 à 17 ans	-2 276	8 846	-0,19	0,73
18 à 19 ans	18 027	9 192	2,20	1,10
20 à 24 ans	228 413	18 752	9,40	0,70
25 à 34 ans	419 936	23 354	7,83	0,40
35 à 44 ans	217 802	19 605	4,25	0,37
45 à 54 ans	129 102	19 430	2,69	0,39
55 à 64 ans	84 286	18 793	1,59	0,35
65 à 74 ans	-10 341	13 361	-0,26	0,33
75 ans et plus	-14 008	12 096	-0,48	0,41
Hommes+¹	758 934	34 974	4,00	0,18
0 à 4 ans	30 137	5 776	3,11	0,58
5 à 14 ans	7 324	9 109	0,34	0,42
15 à 17 ans	-1 177	6 847	-0,19	1,10
18 à 19 ans	9 043	6 155	2,15	1,43
20 à 24 ans	118 670	13 585	9,44	0,98
25 à 34 ans	272 115	19 190	9,88	0,63
35 à 44 ans	158 860	16 321	6,17	0,60
45 à 54 ans	85 976	15 141	3,62	0,61
55 à 64 ans	76 278	14 916	2,90	0,55
65 à 74 ans	8 802	10 430	0,45	0,53
75 ans et plus	-7 095	7 861	-0,56	0,62
Femmes+²	383 307	29 525	2,00	0,15
0 à 4 ans	35 355	5 729	3,81	0,59
5 à 14 ans	-1 516	9 625	-0,07	0,47
15 à 17 ans	-1 099	5 878	-0,19	1,00
18 à 19 ans	8 983	6 920	2,26	1,70
20 à 24 ans	109 743	14 487	9,35	1,12
25 à 34 ans	147 821	15 252	5,67	0,55
35 à 44 ans	58 942	10 408	2,31	0,40
45 à 54 ans	43 126	12 911	1,78	0,52
55 à 64 ans	8 007	11 809	0,30	0,44
65 à 74 ans	-19 143	9 139	-0,91	0,44
75 ans et plus	-6 913	9 066	-0,41	0,54

Rapport technique sur la couverture

Tableau 1.2b

Estimations du sous-dénombrement net de la population et de l'erreur-type pour diverses caractéristiques, Recensement de 2021

Caractéristique	Sous-dénombrement net de la population		Taux de sous-dénombrement net de la population	
	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)
État matrimonial et genre pour les personnes âgées de 15 ans et plus				
Total, genre	1 070 941	37 309	3,34	0,11
Marié(e) et non séparé(e)	104 490	24 944	0,76	0,18
Vivant en union libre	133 518	18 195	3,31	0,44
Célibataire (jamais légalement marié[e])	654 481	32 073	6,76	0,31
Séparé(e)	95 166	12 460	11,38	1,32
Divorcé(e)	70 158	15 392	3,52	0,75
Veuf(ve)	13 128	12 122	0,78	0,72
Hommes+¹	721 473	33 952	4,55	0,20
Marié et non séparé	76 432	18 026	1,10	0,26
Vivant en union libre	100 598	14 643	4,90	0,68
Célibataire (jamais légalement marié)	421 561	25 951	7,95	0,45
Séparé	65 379	10 756	17,01	2,32
Divorcé	50 518	11 062	6,23	1,28
Veuf	6 984	7 661	1,88	2,02
Femmes+²	349 468	28 186	2,16	0,17
Mariée et non séparée	28 058	17 248	0,41	0,25
Vivant en union libre	32 919	10 219	1,66	0,51
Célibataire (jamais légalement mariée)	232 920	22 326	5,32	0,48
Séparée	29 787	6 308	6,59	1,30
Divorcée	19 640	10 867	1,66	0,90
Veuve	6 144	9 498	0,47	0,72
Langue maternelle				
Total	1 142 241	38 241	3,00	0,10
Anglais	523 481	31 784	2,51	0,15
Français	23 971	16 633	0,33	0,23
Langue non officielle	530 185	28 328	6,23	0,31
Français et anglais	10 608	5 006	3,48	1,59
Anglais et langue non officielle	41 970	9 168	4,34	0,91
Français et langue non officielle	8 609	3 563	6,78	2,62
Français, anglais et langue non officielle	3 416	1 854	6,05	3,08
Région métropolitaine de recensement				
St. John's	6 222	1 651	2,84	0,73
Halifax	9 726	2 694	2,05	0,56
Moncton	885	1 530	0,56	0,96
Saint John	2 078	1 573	1,57	1,17
Fredericton	1 362	1 380	1,24	1,24
Saguenay	3 388	3 444	2,05	2,05
Québec	-4 037	5 843	-0,48	0,70

Rapport technique sur la couverture

Tableau 1.2b

Estimations du sous-dénombrement net de la population et de l'erreur-type pour diverses caractéristiques, Recensement de 2021

Caractéristique	Sous-dénombrement net de la population		Taux de sous-dénombrement net de la population	
	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)
Sherbrooke	-5 760	2 928	-2,60	1,36
Trois-Rivières	7 930	4 134	4,68	2,33
Drummondville	-949	1 647	-0,94	1,65
Montréal	21 241	13 439	0,49	0,31
Ottawa–Gatineau	40 654	12 023	2,66	0,77
Kingston	4 402	4 764	2,49	2,63
Belleville–Quinte West	1 686	3 373	1,49	2,94
Peterborough	5 257	3 743	3,93	2,69
Oshawa	13 117	6 958	3,06	1,57
Toronto	329 366	26 924	5,04	0,39
Hamilton	30 219	9 093	3,71	1,07
St. Catharines–Niagara	7 349	7 644	1,67	1,70
Kitchener–Cambridge–Waterloo	11 206	6 768	1,91	1,13
Brantford	5 233	3 765	3,50	2,43
Guelph	-1 191	2 771	-0,72	1,70
London	14 807	7 891	2,65	1,38
Windsor	25 248	7 020	5,64	1,48
Barrie	6 320	4 883	2,88	2,16
Grand Sudbury	5 306	4 455	3,02	2,46
Thunder Bay	4 407	4 787	3,45	3,62
Winnipeg	17 674	4 430	2,07	0,51
Regina	3 929	2 620	1,55	1,02
Saskatoon	10 197	3 286	3,11	0,97
Lethbridge	2 127	2 612	1,69	2,04
Calgary	56 836	8 740	3,69	0,55
Red Deer	4 515	3 642	4,29	3,31
Edmonton	35 376	8 761	2,43	0,59
Kelowna	9 470	4 694	4,09	1,94
Kamloops	8 195	3 763	6,70	2,87
Chilliwack	1 076	2 534	0,94	2,19
Abbotsford–Mission	1 237	3 183	0,63	1,61
Vancouver	101 974	14 437	3,72	0,51
Victoria	20 508	6 768	4,91	1,54
Nanaimo	-966	2 172	-0,84	1,91
Toutes les régions métropolitaines de recensement	817 622	37 380	2,91	0,13
À l'extérieur d'une région métropolitaine de recensement	324 618	26 293	3,23	0,25

1. Le terme « hommes+ » comprend les hommes (et les garçons) et certaines personnes non binaires.

2. Le terme « femmes+ » comprend les femmes (et les filles) et certaines personnes non binaires.

Note : En raison de l'arrondissement, la somme des estimations de la couverture pourrait ne pas correspondre au total.

Sources : Statistique Canada, Recensement de 2021 et études sur la couverture du Recensement de 2021.

L'erreur-type donne une idée de l'exactitude des estimations reposant sur l'échantillonnage. Un intervalle couvrant deux erreurs-types de part et d'autre de l'estimation inclut la vraie valeur 19 fois sur 20. Autrement dit, il existe environ 19 chances sur 20 (95 %) que le taux de sous-dénombrement net réel de la population lors du Recensement de 2021 se situe entre 2,8 % et 3,2 % (c.-à-d. 3,0 % \pm deux erreurs-types) ou environ deux chances sur trois (68 %) que le taux réel se situe entre 2,9 % et 3,1 % (c.-à-d. 3,0 % \pm une erreur-type).

Étant donné que le sous-dénombrement net tient compte du sous-dénombrement et du surdénombrement, le lecteur devrait également consulter les estimations du sous-dénombrement et du surdénombrement présentées au [tableau 1.3](#). Un taux de sous-dénombrement net peu élevé pourrait notamment indiquer un taux de sous-dénombrement peu élevé ou un taux élevé combiné à un taux de surdénombrement élevé.

Les taux de sous-dénombrement net de la population les plus élevés dans le pays ont été observés dans les trois territoires : les Territoires du Nord-Ouest et le Nunavut ont enregistré les taux les plus élevés (respectivement 8,04 % et 7,82 %), suivis du Yukon (5,78 %). Parmi les provinces, l'Île-du-Prince-Édouard a enregistré le taux de sous-dénombrement net le plus élevé (4,28 %), suivi par la Colombie-Britannique (4,08 %) et l'Ontario (3,95 %). Le Québec a enregistré le taux de sous-dénombrement net de la population le plus bas (0,50 %), suivi par le Nouveau-Brunswick (1,73 %). Toutes les autres provinces ont enregistré un taux de sous-dénombrement net d'au moins un point de pourcentage supérieur à celui du Nouveau-Brunswick. En 2016, le taux le plus élevé et le plus bas ont également été observés pour les Territoires du Nord-Ouest et le Québec.

À partir du Recensement de 2021, les estimations de l'erreur de couverture du recensement sont rapportées par genre plutôt que par sexe. Le genre fait référence à l'identité personnelle et sociale d'une personne en tant qu'homme, femme ou personne non binaire (personne n'étant pas exclusivement homme ni femme). Compte tenu de la faible taille de la population non binaire, une agrégation des données dans une variable de genre à deux catégories a été utilisée pour présenter les estimations dans le présent rapport technique. Les personnes appartenant à la catégorie « personnes non binaires » ont été réparties dans les deux autres catégories de genre dénotées par le symbole « + ». Tout au long de ce rapport, la catégorie « hommes+ » comprend les hommes (et les garçons) ainsi que certaines personnes non binaires, tandis que la catégorie « femmes+ » comprend les femmes (et filles), ainsi que certaines personnes non binaires.

Le sous-dénombrement net de la population était généralement plus élevé pour les hommes+ et a atteint un sommet chez les jeunes adultes. Dans la population générale, le taux de sous-dénombrement net le plus élevé a été observé chez les adultes âgés de 20 à 34 ans à la fois pour les hommes+ et les femmes+. Il s'élevait à 9,40 % chez les personnes de 20 à 24 ans et à 7,83 % chez celles de 25 à 34 ans. Pour la catégorie hommes+, les taux de sous-dénombrement net les plus élevés étaient de 9,88 % pour les hommes âgés de 25 à 34 ans et de 9,44 % pour les hommes âgés de 20 à 24 ans. Pour la catégorie femmes+, ces taux ont atteint 9,35 % des femmes âgées de 20 à 24 ans et 5,67 % de celles âgées de 25 à 34 ans. Dans le cas des femmes+ âgées de 65 ans et plus et des hommes+ âgés de 75 ans et plus, le taux de sous-dénombrement net était négatif; ce qui indique davantage de dénombremments en trop que de personnes non dénombrées.

Le taux de sous-dénombrement net de la population âgée de 15 ans et plus était élevé dans le cas des personnes séparées (11,38 %), surtout chez les hommes+ (17,01 %). Il était également élevé dans le cas des personnes célibataires (6,76 %); c.-à-d. des personnes qui n'ont jamais été mariées légalement et qui ne vivent pas en union libre.

Le taux de sous-dénombrement net des personnes dont la langue maternelle était uniquement l'anglais était supérieur à celui des personnes dont la langue maternelle était uniquement le français (2,51 % comparativement à 0,33 %). En excluant les personnes ayant déclaré plusieurs langues maternelles, le taux de sous-dénombrement net pour les allophones (personnes dont la langue maternelle n'est ni l'anglais ni le français) était supérieur (6,23 %).

Le sous-dénombrement net de la population était légèrement plus élevé hors des RMR qu'au sein de celles-ci. Au niveau national, le sous-dénombrement net s'élevait à 3,23 % pour les personnes qui auraient dû être dénombrées hors d'une RMR. Ce taux était légèrement supérieur au taux de sous-dénombrement net des personnes vivant dans une RMR (2,91 %).

1.3 Sous-dénombrement

Un sous-dénombrement désigne généralement des personnes n'ayant pas été dénombrées à titre de résidents habituels dans le questionnaire rempli pour leur lieu habituel de résidence, ou encore des personnes pour lesquelles aucun questionnaire n'a été rempli pour leur lieu habituel de résidence. Ainsi, les personnes qui considèrent leur lieu de résidence comme un lieu de résidence temporaire n'ont peut-être pas été incluses ailleurs comme résidents habituels. Les personnes n'ayant pas de lieu habituel de résidence, les « personnes en situation d'itinérance » par exemple, sont également beaucoup plus sujettes au sous-dénombrement.

La présente section fait état des estimations du sous-dénombrement établies en fonction des variables géographiques et démographiques suivantes :

- [province ou territoire](#) de résidence au moment du recensement;
- [âge](#) et [genre](#);
- [état matrimonial](#) et [genre](#);
- [langue maternelle](#);
- [RMR](#) du [lieu habituel de résidence](#) le jour du recensement.

Le tableau 1.3 fournit une estimation du sous-dénombrement concernant le nombre de personnes non incluses, l'erreur-type liée à l'estimation, le taux de sous-dénombrement estimé correspondant et l'erreur-type connexe.

Tableau 1.3
Estimations du sous-dénombrement et du surdénombrement de la population et de l'erreur-type pour diverses caractéristiques, Recensement de 2021

Caractéristique	Sous-dénombrement de la population				Surdénombrement de la population			
	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)
Canada	1 897 876	37 004	4,98	0,09	755 635	9 648	1,98	0,02
Province ou territoire								
Terre-Neuve-et-Labrador	26 428	1 775	5,02	0,32	10 194	439	1,94	0,08
Île-du-Prince-Édouard	10 190	977	6,32	0,57	3 289	191	2,04	0,12
Nouvelle-Écosse	47 196	3 373	4,73	0,32	19 344	736	1,94	0,07
Nouveau-Brunswick	30 165	2 655	3,82	0,32	16 541	641	2,10	0,08
Québec	225 381	15 756	2,64	0,18	182 513	5 915	2,14	0,07
Ontario	855 470	27 795	5,78	0,18	270 100	6 888	1,82	0,05
Manitoba	62 443	5 089	4,52	0,35	23 160	757	1,68	0,05
Saskatchewan	59 075	4 411	5,06	0,36	24 163	689	2,07	0,06
Alberta	231 249	12 225	5,24	0,26	78 084	2 736	1,77	0,06
Colombie-Britannique	338 678	15 827	6,50	0,28	125 832	2 778	2,41	0,05
Yukon	3 294	170	7,71	0,37	827	38	1,94	0,09
Territoires du Nord-Ouest	4 426	227	9,91	0,46	837	15	1,87	0,03
Nunavut	3 879	250	9,70	0,56	751	17	1,88	0,04
Genre et groupe d'âge								
Total, genre	1 897 876	37 004	4,98	0,09	755 635	9 648	1,98	0,02
0 à 4 ans	98 438	5 809	5,19	0,29	32 946	1 610	1,74	0,08
5 à 14 ans	140 303	12 630	3,35	0,29	134 496	3 737	3,21	0,09
15 à 17 ans	41 091	8 516	3,40	0,68	43 367	2 393	3,58	0,19

Rapport technique sur la couverture

Tableau 1.3

Estimations du sous-dénombrement et du surdénombrement de la population et de l'erreur-type pour diverses caractéristiques, Recensement de 2021

Caractéristique	Sous-dénombrement de la population				Surdénombrement de la population			
	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)
18 à 19 ans	48 426	8 933	5,91	1,03	30 399	2 168	3,71	0,26
20 à 24 ans	313 912	18 381	12,91	0,66	85 499	3 713	3,52	0,15
25 à 34 ans	530 315	22 964	9,89	0,39	110 379	4 251	2,06	0,08
35 à 44 ans	281 659	19 332	5,49	0,36	63 857	3 261	1,25	0,06
45 à 54 ans	192 317	19 130	4,01	0,38	63 215	3 398	1,32	0,07
55 à 64 ans	164 211	18 459	3,10	0,34	79 926	3 530	1,51	0,07
65 à 74 ans	49 955	13 112	1,23	0,32	60 296	2 567	1,49	0,06
75 ans et plus	37 248	11 803	1,26	0,40	51 256	2 646	1,74	0,09
Hommes+¹	1 140 736	34 194	6,01	0,17	381 802	7 344	2,01	0,04
0 à 4 ans	46 527	5 668	4,80	0,56	16 390	1 113	1,69	0,11
5 à 14 ans	76 859	8 684	3,57	0,39	69 535	2 750	3,23	0,12
15 à 17 ans	22 004	6 595	3,53	1,02	23 180	1 840	3,72	0,29
18 à 19 ans	23 968	5 962	5,70	1,34	14 925	1 530	3,55	0,35
20 à 24 ans	163 089	13 315	12,97	0,92	44 420	2 698	3,53	0,21
25 à 34 ans	330 497	18 903	12,00	0,60	58 381	3 306	2,12	0,12
35 à 44 ans	192 459	16 057	7,48	0,58	33 599	2 925	1,31	0,11
45 à 54 ans	116 139	15 044	4,89	0,60	30 164	1 710	1,27	0,07
55 à 64 ans	116 176	14 720	4,42	0,54	39 897	2 407	1,52	0,09
65 à 74 ans	37 494	10 291	1,92	0,52	28 692	1 699	1,47	0,09
75 ans et plus	15 524	7 677	1,22	0,60	22 619	1 694	1,78	0,13
Femmes+²	757 140	28 656	3,95	0,14	373 833	7 113	1,95	0,04
0 à 4 ans	51 911	5 609	5,60	0,57	16 556	1 164	1,78	0,12
5 à 14 ans	63 444	9 289	3,12	0,44	64 961	2 524	3,20	0,12
15 à 17 ans	19 088	5 677	3,26	0,94	20 186	1 527	3,44	0,25
18 à 19 ans	24 458	6 747	6,14	1,59	15 475	1 539	3,89	0,38
20 à 24 ans	150 822	14 260	12,86	1,06	41 079	2 554	3,50	0,21
25 à 34 ans	199 818	15 010	7,66	0,53	51 998	2 705	1,99	0,10
35 à 44 ans	89 199	10 306	3,49	0,39	30 257	1 448	1,18	0,06
45 à 54 ans	76 178	12 572	3,14	0,50	33 052	2 938	1,36	0,12
55 à 64 ans	48 036	11 524	1,80	0,42	40 029	2 577	1,50	0,10
65 à 74 ans	12 462	8 934	0,60	0,42	31 605	1 928	1,51	0,09
75 ans et plus	21 724	8 843	1,30	0,52	28 637	1 997	1,71	0,12

Rapport technique sur la couverture

Tableau 1.3

Estimations du sous-dénombrement et du surdénombrement de la population et de l'erreur-type pour diverses caractéristiques, Recensement de 2021

Caractéristique	Sous-dénombrement de la population				Surdénombrement de la population			
	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)
État matrimonial et genre pour les personnes âgées de 15 ans et plus								
Total, genre	1 659 135	36 154	5,18	0,11	588 194	9 210	1,84	0,03
Marié(e) et non séparé(e)	296 693	24 616	2,15	0,17	192 203	4 032	1,39	0,03
Vivant en union libre	194 311	18 002	4,82	0,42	60 793	2 640	1,51	0,06
Célibataire (jamais légalement marié[e])	917 850	31 336	9,48	0,29	263 369	6 836	2,72	0,07
Séparé(e)	106 837	12 428	12,77	1,30	11 671	897	1,40	0,11
Divorcé(e)	100 367	15 200	5,04	0,72	30 209	2 425	1,52	0,12
Veuf(ve)	43 077	11 956	2,57	0,69	29 949	2 002	1,79	0,12
Hommes+¹	1 017 350	33 273	6,41	0,20	295 877	6 754	1,87	0,04
Marié et non séparé	174 151	17 794	2,51	0,25	97 718	2 884	1,41	0,04
Vivant en union libre	130 893	14 521	6,37	0,66	30 295	1 884	1,47	0,09
Célibataire (jamais légalement marié)	565 335	25 393	10,66	0,43	143 774	5 350	2,71	0,10
Séparé	70 567	10 737	18,36	2,28	5 188	638	1,35	0,17
Divorcé	61 223	11 009	7,55	1,26	10 704	1 075	1,32	0,13
Veuf	15 182	7 576	4,08	1,95	8 198	1 137	2,20	0,30
Femmes+²	641 785	27 411	3,96	0,16	292 317	6 562	1,81	0,04
Mariée et non séparée	122 543	17 010	1,78	0,24	94 485	2 855	1,37	0,04
Vivant en union libre	63 417	10 048	3,20	0,49	30 498	1 860	1,54	0,09
Célibataire (jamais légalement mariée)	352 515	21 877	8,05	0,46	119 595	4 455	2,73	0,10
Séparée	36 270	6 277	8,03	1,28	6 483	631	1,43	0,14
Divorcée	39 145	10 647	3,31	0,87	19 505	2 175	1,65	0,18
Veuve	27 895	9 357	2,14	0,70	21 751	1 627	1,67	0,12
Langue maternelle								
Total	1 897 876	37 004	4,98	0,09	755 635	9 648	1,98	0,02
Anglais	935 492	31 088	4,48	0,14	412 011	6 619	1,97	0,03
Français	181 362	15 731	2,48	0,21	157 391	5 402	2,15	0,07
Langue non officielle	688 896	27 968	8,10	0,30	158 711	4 501	1,87	0,05
Français et anglais	18 386	4 933	6,04	1,52	7 778	856	2,55	0,28
Anglais et langue non officielle	58 806	9 125	6,08	0,89	16 836	894	1,74	0,09
Français et langue non officielle	10 535	3 550	8,30	2,57	1 925	302	1,52	0,24
Français, anglais et langue non officielle	4 400	1 843	7,79	3,01	983	201	1,74	0,35

Rapport technique sur la couverture

Tableau 1.3

Estimations du sous-dénombrement et du surdénombrement de la population et de l'erreur-type pour diverses caractéristiques, Recensement de 2021

Caractéristique	Sous-dénombrement de la population				Surdénombrement de la population			
	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)
Région métropolitaine de recensement								
St. John's	11 044	1 618	5,05	0,70	4 822	329	2,20	0,15
Halifax	18 113	2 639	3,81	0,53	8 387	541	1,76	0,11
Moncton	3 694	1 512	2,33	0,93	2 809	229	1,77	0,14
Saint John	4 595	1 560	3,46	1,13	2 516	198	1,90	0,15
Fredericton	3 813	1 350	3,47	1,19	2 452	284	2,23	0,25
Saguenay	6 799	3 353	4,12	1,95	3 411	790	2,07	0,47
Québec	15 048	5 500	1,80	0,65	19 084	1 970	2,28	0,23
Sherbrooke	1 031	2 718	0,47	1,22	6 791	1 088	3,06	0,48
Trois-Rivières	11 233	4 067	6,63	2,24	3 304	741	1,95	0,43
Drummondville	1 096	1 539	1,09	1,51	2 044	587	2,03	0,57
Montréal	108 911	12 869	2,53	0,29	87 670	3 872	2,03	0,09
Ottawa–Gatineau	67 713	11 882	4,43	0,74	27 060	1 836	1,77	0,12
Kingston	7 590	4 724	4,29	2,56	3 188	611	1,80	0,34
Belleville–Quinte West	3 807	3 324	3,37	2,85	2 120	573	1,88	0,50
Peterborough	7 555	3 718	5,64	2,62	2 298	435	1,72	0,32
Oshawa	20 929	6 892	4,89	1,53	7 812	956	1,82	0,22
Toronto	455 717	26 436	6,98	0,38	126 350	5 102	1,93	0,08
Hamilton	42 294	9 009	5,19	1,05	12 075	1 234	1,48	0,15
St. Catharines–Niagara	15 168	7 573	3,44	1,66	7 819	1 043	1,77	0,23
Kitchener–Cambridge–Waterloo	20 172	6 693	3,44	1,10	8 966	1 006	1,53	0,17
Brantford	7 767	3 667	5,20	2,33	2 534	856	1,70	0,57
Guelph	2 188	2 708	1,33	1,63	3 378	586	2,05	0,35
London	24 481	7 798	4,38	1,34	9 674	1 209	1,73	0,21
Windsor	32 144	6 956	7,18	1,44	6 896	943	1,54	0,21
Barrie	10 079	4 847	4,60	2,11	3 759	597	1,72	0,27
Grand Sudbury	9 323	4 398	5,30	2,37	4 017	711	2,28	0,40
Thunder Bay	7 022	4 755	5,50	3,52	2 615	553	2,05	0,43
Winnipeg	29 805	4 401	3,50	0,50	12 131	510	1,42	0,06
Regina	9 078	2 593	3,59	0,99	5 150	378	2,03	0,15
Saskatoon	16 775	3 266	5,12	0,95	6 578	357	2,01	0,11
Lethbridge	4 581	2 584	3,64	1,98	2 454	383	1,95	0,30
Calgary	83 937	8 636	5,46	0,53	27 101	1 344	1,76	0,09
Red Deer	7 107	2 650	6,75	2,35	2 592	2 498	2,46	2,31
Edmonton	59 952	8 696	4,12	0,57	24 575	1 064	1,69	0,07
Kelowna	16 229	4 625	7,01	1,86	6 760	802	2,92	0,34
Kamloops	10 217	3 752	8,35	2,81	2 022	293	1,65	0,24
Chilliwack	3 603	2 502	3,14	2,11	2 527	404	2,20	0,35
Abbotsford–Mission	6 273	3 148	3,18	1,55	5 036	472	2,56	0,24

Tableau 1.3

Estimations du sous-dénombrement et du surdénombrement de la population et de l'erreur-type pour diverses caractéristiques, Recensement de 2021

Caractéristique	Sous-dénombrement de la population				Surdénombrement de la population			
	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)
Vancouver	170 741	14 302	6,22	0,49	68 767	1 973	2,51	0,07
Victoria	31 509	6 711	7,54	1,49	11 001	876	2,63	0,21
Nanaimo	2 916	2 053	2,55	1,75	3 881	707	3,39	0,60
Toutes les régions métropolitaines de recensement	1 372 046	36 333	4,88	0,12	554 424	8 787	1,97	0,03
À l'extérieur d'une région métropolitaine de recensement	525 830	25 907	5,24	0,24	201 212	4 492	2,01	0,04

1. Le terme « hommes+ » comprend les hommes (et les garçons) et certaines personnes non binaires.

2. Le terme « femmes+ » comprend les femmes (et les filles) et certaines personnes non binaires.

Note : En raison de l'arrondissement, la somme des estimations de la couverture pourrait ne pas correspondre au total.

Sources : Statistique Canada, Recensement de 2021 et études sur la couverture du Recensement de 2021.

On dénote certaines tendances démographiques en ce qui a trait au sous-dénombrement. Le taux de sous-dénombrement pour les hommes+ était une fois et demie supérieur à celui des femmes+ (6,01 % par rapport à 3,95 %). Le taux de sous-dénombrement le plus élevé était celui des jeunes adultes âgés de 18 à 34 ans (hommes+ et femmes+). Parmi les jeunes hommes+, le sous-dénombrement était de 12,97 % pour ceux âgés de 20 à 24 ans et de 12,00 % pour ceux âgés de 25 à 34 ans, mais il était plus élevé pour ceux âgés de 35 à 44 ans que pour ceux âgés de 18 à 19 ans, contrairement à ce qui a été habituellement observé par le passé.

Du côté de l'état matrimonial, le taux de sous-dénombrement était le plus élevé pour les personnes âgées de 15 ans ou plus séparées et qui ne vivaient pas en union libre, s'élevant à 12,77 %. Il était également élevé pour les personnes célibataires ne vivant pas en union libre, se situant à 9,48 %. Dans les deux cas, ces taux étaient plus élevés chez les hommes+ que chez les femmes+.

Le [tableau 1.3](#) met en évidence le profil des personnes les plus susceptibles d'avoir été omises lors du Recensement de 2021; c.-à-d. les hommes+ de 20 à 34 ans célibataires (n'ayant jamais été mariés et ne vivant pas en union libre) et les personnes séparées. La langue maternelle a également été significative pour les taux de sous-dénombrement; en effet, pour les personnes n'ayant déclaré qu'une langue maternelle, les taux de sous-dénombrement étaient inférieurs pour celles dont la langue maternelle était le français (2,48 %), suivi de celles dont la langue maternelle était l'anglais (4,48 %). Chez les personnes pour lesquelles la seule langue maternelle n'était ni le français ni l'anglais, le taux était encore plus élevé : 8,10 %.

1.4 Surdénombrement

Le surdénombrement de la population désigne le nombre de dénombremments en trop dans les chiffres du recensement pour les personnes dénombrées plus d'une fois (généralement deux fois). Cette erreur introduit un biais, étant donné que ces personnes auraient dû être dénombrées une seule fois. Parmi les exemples de surdénombrement figurent le cas des enfants dont les parents ne vivent pas ensemble et qui sont inclus dans le questionnaire de recensement de chacun des parents, celui des personnes qui s'absentent de la maison pour des raisons professionnelles et dont le nom figure dans le questionnaire de la famille et dans le questionnaire du logement où elles vivent pendant qu'elles travaillent ainsi que celui des étudiants qui suivent leurs études et qui sont inclus à la fois dans le questionnaire du recensement de leurs colocataires et dans celui de leurs parents.

La présente section fait état des estimations du surdénombrement établies en fonction des variables géographiques et démographiques suivantes :

- [province ou territoire](#) de résidence au moment du recensement;
- [âge](#) et [genre](#);
- [état matrimonial](#) et [genre](#);
- [langue maternelle](#);
- [RMR](#) du [lieu habituel de résidence](#) le jour du recensement.

Le [tableau 1.3](#) contient également une estimation du nombre de dénombremments en trop ainsi que l'estimation de l'erreur-type qui y est associée.

Les erreurs-types des estimations du surdénombrement figurant au [tableau 1.3](#) étaient moins élevées que celles du sous-dénombrement. Comme indiqué ci-dessous, certaines tendances démographiques se sont dégagées des estimations du surdénombrement.

Le taux estimé de surdénombrement variait moins que le taux de sous-dénombrement, et ce, pour l'ensemble des provinces et des territoires. En effet, l'écart entre le taux le plus bas et le taux le plus élevé parmi l'ensemble des provinces et des territoires était de 0,73 point de pourcentage dans le cas du surdénombrement (taux variant entre 1,68 % et 2,41 %) alors que cet écart était de 7,27 points de pourcentage pour le sous-dénombrement (taux variant entre 2,64 % et 9,91 %). Les catégories hommes+ et femmes+ enregistraient des taux de surdénombrement similaires : 2,01 % par rapport à 1,95 %. Le surdénombrement était supérieur pour les enfants et les jeunes adultes âgés de 5 à 24 ans; ces taux variant de 3,21 % à 3,71 % pour ces groupes d'âge. Comme mentionné, ces taux élevés étaient attribuables en grande partie au dénombrement multiple d'enfants en garde partagée et d'étudiants.

Pour l'état matrimonial, le taux de surdénombrement était plus élevé pour les personnes n'ayant jamais été mariées et ne vivant pas en union libre (2,72 %). Ce phénomène s'observait autant chez les hommes+ que chez les femmes+ et reflétait les résultats selon l'âge.

En résumé, le [tableau 1.3](#) dresse le profil des personnes les plus susceptibles d'avoir été dénombrées plus d'une fois; il y avait autant de chances qu'il s'agisse d'hommes+ que de femmes+ et plus de chances qu'il s'agisse d'enfants et de jeunes adultes. Dans le cas des personnes âgées de 15 ans et plus, les célibataires étaient plus susceptibles d'être dénombrés plus d'une fois.

2. Univers du recensement

2.1 Introduction

Alors que le Recensement de 2021 permet d'obtenir des renseignements sur la population, les logements, les ménages et les familles, dans le cadre du Programme de mesure de l'erreur de couverture du Recensement de 2021, on estime les erreurs de couverture de l'univers de la population seulement. Cependant, les définitions des concepts de logement et les règles à suivre pour déterminer la liste des personnes qui devraient être dénombrées dans chaque logement ont une incidence sur la couverture de la population cible du recensement. Par conséquent, la présente section décrit les concepts de population et de logement. De plus, étant donné qu'une erreur de couverture peut être causée par une mauvaise interprétation du concept de lieu habituel de résidence, tel qu'il est défini dans les questionnaires du recensement, la section fournit également les renseignements figurant dans les questionnaires du recensement et la définition de [lieu habituel de résidence](#) du Recensement de 2021.

2.2 Univers de la population

La population cible du Recensement de 2021 comprend les groupes suivants :

- les citoyens canadiens et les immigrants reçus (résidents permanents) ayant un lieu habituel de résidence au Canada;
- les citoyens canadiens et les immigrants reçus (résidents permanents) qui sont à l'étranger, dans une base militaire ou en mission diplomatique;
- les citoyens canadiens et les immigrants reçus (résidents permanents) qui sont en mer ou dans des ports à bord de navires marchands immatriculés au Canada et de navires gouvernementaux canadiens;
- les résidents non permanents ayant un lieu habituel de résidence au Canada :
 - qui demandent le statut de réfugié (demandeurs d'asile);
 - qui sont titulaires d'un permis d'études (valide le jour du recensement);
 - qui sont titulaires d'un permis de travail (valide le jour du recensement);et les membres de leur famille vivant avec eux.

L'univers de la population du Recensement de 2021 ne comprend pas les résidents étrangers, mais, depuis 1991, il comprend les résidents non permanents.

La définition de la population cible précise quelles personnes doivent être dénombrées dans le recensement, mais non l'endroit où ces personnes devraient l'être. Le Recensement du Canada repose sur une méthode de dénombrement *de jure* modifiée, selon laquelle les personnes doivent être dénombrées à leur lieu habituel de résidence, même si elles sont temporairement absentes le jour du recensement. Les personnes qui ne sont pas à leur lieu habituel de résidence et qui résident ailleurs au Canada doivent être dénombrées à leur lieu habituel de résidence et sont considérées comme présentes, mais temporairement à l'autre endroit. Les personnes qui n'ont pas de lieu habituel de résidence doivent être dénombrées à l'endroit où elles se trouvent le jour du recensement. Certains pays utilisent la méthode de dénombrement *de facto*, selon laquelle toutes les personnes doivent être dénombrées à l'endroit où elles se trouvent le jour du recensement, quel que soit leur lieu habituel de résidence.

2.3 Univers des logements

Un [logement](#) est défini comme un ensemble de pièces d'habitation. Dans le cadre du recensement, on reconnaît deux genres de logements, soit les logements collectifs et les logements privés. Les études sur la couverture du recensement englobent ces deux genres de logements, sans distinction.

Rapport technique sur la couverture

Un [logement privé](#) s'entend d'un ensemble séparé de pièces d'habitation possédant une entrée privée soit à partir de l'extérieur de l'immeuble, soit à partir d'un hall, d'un foyer, d'un vestibule ou d'un escalier commun situé à l'intérieur de l'immeuble. Il faut qu'on puisse emprunter l'entrée menant au logement sans passer par les pièces d'habitation d'une autre personne ou d'un autre groupe de personnes.

Le logement doit répondre aux deux conditions qui le rendent propre à l'habitation durant toute l'année :

1. avoir une source de chauffage ou d'énergie (ce dont témoignent les cheminées, les lignes électriques, les conduites ou les compteurs d'huile ou de gaz, les génératrices, les cordes de bois, les ampoules électriques, les thermopompes, ou les panneaux solaires);
2. fournir un espace clos permettant de s'abriter des intempéries comme en témoigne la présence de murs d'enceinte et d'un toit ainsi que de portes et fenêtres qui protègent contre le vent, la pluie et la neige.

Les logements qui ne répondent pas aux conditions les rendant propres à l'habitation durant toute l'année sont des logements marginaux. Les logements privés sont classés parmi les logements privés ordinaires et les logements marginaux occupés. Les logements privés ordinaires se subdivisent à leur tour en trois grandes catégories : les logements occupés (par des résidents habituels), les logements occupés uniquement par des résidents étrangers ou par des personnes présentes temporairement, et les logements inoccupés. Les logements marginaux sont classés parmi les logements occupés par des résidents habituels ou parmi les logements occupés par des résidents étrangers ou par des personnes présentes temporairement. Les logements marginaux inoccupés le 11 mai 2021 ne font pas partie du parc immobilier.

Un [logement collectif](#) désigne un logement de nature commerciale, institutionnelle ou communautaire dans lequel une personne ou un groupe de personnes réside ou pourrait résider. Il doit fournir des soins ou des services ou disposer de certaines installations communes, telles qu'une cuisine ou une salle de bain, qui sont partagées par les occupants. Les logements collectifs comprennent, par exemple, les pensions et maisons de chambres, les hôtels, les motels, les établissements pour touristes, les établissements de soins infirmiers, les résidences pour personnes âgées, les hôpitaux, les résidences de personnel, les bases militaires, les campements de travailleurs, les établissements correctionnels et les foyers collectifs.

Les logements collectifs sont soit des logements occupés ou soit des logements inoccupés. Les logements occupés sont soit occupés par des résidents habituels, soit occupés uniquement par des résidents étrangers ou par des personnes présentes temporairement. Dans le cas des logements collectifs inoccupés, des données sur le logement, comme les types de services offerts, ont été recueillies, mais ne sont pas comprises dans les produits du recensement.

En bref, l'univers des logements inclut :

- les logements privés occupés par des résidents habituels
- les logements privés occupés uniquement par des résidents étrangers ou par des personnes présentes temporairement
- les logements privés inoccupés
- les logements marginaux occupés le jour du recensement
- les logements collectifs occupés par des résidents habituels
- les logements collectifs occupés uniquement par des résidents étrangers ou par des personnes présentes temporairement.

L'univers des logements n'inclut pas :

- les logements marginaux inoccupés le jour du recensement
- les logements collectifs inoccupés le jour du recensement
- les logements situés à l'extérieur du Canada.

2.4 Lieu habituel de résidence

La méthode de dénombrement *de jure* utilisée lors du Recensement de la population du Canada permet le dénombrement de la population en fonction du « [lieu habituel de résidence](#) », c'est-à-dire le logement où une personne vit habituellement. La plupart des personnes n'ont qu'une résidence, et il est facile de les dénombrer à leur lieu habituel de résidence. Le dénombrement consiste à dresser la liste de toutes les personnes ayant un logement précis comme lieu habituel de résidence le jour du recensement, en suivant les instructions étape par étape au début du questionnaire du recensement : « **En vous incluant**, combien de personnes vivent habituellement à cette adresse le 11 mai 2021? **Inclure** : toutes les personnes qui ont leur résidence principale à cette adresse, même si elles sont temporairement absentes. **Voir les instructions à la page 3** (garde partagée, étudiants, immigrants reçus, résidence secondaire, etc.) ». Les instructions de la page 3 du [questionnaire du Recensement de 2021](#) sont présentées à l'[annexe A](#).

Il est parfois difficile de déterminer le lieu habituel de résidence d'une personne. C'est la raison pour laquelle des règles spéciales ont été élaborées pour définir le lieu habituel de résidence dans certains cas :

1. Les personnes ayant plus d'une résidence

Cette catégorie englobe toutes les personnes ayant plusieurs logements au Canada qui pourraient être considérés comme leur lieu habituel de résidence. En pareil cas, le [lieu habituel de résidence](#) est l'endroit où la personne passe la majeure partie de l'année. Si elle passe autant de temps à chaque endroit ou si elle ne sait pas quel logement choisir, elle devrait choisir le lieu de résidence où elle a passé la nuit du 10 au 11 mai 2021. Cette règle prévoit deux exceptions :

- a. Les enfants qui habitent ailleurs à cause de leurs études ou de leur emploi d'été, mais qui reviennent vivre chez leurs parents pendant une partie de l'année devraient considérer le lieu de résidence qu'ils partagent avec leurs parents comme leur lieu habituel de résidence même s'ils passent la majeure partie de l'année ailleurs.
- b. Les conjoints ou les partenaires en union libre qui n'habitent pas avec leur famille à cause de leur travail ou de leurs études, mais qui reviennent à la maison périodiquement devraient considérer le lieu de résidence qu'ils partagent avec leur conjoint comme leur lieu habituel de résidence même s'ils passent la majeure partie de l'année ailleurs.

2. Les personnes vivant en établissement, notamment un hôpital, un foyer pour personnes âgées, une prison ou un établissement correctionnel

Les personnes ayant vécu dans un ou plusieurs établissements pendant une période continue de six mois ou plus au moment du recensement sont considérées comme des résidents habituels de l'établissement.

3. Les personnes n'ayant pas de lieu habituel de résidence

Les personnes qui n'ont pas de lieu habituel de résidence devraient être dénombrées à l'endroit où elles ont passé la nuit du 10 au 11 mai 2021.

4. Les personnes habitant à l'extérieur du Canada

Les citoyens canadiens et les immigrants reçus (résidents permanents) habitant à l'extérieur du Canada le jour du recensement comprennent notamment :

- les personnes à bord de navires canadiens ou de navires marchands;
- les employés du gouvernement fédéral et des gouvernements provinciaux ou territoriaux, ainsi que les membres de leur famille;
- les membres des Forces armées canadiennes et les membres de leur famille.

Ces personnes devraient indiquer dans le questionnaire du recensement l'adresse utilisée à des fins électorales ou leur dernière adresse permanente au Canada, si elles ne sont pas déjà prises en compte à la résidence de leur famille. Ces renseignements sont utilisés pour leur associer un lieu géographique au Canada aux fins de diffusion.

3. Erreur de couverture de la population

3.1 Sources

Bien que la collecte et le traitement des données du recensement doivent répondre à des normes de qualité élevées, il est très difficile d'éliminer toutes les erreurs potentielles. Il existe deux types d'erreurs de couverture de la population. Le sous-dénombrement de la population correspond à l'exclusion de personnes qui auraient dû être dénombrées, alors que le surdénombrement de la population désigne l'inclusion de personnes dénombrées plus d'une fois (en général deux fois). Le surdénombrement comprend aussi les personnes dénombrées qui n'auraient pas dû l'être. Toutefois, on ne mesure pas ce type d'erreur, car il est jugé négligeable.

Un sous-dénombrement peut se produire lors de la première étape du recensement si la liste des logements utilisée pour l'univers des logements est incomplète. Ce risque est plus élevé, par exemple, si un logement est en cours de construction. Inversement, un surdénombrement peut se produire si un logement est dénombré à deux reprises.

Il y a également des risques d'erreur de couverture à l'étape de la collecte des données sur le terrain. Les erreurs commises par les répondants sont à l'origine d'erreurs de couverture lorsque la personne qui remplit le questionnaire du recensement omet une personne dont le lieu habituel de résidence est le logement en question selon les règles du recensement; il s'agit alors d'un sous-dénombrement. La personne peut également inclure une personne dont le lieu habituel de résidence n'est pas le logement en question; il s'agit alors d'un surdénombrement si la personne a déjà été dénombrée à son lieu habituel de résidence ou à un autre endroit. Il est généralement facile de déterminer le lieu habituel de résidence d'une personne. Cependant, comme on le mentionne à la section précédente, le processus est parfois plus complexe, et des règles spéciales ont été élaborées pour déterminer le lieu habituel de résidence d'une personne. Ces règles sont énoncées dans le questionnaire du recensement, mais la liste est longue et elles peuvent être difficiles à comprendre. Des erreurs de couverture peuvent survenir lorsque les règles ne sont pas consultées ou lorsqu'elles sont appliquées de façon inappropriée. La notion de jour du recensement comme date de référence pour déterminer le lieu habituel de résidence peut également être mal comprise, ce qui peut entraîner une erreur de couverture.

Des erreurs de couverture peuvent également être commises à tout moment au cours de l'étape de traitement, lorsque les renseignements indiqués dans les dossiers des personnes ou des ménages sont entrés dans la base de données du recensement ou extraits de celle-ci. Les dossiers peuvent être supprimés par erreur. Les questionnaires peuvent être associés à un dossier erroné ou retournés trop tard pour être ajoutés.

Même si des efforts sont faits pour recenser la population des personnes en situation d'itinérance, le risque de sous-dénombrement est élevé. Certains autres modes de vie sont également susceptibles de donner lieu à des erreurs de couverture. Par exemple, les jeunes adultes qui viennent de quitter le domicile familial peuvent être soit sous-dénombrés, parce que ni leurs colocataires ni leurs parents ne les ont inscrits dans le questionnaire du recensement, soit surdénombrés, parce qu'ils ont été inscrits dans les deux questionnaires du recensement. Les personnes qui conservent une résidence secondaire en raison de leur emploi peuvent également être à l'origine d'une erreur de couverture.

Les utilisateurs doivent également savoir dans quelle mesure les réserves et les établissements ont participé au Recensement de 2021. Dans certains cas, le dénombrement n'a pas été autorisé par la communauté ou a été interrompu avant d'être terminé. Ces régions géographiques (63 en tout en 2021, en hausse par rapport aux 14 régions enregistrées en 2016) sont considérées comme des réserves et des établissements partiellement dénombrés. Il n'y a pas de données en 2021 pour les réserves et les établissements partiellement dénombrés, et ces régions ne sont pas prises en compte dans les totaux. Des problèmes semblables sont survenus au cours de recensements antérieurs. Par exemple, 22 réserves et établissements ont été partiellement dénombrés lors du Recensement de 2006, et 31 lors du Recensement de 2011.

Les estimations démographiques pour les 63 réserves et établissements partiellement dénombrés sont basées sur un modèle. Cependant, étant donné qu'il n'y a pas de source fiable permettant de vérifier les hypothèses du modèle, les estimations doivent être utilisées avec circonspection. Pour obtenir plus de renseignements, voir la [section 12.2](#).

3.2 Contrôle

Les sources possibles d'erreurs de couverture ont été définies au cours de l'étape de planification du Recensement de 2021, et les mesures suivantes ont été prises pour réduire au minimum les risques connexes :

- Les limites des unités de collecte (UC) ont été soigneusement définies et configurées, de façon à ce qu'aucune des régions géographiques ne soit omise ou incluse deux fois.
- Les secteurs de listage/livraison : Le manuel de l'agent recenseur renfermait des instructions quant à la façon de dénombrer une UC tout en minimisant les risques d'oublier des logements. Le nombre total de logements selon le Recensement de 2016 a été communiqué aux chargés des opérations sur le terrain afin de les aider à déceler les modifications notables. Par ailleurs, lorsque les opérations de listage donnaient lieu à une différence marquée du nombre de logements par rapport à celui du Recensement de 2016, une vérification de la liste était effectuée. Enfin, des procédures précises de contrôle qualitatif ont été appliquées à l'UC afin d'évaluer puis de corriger les modifications apportées à la liste.
- Les secteurs d'envoi par la poste : L'envoi par la poste a été effectué en fonction d'une liste d'adresses provenant du Registre des adresses de Statistique Canada. Cette liste était mise à jour régulièrement et des activités de listage ont été effectuées (sur le terrain et à distance), principalement dans les zones où les erreurs de base de sondage étaient plus probables. Les activités de listage ont été effectuées en continu, mais de façon plus intensive dans les deux années précédant le recensement. Le travail des agents recenseurs était étroitement surveillé.
- En 2021, la méthode d'envoi par la poste et de livraison à la porte (EPLP) a été mise en place. Les secteurs d'EPLP sont ceux où tous les logements ont une adresse et où la majorité d'entre eux peuvent recevoir du courrier. Dans ces secteurs mixtes, les logements disposant d'une adresse postale valide ont reçu le matériel d'envoi régulier par la poste (comme dans les secteurs d'envoi par la poste), tandis que ceux qui n'avaient pas d'adresse postale valide (correspondant à l'adresse municipale) ont reçu une lettre d'invitation qui a été déposée à leur porte par un employé du recensement. Les secteurs d'EPLP ont été mis en place pour maximiser le nombre de logements admissibles à l'envoi par la poste pour le recensement. Comme pour les secteurs d'envoi par la poste, les secteurs d'EPLP étaient basés sur une liste d'adresses provenant du Registre des adresses de Statistique Canada, et la liste était mise à jour régulièrement.
- Des procédures spéciales ont été définies pour assurer le dénombrement de la population vivant dans les réserves.
- Des messages publicitaires informaient les Canadiens à propos du recensement et indiquaient ce qu'ils devaient faire s'ils n'avaient pas reçu de questionnaire.
- L'Assistance téléphonique du recensement (ATR) était offerte pour obtenir une réponse aux questions à propos du recensement, y compris celles à propos de la couverture.
- Lorsque le service d'ATR recevait des appels concernant un logement susceptible d'avoir été omis par le recensement, une procédure était mise en place pour examiner les environs à la recherche d'autres logements susceptibles d'avoir été omis.
- Le questionnaire comprenait une partie intitulée « Qui inscrire », ce qui permettait aux répondants de déterminer les personnes qui devaient être incluses. De plus, un peu plus de 84 % des réponses au Recensement de 2021 ont été obtenues par Internet, et le questionnaire électronique comprenait des questions de vérification supplémentaires lorsque les répondants déclaraient qu'un logement était inoccupé ou inexistant, ou s'ils avaient de la difficulté à déterminer si une personne devait être incluse ou non.
- Dans le questionnaire, on demandait aux répondants d'indiquer s'il y avait des personnes qui n'avaient pas été inscrites parce qu'ils n'étaient pas certains qu'il fallait les inclure. Le questionnaire électronique guidait les répondants pour qu'ils prennent la bonne décision. Dans les autres cas, un suivi téléphonique était effectué par la suite auprès des répondants pour déterminer si les personnes en question devaient être inscrites ou non dans le questionnaire.
- Le suivi téléphonique a été effectué après l'examen des questionnaires en cas d'incohérence quant à la couverture ou pour vérifier la situation des ménages, y compris les questionnaires ne concernant que des résidents étrangers ou des personnes présentes temporairement.

Ces procédures ainsi qu'une formation appropriée du personnel, des vérifications de supervision et des contrôles qualitatifs durant les étapes de collecte et de traitement ont permis de réduire le nombre d'erreurs de couverture.

3.3 Définitions

La présente section porte sur la définition algébrique des erreurs de couverture. La lettre P représente le total ou le nombre « réel » de personnes ciblées par le Recensement de la population. La lettre C représente le chiffre publié du recensement du nombre de personnes dans la population cible. L'erreur associée à l'utilisation de C plutôt que de P est la suivante :

$$N = P - C$$

Cette erreur, désignée au moyen de la lettre N , représente l'**erreur de couverture nette de la population**.

La lettre S désigne le **sous-dénombrement de la population**, c'est-à-dire le nombre de personnes qui auraient dû être prises en compte dans C , mais qui ne l'ont pas été.

Le nombre de personnes dans le recensement, C , comporte deux éléments, soit :

$$C = D + I$$

où :

D est le nombre de personnes dénombrées, c'est-à-dire le nombre de personnes inscrites dans un questionnaire du recensement.

I est le nombre de personnes imputées. Il s'agit d'une estimation du nombre de personnes omises parce que leur logement a été classé comme étant occupé, mais non répondant ou classé par erreur comme étant inoccupé (donc pour lequel aucun suivi n'a été fait). Pour obtenir plus de renseignements sur l'imputation des ménages entiers (IME), voir la [section 3.6](#) du *Rapport technique sur l'échantillonnage et la pondération, Recensement de la population, 2021*, n° 98-306-X au catalogue de Statistique Canada.

Le sous-dénombrement par rapport au nombre de personnes publié dans le recensement C est donc ce qui reste des personnes qui auraient dû être inscrites dans un questionnaire du recensement et qui n'ont pas été prises en compte par l'IME. Autrement dit, le sous-dénombrement ne comprend pas l'estimation du nombre de personnes qui n'ont pas été dénombrées soit parce qu'un questionnaire du recensement rempli n'a pas été retourné pour le logement (logement non répondant), soit parce que le logement n'a pas reçu de questionnaire puisqu'il a été classé par erreur comme étant inoccupé (erreur de classification).

Le concept de sous-dénombrement avant l'IME existe aussi. C'est ce que l'on appelle le sous-dénombrement de la collecte du Recensement de la population. Pour obtenir plus de renseignements, voir la [section 12.1](#).

La lettre R désigne le **surdénombrement de la population**, c'est-à-dire le nombre de dénombremments excédentaires compris dans C , qui n'auraient pas dû l'être.

R comporte deux éléments. Le premier est le dénombrement excédentaire de personnes, c'est-à-dire qu'elles ont été prises en compte plus d'une fois. Les études sur la couverture portent sur ces dénombremments excédentaires. Le deuxième concerne les personnes qui ont été dénombrées, mais qui ne faisaient pas partie de la population cible du recensement. Par exemple, les résidents étrangers en visite au Canada dont le nom figure dans un questionnaire du recensement en tant que résidents habituels d'un logement ne devraient pas faire partie de C . Les personnes fictives sont un autre exemple. Selon des études antérieures, le nombre de personnes qui sont dénombrées, mais qui ne font pas partie de la population cible du recensement est en général très faible et peut être ignoré. Par conséquent, les études sur la couverture du recensement n'évaluent pas cet élément des erreurs de couverture.

Rapport technique sur la couverture

Puisque S désigne les personnes non dénombrées qui devraient être comprises dans C et que R désigne les dénombrements qui ne devraient pas être compris dans C , l'écart entre P et C correspond à S moins R , c'est-à-dire :

$$N = S - R$$

Le nombre réel de personnes dans la population cible du recensement est alors :

$$P = C + N = C + S - R$$

En pratique, pour des raisons de coût et d'actualité des données produites, on établit une estimation de P , dénotée \hat{P} , fondée sur des études par échantillon, où :

$$\hat{P} = C + \hat{N} = C + \hat{S} - \hat{R}$$

\hat{S} est une estimation du nombre de personnes non incluses dans C qui auraient dû l'être et \hat{R} est une estimation du nombre de personnes incluses dans C qui n'auraient pas dû l'être. On peut supposer que le surdénombrement des personnes comprises dans C alors qu'elles ne font pas partie de la population visée par le recensement est égal à zéro, car il est négligeable. \hat{R} se limite donc à une estimation du nombre de dénombrements en double. Les études sur la couverture du recensement visent à établir les valeurs \hat{S} et de \hat{R} .

En bref, la population réelle P est formée du nombre de personnes publiés dans le recensement C et du sous-dénombrement net N . On parle de sous-dénombrement net, car S est généralement plus grand que R dans le contexte du recensement actuel au Canada. Toutefois, l'inverse est possible et N serait alors négatif. La variable C correspond à D plus le nombre de personnes ajoutées lors de l'IME, et cette imputation I vise les personnes habitant des logements non répondants ou des logements occupés classés par erreur comme étant des logements inoccupés.

Les erreurs de couverture de la population du recensement peuvent généralement être exprimées en fonction de taux relatifs à la population réelle. Le taux de sous-dénombrement T_S est S exprimé en un pourcentage de P . Le taux de surdénombrement T_R est R exprimé en un pourcentage de P . Le taux de sous-dénombrement net T_N correspond à l'écart entre S et R , exprimé en un pourcentage de la population cible du recensement. Ces trois taux peuvent être estimés au moyen de \hat{T}_S , \hat{T}_R et \hat{T}_N comme suit :

$$\hat{T}_S = 100 * \frac{\hat{S}}{\hat{P}} = 100 * \frac{\hat{S}}{C + \hat{N}}$$

$$\hat{T}_R = 100 * \frac{\hat{R}}{\hat{P}} = 100 * \frac{\hat{R}}{C + \hat{N}}$$

$$\hat{T}_N = 100 * \frac{\hat{N}}{\hat{P}} = 100 * \left(\frac{\hat{S} - \hat{R}}{C + \hat{N}} \right)$$

Un taux de sous-dénombrement net positif indique que le taux de sous-dénombrement est plus élevé que le taux de surdénombrement. Autrement dit, le nombre de personnes non comprises dans les chiffres publiés du recensement C est plus élevé que le nombre de personnes dénombrées en trop. C'est généralement le cas pour tous les recensements du Canada. Cependant, un taux de sous-dénombrement net négatif est parfois observé pour certains domaines d'intérêt.

3.4 Évaluation

Deux études postcensitaires ont été menées pour estimer les erreurs de couverture de la population du Recensement de 2021. L'Étude sur le sous-dénombrement du recensement (ESoR) a fourni des estimations du sous-dénombrement de la population, tandis que l'Étude sur le surdénombrement du recensement (ESuR) a permis d'estimer le surdénombrement de la population. Comme il est mentionné précédemment, l'Enquête sur la classification des logements (ECL) ne contribue pas à l'estimation des erreurs de couverture du recensement, puisque les chiffres du recensement sont déjà corrigés pour tenir compte des résultats de l'ECL.

L'ESoR et l'ESuR ont été réalisées après les opérations de collecte sur le terrain et de traitement du recensement. Les estimations provisoires des erreurs de couverture de la population du Recensement de 2021 ont été publiées le 28 avril 2023. Les estimations définitives ont été publiées le 27 septembre 2023, après un exercice de validation approfondi effectué en collaboration avec le Centre de démographie et les coordonnateurs statistiques provinciaux et territoriaux. Ces données ont été publiées en même temps que les nouvelles estimations démographiques officielles tenant compte de la mise à jour de la population de base du Recensement de 2021. Les chiffres de population du recensement corrigés pour tenir compte du sous-dénombrement net de la population constituaient les estimations à jour de la population de base.

Une brève description de la méthodologie utilisée dans les deux études sur la couverture du recensement est présentée ci-dessous :

Étude sur le sous-dénombrement du recensement

Dans l'ESoR, un échantillon aléatoire de personnes représentant la population cible du Recensement de 2021 a été sélectionné à partir de bases de sondage indépendantes du recensement. Ces bases de sondage sont décrites à la [section 7.1](#). La base de données du Recensement de 2021 a ensuite été consultée pour déterminer si ces personnes avaient bien été dénombrées.

On a également effectué, au besoin, des interviews, dont la majorité a été des interviews téléphoniques assistées par ordinateur, à partir des bureaux régionaux (BR), afin de recueillir des renseignements pouvant servir à des recherches supplémentaires dans la base de données du Recensement de 2021. Une interview a été réalisée dans 60,2 % des 12 787 cas transmis aux BR. Le poids d'échantillonnage a été corrigé pour tenir compte de la non-réponse. Plus précisément, le poids d'échantillonnage total des non-répondants a été réparti entre des groupes de répondants ressemblant le plus possible aux non-répondants quant à leur probabilité de réponse.

L'estimation du sous-dénombrement de la population est basée sur le nombre de personnes dans l'échantillon de l'ESoR qui ont été classées comme des personnes « omises ». Ces personnes faisaient partie de la population cible du Recensement de 2021, mais aucune preuve de dénombrement n'a pu être trouvée dans la base de données des réponses du Recensement de 2021. À l'échelle nationale, 6 212 personnes faisant partie de l'échantillon de l'ESoR ont été classées comme des personnes omises dans les provinces, et 1 241, dans les territoires.

Étude sur le surdénombrement du recensement

Le surdénombrement a été mesuré en créant une liste d'enregistrements potentiellement en double dans la base de données du Recensement de 2021 à l'aide de méthodes de couplage probabiliste et déterministe pour effectuer un appariement de la base de données définitive du Recensement de 2021 avec elle-même, suivi d'un appariement entre la base de données définitive du Recensement de 2021 et une liste de personnes qui auraient dû être dénombrées d'après des sources de données administratives. Le couplage probabiliste permet d'estimer la probabilité que les paires liées soient de vraies concordances en se basant sur un modèle de concordance obtenu en comparant les variables de couplage entre les enregistrements. Le couplage déterministe permet de classer les paires d'enregistrements comme concordances lorsque les variables de couplage sont généralement en accord et comme non-concordances dans le cas contraire. Aucun poids n'est attribué à la force de la concordance.

Un échantillon aléatoire de doublons potentiels a été sélectionné à partir de la liste créée à l'aide des couplages, et les paires ont été vérifiées manuellement en comparant les caractéristiques démographiques ainsi que les noms afin de repérer les véritables cas de surdénombrement. Le surdénombrement de la population a été estimé à partir de cet échantillon.

4. Collecte des données du recensement

L'étape de la collecte des données du Recensement de 2021 visait à faire en sorte que chacun des 16,3 millions de logements au Canada soit dénombré et que, pour chaque logement occupé, le ménage correspondant remplisse un questionnaire du recensement.

En 2021, les ménages canadiens avaient la possibilité de répondre en ligne ou de communiquer avec le service d'Assistance téléphonique du recensement pour demander un questionnaire papier, s'il le souhaitait. Les ménages qui n'ont pas utilisé l'un de ces deux modes de réponse ont reçu un suivi par téléphone ou en personne. Les quatre méthodes de collecte utilisées pour le recensement canadien étaient l'envoi par la poste, le listage/livraison, l'envoi par la poste et livraison à la porte (EPLP), et le recensement par interview. Afin que la collecte du recensement se fasse le plus efficacement possible, le territoire du Canada a été divisé en petites unités géographiques appelées unités de collecte (UC). Lors du Recensement de 2021, on comptait environ 49 000 UC au Canada.

- Dans les UC dont la méthode de collecte était l'envoi par la poste, on a eu recours au service postal pour envoyer le matériel du recensement. Ce type d'UC se trouve généralement dans les régions urbaines. Bien qu'il représente maintenant environ 86 % des logements privés au Canada, il ne représente qu'une infime partie de la superficie du pays.
- Les UC dont la méthode de collecte est le listage/livraison se trouvent généralement dans les régions rurales. Dans ces régions, les employés du recensement avaient préparé une liste de logements et déposé une lettre d'invitation à leur porte. Environ 7 % des logements au Canada se trouvent dans une UC de type listage/livraison, et ce type d'UC couvre une grande partie de la superficie du pays.
- En 2021, la méthode d'EPLP a été mise en place pour les secteurs où tous les logements disposent d'une adresse et dont la majorité peut recevoir du courrier. Dans ces secteurs mixtes, les logements ayant une adresse postale valide ont reçu le matériel d'envoi postal habituel, tandis que ceux n'ayant pas d'adresse postale valide (correspondant à l'adresse municipale) ont reçu une lettre d'invitation déposée à leur porte par un employé du recensement. Cela a permis d'étendre l'utilisation de la méthode d'envoi par la poste à environ 90 % des logements.
- Le dénombrement par interview avec un agent recenseur a lieu traditionnellement dans les communautés des Premières Nations, les établissements métis, les régions inuites et d'autres régions éloignées. Cependant, pour la première fois, en 2021, l'ensemble des communautés des Premières Nations, des établissements métis, des régions inuites et des autres régions éloignées se sont vu offrir la possibilité de répondre au recensement par eux-mêmes, à condition que cela soit possible sur le plan opérationnel (c.-à-d. qu'Internet soit accessible dans la communauté). Les ménages vivant dans des régions où il n'était pas possible, d'un point de vue opérationnel, d'offrir l'autodéclaration ont rempli leur questionnaire du recensement avec un employé du recensement (en personne ou par téléphone). En 2021, les logements des communautés éloignées, nordiques et autochtones représentaient environ 1 % des logements au Canada.

Depuis le Recensement de 2011, une méthode de collecte par vagues est utilisée pour la collecte du recensement. Elle consiste à communiquer avec les ménages non répondants à des moments clés afin de leur rappeler de participer au recensement et de les convaincre de remplir le questionnaire. À chaque vague, les renseignements nécessaires pour répondre au questionnaire sont fournis aux ménages. Cette méthode diffère selon la méthode de collecte utilisée pour distribuer le matériel du recensement pour une région donnée.

La méthode par vagues, qui s'inscrit dans un contexte où chaque ménage canadien est tenu par la loi de répondre aux questions du recensement, a été conçue de façon à favoriser la réponse hâtive par Internet tout en offrant une solution de rechange aux ménages qui ne souhaitent pas remplir leur questionnaire par Internet. Cette méthode comporte de nombreux avantages. Elle augmente le nombre de réponses par Internet, qui est une méthode facilitant l'enchaînement des questions et la modification en temps réel des réponses. Elle réduit la non-réponse, la nécessité d'effectuer des suivis coûteux et le nombre de questionnaires à enregistrer et de données à saisir. En somme, elle augmente la qualité du recensement.

Rapport technique sur la couverture

La pandémie de COVID-19 est apparue au Canada au début de 2020 et a touché toutes les étapes du processus du Recensement de 2021, de la collecte des données à la diffusion.

Aucune question n'a été ajoutée au questionnaire du Recensement de 2021 pour recueillir des renseignements sur la COVID-19. Néanmoins, Statistique Canada est conscient que la pandémie de COVID-19 peut avoir eu une incidence sur les réponses à certaines questions du recensement, notamment celles se rapportant à l'emploi, à la scolarité, au navettage et aux dépenses. Au moment de fournir des réponses aux questions du recensement, les répondants ont reçu pour consigne de choisir les réponses qui reflètent le mieux leur situation ou celle des membres de leur ménage pour la date ou la période en question. Des instructions supplémentaires étaient également accessibles dans les fonctions d'aide des questionnaires en ligne.

Dans le contexte de la pandémie de COVID-19, Statistique Canada a embauché 15 conseillers supplémentaires au Programme de liaison avec les Autochtones pour former le Groupe de travail sur la participation des Autochtones au recensement afin d'appuyer les efforts de mobilisation des communautés autochtones et nordiques pour le Recensement de 2021.

Il a fallu adapter la stratégie de collecte du Recensement de la population de 2021 pour la transformer en un processus complet sans contact qui garantit la sécurité des répondants et des employés du recensement. Par exemple, le dénombrement anticipé a été annulé pour les communautés des Premières Nations, les établissements métis, les régions inuites et d'autres régions éloignées, et l'on a insisté pour que le questionnaire du recensement soit rempli en ligne. Aussi, plus de ressources ont été affectées au suivi téléphonique pour réduire le nombre de visites en personne, et une lettre de rappel supplémentaire a été envoyée aux non-répondants en juillet.

Les procédures de collecte de données du Recensement de 2021 ont été remaniées pour veiller à ce que les répondants et les employés du recensement soient en sécurité en limitant le nombre de contacts nécessaires pour participer au recensement. Aucun employé du recensement de Statistique Canada n'a été autorisé à visiter les logements collectifs institutionnels, en particulier les logements où résidaient des personnes vulnérables à la COVID-19, comme les résidences pour personnes âgées et les hôpitaux. Pour la majorité des logements collectifs, le questionnaire électronique du [Recensement de 2021 : logements collectifs](#) a été utilisé pour recueillir des renseignements sur les établissements et leurs résidents habituels. En avril 2021, des lettres d'invitation ou des courriels contenant un code d'accès sécurisé ont été envoyés aux administrateurs des logements collectifs pour qu'ils répondent au questionnaire en ligne.

La pandémie a présenté certains défis pour la réalisation du Recensement de la population de 2021, mais malgré ces défis, le taux de réponse de la collecte pour le pays dans son ensemble constitue un franc succès, ayant atteint 98,0 % grâce aux Canadiens qui ont rempli leur questionnaire du recensement en pleine troisième vague de la pandémie. De plus, d'autres défis particuliers se sont posés dans les régions nordiques ou éloignées du pays, comme des restrictions relatives aux déplacements, des fermetures de frontières, des périodes de collecte plus courtes et décalées, des incendies de forêt et l'indisponibilité du personnel local. Assurer la santé et la sécurité des Canadiens et des employés de Statistique Canada en adaptant les opérations de collecte de l'organisme tout en préservant la grande qualité et la fiabilité des données du recensement était une priorité importante pour Statistique Canada. Pour obtenir plus de renseignements sur la collecte des données du recensement, veuillez consulter le [chapitre 1](#) du *Rapport technique sur l'échantillonnage et la pondération, Recensement de la population, 2021*, n° 98-306-X au catalogue de Statistique Canada.

5. Traitement des données du recensement

Le traitement de tous les questionnaires remplis, quel que soit leur type, comporte plusieurs étapes, de la réception des questionnaires à la création d'une base de données du recensement exacte et complète. Ces étapes comprennent l'enregistrement des questionnaires, l'imagerie et la saisie des données des questionnaires, le contrôle, la correction des erreurs, le suivi des questionnaires rejetés au contrôle, le codage, la classification des logements et les ajustements pour la non-réponse, le couplage des données administratives, l'imputation et la pondération (pour l'échantillon du questionnaire détaillé).

Les processus automatisés mis en œuvre pour le Recensement de 2021 ont fait l'objet d'une surveillance de manière à ce que toutes les résidences canadiennes ne soient dénombrées qu'une seule fois. Le Système de contrôle principal (SCP) a été mis au point afin de contrôler et de surveiller l'enchaînement des opérations, de la collecte au traitement des données. Le SCP contenait une liste principale de tous les logements au Canada, sur laquelle figurait pour chacun d'eux un identificateur unique. Ce système était mis à jour quotidiennement, en fonction des renseignements sur la situation de chaque logement dans le processus d'enchaînement des opérations du recensement (p. ex. livraison, réception ou traitement). Des rapports étaient produits quotidiennement et mis en ligne à la disposition des gestionnaires afin que les opérations du recensement se déroulent de façon efficace et efficiente.

Pour obtenir plus de renseignements sur le traitement des données du recensement, veuillez consulter le [chapitre 3](#) du *Rapport technique sur l'échantillonnage et la pondération, Recensement de la population, 2021*, n° 98-306-X au catalogue de Statistique Canada.

6. Enquête sur la classification des logements

6.1 Introduction

L'Enquête sur la classification des logements (ECL) mesure les erreurs de classification de logements lors du recensement. L'ECL permet également d'estimer le nombre de résidents habituels vivant dans les logements non répondants occupés. Au moyen de ces résultats, l'occupation ou l'inoccupation des logements non répondants est corrigée, et les personnes sont imputées dans les logements occupés lors de la procédure d'imputation des ménages entiers (IME), en fonction des fichiers des totaux de contrôle de l'ECL. Dans le présent chapitre, le terme « logement » désigne toujours des logements privés³.

L'une des sources possibles d'erreur lors d'un recensement est la classification erronée d'un logement. Lorsqu'un ménage omet de retourner son questionnaire de recensement, l'agent recenseur doit déterminer si le logement est occupé. Deux types d'erreurs de classification de logement peuvent survenir dans cette situation. **Premièrement, un logement occupé peut être classé par erreur comme étant inoccupé.** Si ces logements étaient en fin de compte considérés comme inoccupés, cela entraînerait un sous-dénombrement de la population du recensement, car les occupants du logement ne seraient pas dénombrés. **Deuxièmement, un logement inoccupé peut être classé par erreur comme étant occupé.** Si ces logements étaient en fin de compte considérés comme occupés, cela entraînerait un surdénombrement de la population du recensement, car des personnes seraient dénombrées, alors qu'en fait, personne ne vit dans le logement. Les estimations découlant de l'ECL servent à corriger les données du recensement pour tenir compte de ces deux erreurs de couverture.

De plus, l'ECL mesure un troisième type d'erreur de classification de logement : les erreurs survenant lorsque des logements marginaux ou en construction sont classés par erreur comme étant des logements. Cela entraîne un surdénombrement des logements et pourrait donner lieu à un surdénombrement de la population si le logement est classé comme étant occupé. Les données du recensement ne sont pas corrigées pour tenir compte de ces logements, de sorte que les estimations du parc immobilier au recensement comprennent un certain degré de surdénombrement. Ce sujet est abordé plus en profondeur à la [section 6.3.1.2](#) et à la [section 6.3.2.2](#).

6.2 Méthodologie

6.2.1 Stratification et sélection de l'échantillon

La population visée par l'ECL comprenait tous les logements privés classés comme étant des logements inoccupés ou non répondants, sauf les logements dans les communautés des Premières Nations, les établissements métis, les régions inuites et d'autres régions éloignées, ainsi que les logements privés attenants à un logement collectif⁴. Ces secteurs sont exclus pour des raisons de coûts et de facteurs opérationnels.

La taille de l'échantillon de l'ECL a été fixée à 1903 unités de collecte (UC). La base de sondage comprenait toutes les UC non éloignées et hors réserves (c.-à-d. les UC dont la méthode de collecte est celle de type envoi par la poste, envoi par la poste et livraison à la porte [EPLP], listage/livraison et saisonnier)⁵. Par conséquent, aucune UC du Nunavut ne faisait partie du champ de l'enquête, si bien que l'ECL n'y a pas eu lieu. Le plan d'échantillonnage était le suivant : toutes les UC faisant partie du champ de l'enquête au Yukon (42 UC) et dans les Territoires du Nord-Ouest (21 UC) formaient deux strates, une pour chaque territoire. Toutes leurs UC faisant partie du champ de l'enquête ont été sélectionnées pour l'échantillon de l'ECL avec certitude. Toutes les UC de l'Île-du-Prince-Édouard formaient une troisième strate, à partir de laquelle un échantillon aléatoire simple de 51 UC a été sélectionné.

3. Voir la [section 2.3](#) pour obtenir la définition d'un logement privé.

4. Voir le [chapitre 7](#) du *Guide du Recensement de la population, 2021*, produit N° 98-304-X au catalogue de Statistique Canada.

5. La méthode de collecte saisonnière est la même que la méthode de collecte par listage/livraison. Cependant, les secteurs pour lesquels ces méthodes ont été utilisées comprennent des logements qui sont principalement utilisés de façon saisonnière et par conséquent, ont tendance à avoir de faibles taux d'occupation. Cela est important lorsque vient le temps de décider du nombre de logements à échantillonner dans ces régions.

Rapport technique sur la couverture

Les UC restantes dans chaque province ont été regroupées en strates urbaines et rurales. Une UC était initialement considérée comme urbaine si elle avait fait partie d'une région métropolitaine de recensement (RMR) ou d'une agglomération de recensement (AR) comptant 40 000 logements ou plus. Par ailleurs, toutes les UC relevant d'un district de chef d'équipe (DCE) étaient considérées comme urbaines si plus de 50 % des UC relevant du DCE étaient urbaines. Toutes les UC restantes formaient les strates rurales. Les UC urbaines ont été stratifiées par RMR et AR. Un échantillon aléatoire simple d'au moins cinq UC a été choisi pour chaque strate. Les données des recensements précédents ont permis d'établir que cinq UC constituaient une charge de travail appropriée pour un intervieweur. L'échantillon comptait 1 037 UC urbaines. Pour contrôler les coûts des opérations sur le terrain, des UC géographiquement voisines ont été choisies pour l'échantillon des UC rurales. La sélection a été effectuée au moyen d'un plan de sondage aléatoire stratifié à deux degrés. Pour le premier degré, des DCE ont été choisis dans chaque province en utilisant un échantillonnage aléatoire simple. Pour le deuxième degré, cinq UC ont été choisies de manière aléatoire dans chacun des DCE sélectionnés. L'échantillon comptait 866 UC rurales.

Un sous-échantillonnage de logements au sein d'une UC échantillonnée a eu lieu lorsque le nombre de logements inoccupés et non répondants dépassait un paramètre maximal de logements, qui était de 35 pour les UC d'envoi par la poste et les UC de listage/livraison, de 40 pour les UC d'EPLP, et de 135 pour les UC saisonnières. Un sous-échantillonnage de logements faisant partie du champ de l'enquête a eu lieu pour 584 UC. Sinon, tous les logements inoccupés et les logements non répondants⁶ faisant partie des UC échantillonnées formaient l'échantillon des logements de l'ECL. Au total, 34 459 logements inoccupés et 11 346 logements non répondants ont été échantillonnés en 2021. Le tableau 6.2.1 présente la répartition de l'échantillon par province et territoire.

Tableau 6.2.1
Taille de l'échantillon pour le Canada, les provinces et les territoires

Province et territoires	Nombre d'unités de collecte	Nombre de logements inoccupés	Nombre de logements non répondants
Canada	1 903	34 459	11 346
Terre-Neuve-et-Labrador	159	4 404	466
Île-du-Prince-Édouard	51	1 311	175
Nouvelle-Écosse	125	2 831	501
Nouveau-Brunswick	109	1 651	710
Québec	316	4 987	2 145
Ontario	394	6 700	2 281
Manitoba	164	2 186	957
Saskatchewan	135	2 670	770
Alberta	190	3 118	1 354
Colombie-Britannique	197	3 646	1 434
Yukon	42	757	245
Territoires du Nord-Ouest	21	198	308
Nunavut	0	0	0

Source : Statistique Canada, Enquête sur la classification des logements de 2021.

6. L'état du logement est fondé sur l'état au moment de la sélection de l'échantillon de l'ECL. Il peut changer avant la fin de la période de collecte du recensement (p. ex. si un logement répond au recensement).

6.2.2 Interviews sur le terrain

Un questionnaire de l'ECL a été utilisé pour vérifier l'occupation/inoccupation réelle le jour du recensement des logements échantillonnés faisant partie des UC échantillonnées. L'occupation/inoccupation a été vérifiée de la mi-juin à la mi-août 2021. Les opérations de l'ECL commençaient idéalement dès qu'un seuil de taux de réponse prédéterminé pour le Recensement de la population était obtenu pour l'UC échantillonnée, mais parfois, elles devaient commencer plus tôt en raison de contraintes opérationnelles.

Pour déterminer l'occupation/inoccupation et recueillir d'autres renseignements, les agents recenseurs avaient pour instruction de communiquer avec les occupants, les voisins, les propriétaires ou toute autre personne connaissant le logement. On a effectué jusqu'à trois tentatives de prise de contact pour chaque logement. Si le logement s'avérait occupé le jour du recensement, on obtenait également le nombre d'occupants le jour du recensement, ainsi que le sexe et l'âge des occupants, si possible.

6.2.3 Traitement et estimation

Tous les questionnaires remplis ont été envoyés pour traitement au bureau central à Ottawa.

La première étape de traitement était la saisie des données. Au terme de cette étape, les questionnaires ont été soumis à une série exhaustive de contrôles de la cohérence. Les questionnaires pour lesquels les contrôles ont échoué ont été examinés manuellement afin de résoudre les incohérences.

À ce stade du processus de traitement, les logements inoccupés et les logements non répondants figurant dans l'échantillon ont été séparés et la classification de ces logements a été confirmée en fonction de la base de données du recensement. Les questionnaires remplis pour chaque UC échantillonnée ont été appariés aux listes définitives des logements inoccupés et des logements non répondants du recensement (c.-à-d. aucun questionnaire du recensement n'a été reçu pour le logement). En cas de non-appariement, ce qui signifie que l'état du logement avait changé après la sélection de l'échantillon de l'ECL, le logement échantillonné était éliminé et aucun autre traitement n'était requis. Les logements inscrits comme étant inoccupés ou non répondants sur les listes du recensement des UC échantillonnées et pour lesquels aucun questionnaire de l'ECL n'avait été reçu ont été considérés comme des cas de non-réponse totale à l'ECL et sont passés à l'étape suivante du traitement.

Les cas de non-réponse totale à l'ECL ont donné lieu à une correction par pondération, tandis qu'une imputation partielle a été appliquée aux cas de non-réponse partielle. La procédure était la même pour les logements inoccupés et les logements non répondants. En l'absence de renseignements sur un logement, les poids de sondage des répondants ont été corrigés, afin de tenir compte du poids de sondage des non-répondants. La correction a été faite séparément par strate géographique de second niveau. Voir [Hong \(2023\)](#) pour obtenir plus de précisions à propos des strates de second niveau utilisées. Les poids de sondage ont ensuite été corrigés, de sorte que la somme des poids corrigés pour chaque strate géographique de second niveau soit égale au nombre total de logements inoccupés ou de logements non répondants. Les cas de non-réponse partielle concernant l'occupation/inoccupation et le nombre de résidents habituels ont été traités par imputation. L'occupation/inoccupation a été imputée en premier, puis utilisée dans l'imputation des autres variables.

Finalement, la taille du ménage, recueillie au moyen du questionnaire de l'ECL lorsqu'un logement s'avérait occupé, a servi à produire une estimation, par taille de ménage, du taux d'occupation des logements privés classés comme étant inoccupés ou non répondants lors du recensement.

6.2.4 Imputation des ménages entiers du recensement

La procédure d'IME vise à représenter les logements privés non répondants du recensement; on procède ainsi à une imputation de la non-réponse totale au recensement. La stratégie varie selon que l'ECL a été menée ou pas dans la région. Le résultat de la procédure d'IME est que tous les logements privés ont été classés comme étant occupés ou inoccupés dans la base de données du recensement (c.-à-d. qu'il ne reste plus de logements non répondants).

Dans les secteurs où l'ECL a été menée (c.-à-d. les UC de types envoi par la poste, EPLP, listage/livraison et collecte saisonnière, à l'exclusion des logements privés attenants à un logement collectif), les estimations de l'ECL ont été utilisées comme intrants dans l'algorithme de la procédure d'IME, afin d'imputer les logements privés occupés en fonction de la répartition des tailles des ménages et des strates de second niveau. Tout d'abord, au sein d'une strate géographique de second niveau de l'ECL, tous les logements sans réponse ont été relevés (cette étape s'est déroulée séparément pour les univers des logements inoccupés et non répondants). Ensuite, pour l'univers des logements non répondants seulement, tout logement non répondant pour lequel la collecte sur le terrain avait permis d'obtenir le nombre de résidents habituels a été considéré comme occupé et s'est vu attribuer la taille du ménage enregistrée. Enfin, des logements sans réponse ont été sélectionnés aléatoirement et imputés comme étant occupés. La sélection s'est effectuée de sorte que le nombre définitif de logements non répondants et inoccupés convertis en logements occupés pour la strate de second niveau soit égal à l'estimation des logements occupés de l'ECL dans les univers de logements non répondants et inoccupés. De plus, comme nouveauté pour le cycle de 2021, les fichiers des totaux de contrôle ont été utilisés, de sorte que l'estimation de l'ECL du taux d'occupation dans chaque strate de second niveau a été appliquée dans chaque subdivision de recensement. Une fois qu'un logement privé inoccupé ou non répondant était imputé comme étant occupé, une procédure soumise aux contraintes des estimations de l'ECL par strate de second niveau et taille du ménage a servi à imputer la taille de logement du ménage. Puis, le logement le plus proche sur le plan géographique de la taille du ménage prédéterminée a été choisi comme donneur. Les réponses au questionnaire abrégé du recensement de ce ménage donneur ont ensuite été attribuées au logement inoccupé ou non répondant imputé comme étant occupé.

Dans les régions géographiques où l'ECL n'a pas été menée (c.-à-d. les UC ayant des types de collecte pour les régions éloignées et les réserves, ainsi que les logements privés attenants à un logement collectif), la vaste majorité des logements privés classés comme étant non répondants ont été imputés comme étant occupés, la taille du ménage provenant de la collecte sur le terrain ou d'un ménage donneur. Le logement le plus proche sur le plan géographique de la taille du ménage prédéterminée a été choisi comme donneur. Si le logement ne s'était pas encore fait attribuer de taille de ménage, on choisissait la taille du ménage du donneur comme taille de ménage. Les renseignements de ce ménage donneur ont ensuite été attribués au logement inoccupé ou non répondant imputé comme étant occupé. Pour les logements situés dans les régions de collecte éloignées ou dans les réserves, l'enregistrement complet leur a été attribué. Pour les logements privés attenants à un logement collectif, seules les réponses au questionnaire abrégé du recensement leur ont été attribuées.

Plusieurs adaptations ont été introduites dans le plan de collecte de données du Recensement de 2021 afin d'atténuer l'incidence de la pandémie de COVID-19 et d'autres risques. Un plan de contingence statistique a été dressé en fonction de l'utilisation sécuritaire, responsable et appropriée des données administratives afin d'appuyer le Recensement de 2021 en cas de perturbations dans la collecte de données. Des données administratives couplées provenant de sources de données fédérales et provinciales ont été utilisées dans certaines régions pour améliorer l'imputation des ménages non répondants à l'étape de l'imputation après la fin de la collecte de données du Recensement de 2021⁷. Les ménages créés à partir de données administratives ont été ajoutés au bassin de ménages donneurs dans ces régions.

De plus amples renseignements sur la procédure d'IME figurent dans le rapport de [Hong \(2023\)](#).

6.3 Estimations

Les données du recensement ont été corrigées à l'aide des estimations de l'ECL, au moyen de la procédure d'IME, pour les logements non répondants et les logements occupés ayant été classés par erreur comme étant inoccupés. Les estimations sont présentées aux sections [6.3.1.1](#) et [6.3.2.1](#). Les données du recensement n'ont pas été corrigées pour tenir compte des logements marginaux ou en construction classés par erreur comme étant des logements. Les sections [6.3.1.2](#) et [6.3.2.2](#) présentent les estimations du nombre de logements marginaux et de logements en construction classés par erreur comme étant des logements et donc ajoutés par erreur au parc immobilier.

7. Voir l'[annexe 1.7](#) du *Guide du Recensement de la population, 2021*, produit n° 98-304-X au catalogue de Statistique Canada.

6.3.1 Logements inoccupés

6.3.1.1 Logements occupés classés par erreur comme étant inoccupés

Le [tableau 6.3.1.1.1](#) présente le nombre estimé de logements occupés qui ont été classés par erreur comme étant inoccupés ainsi que le taux d'erreur correspondant pour les logements inoccupés par région urbaine et rurale⁸, et par province et territoire. Aux fins de comparaison, le [tableau 6.3.1.1.2](#) fournit les estimations pour les mêmes variables pour le Recensement de 2016.

Le [tableau 6.3.1.1.1](#) indique que 17,3 % de tous les logements classés comme étant inoccupés en 2021 étaient en fait des logements occupés. Ce taux est en légère hausse par rapport au taux de 15,0 % observé en 2016 et constitue une variation statistiquement significative. La classification erronée des logements était plus fréquente dans les régions urbaines (23,8 %) que dans les régions rurales (10,9 %). Les régions urbaines et rurales ont enregistré une augmentation par rapport à 2016. Le taux de classification erronée a augmenté ou est resté pratiquement inchangé dans l'ensemble des provinces et des territoires, à l'exception des Territoires du Nord-Ouest, où il a diminué.

Parmi les provinces et les territoires, le Nouveau-Brunswick a enregistré le taux de classification erronée le plus élevé (20,2 %), suivi de l'Ontario et de la Colombie-Britannique (18,9 %) et du Québec (17,7 %). Les taux des autres provinces et des autres territoires variaient de 8,1 % à l'Île-du-Prince-Édouard à 16,6 % en Alberta.

Le [tableau 6.3.1.1.3](#) présente le nombre estimé de logements occupés qui ont été classés par erreur comme étant inoccupés, le nombre estimatif de personnes vivant dans ces logements et l'erreur-type. Ces estimations ont été utilisées pour décider du nombre de logements à imputer en tant que logements occupés dans la procédure d'IME. Le [tableau 6.3.1.1.4](#) fournit les estimations pour les mêmes variables pour le Recensement de 2016. Du fait d'erreurs dans la classification initiale des logements, environ 218 042 ménages n'ont pas été dénombrés dans le Recensement de 2021. Il s'agit du nombre de ménages ajoutés au recensement durant la procédure d'IME.

Tableau 6.3.1.1.1

Nombres de logements occupés classés par erreur comme étant des logements inoccupés selon diverses régions géographiques, Recensement de 2021

Région géographique	Nombre de logements inoccupés	Logements occupés classés par erreur comme étant inoccupés			
		Nombre estimé	Erreur-type	Taux estimé ¹ (%)	Erreur-type (%)
Canada	1 259 149	218 042	10 097	17,3	0,8
Urbaine	626 525	148 973	5 918	23,8	0,9
Rurale	632 624	69 069	5 565	10,9	0,9
Provinces de l'Atlantique	132 141	16 504	1 359	12,5	1,0
Terre-Neuve-et-Labrador	44 142	3 818	586	8,6	1,3
Île-du-Prince-Édouard	10 345	836	127	8,1	1,2
Nouvelle-Écosse	47 584	5 778	743	12,1	1,6
Nouveau-Brunswick	30 070	6 073	764	20,2	2,5
Québec	285 920	50 477	5 283	17,7	1,8
Ontario	433 952	82 077	7 862	18,9	1,8
Provinces des Prairies	237 232	36 947	2 875	15,6	1,2
Manitoba	49 464	6 346	1 445	12,8	2,9
Saskatchewan	60 916	9 575	1 373	15,7	2,3

8. Les termes « urbain » et « rural » sont définis par la strate géographique de second niveau utilisée pour l'estimation. Voir [Hong \(2023\)](#) pour obtenir plus de renseignements.

Rapport technique sur la couverture

Tableau 6.3.1.1.1

Nombres de logements occupés classés par erreur comme étant des logements inoccupés selon diverses régions géographiques, Recensement de 2021

Région géographique	Nombre de logements inoccupés	Logements occupés classés par erreur comme étant inoccupés			
		Nombre estimé	Erreur-type	Taux estimé ¹ (%)	Erreur-type (%)
Alberta	126 852	21 026	2 005	16,6	1,6
Colombie-Britannique	168 515	31 810	2 774	18,9	1,6
Territoires²	1 389	226	0	16,3	0,0
Yukon	1 016	166	0	16,4	0,0
Territoires du Nord-Ouest	373	59	0	15,9	0,0

1. Le taux est le nombre estimé de logements occupés classés par erreur comme étant inoccupés exprimé en pourcentage de l'ensemble des logements classés comme inoccupés.

2. Le Nunavut n'est pas inclus car il n'existe aucun logement faisant partie du champ de l'enquête dans le territoire.

Source : Statistique Canada, Enquête sur la classification des logements de 2021.

Tableau 6.3.1.1.2

Nombre de logements occupés classés par erreur comme étant des logements inoccupés selon diverses régions géographiques, Recensement de 2016

Région géographique	Nombre de logements inoccupés	Logements occupés classés par erreur comme étant inoccupés			
		Nombre estimé	Erreur-type	Taux estimé ¹ (%)	Erreur-type (%)
Canada	1 187 392	178 219	5 520	15,0	0,5
Urbaine	680 629	145 308	4 819	21,3	0,7
Rurale	506 763	32 911	2 692	6,5	0,5
Provinces de l'Atlantique	144 073	12 577	816	8,7	0,6
Terre-Neuve-et-Labrador	43 770	2 830	310	6,5	0,7
Île-du-Prince-Édouard	10 909	646	117	5,9	1,1
Nouvelle-Écosse	51 940	5 351	633	10,3	1,2
Nouveau-Brunswick	37 454	3 750	395	10,0	1,1
Québec	289 593	41 544	2 240	14,3	0,8
Ontario	378 298	62 536	4 309	16,5	1,1
Provinces des Prairies	204 641	30 779	1 637	15,0	0,8
Manitoba	36 147	4 680	460	12,9	1,3
Saskatchewan	54 414	6 839	1 022	12,6	1,9
Alberta	114 080	19 260	1 193	16,9	1,0
Colombie-Britannique	169 340	30 526	1 881	18,0	1,1
Territoires²	1 447	256	0	17,7	0,0
Yukon	874	143	0	16,3	0,0
Territoires du Nord-Ouest	573	113	0	19,8	0,0

1. Le taux est le nombre estimé de logements occupés classés par erreur comme étant inoccupés exprimé en pourcentage de l'ensemble des logements classés comme inoccupés.

2. Le Nunavut n'est pas inclus car il n'existe aucun logement faisant partie du champ de l'enquête dans le territoire.

Source : Statistique Canada, Enquête sur la classification des logements de 2016.

Rapport technique sur la couverture

Tableau 6.3.1.1.3

Nombre de logements occupés classés par erreur comme étant inoccupés et nombre de personnes qui y vivaient, selon diverses régions géographiques, Recensement de 2021

Région géographique	Logements occupés classés par erreur comme étant inoccupés		Personnes vivant dans des logements occupés classés par erreur comme étant inoccupés	
	Nombre estimé	Erreur-type	Nombre estimé	Erreur-type
Canada	218 042	10 097	406 008	19 153
Urbaine	148 973	5 918	286 311	12 305
Rurale	69 069	5 565	119 698	9 311
Provinces de l'Atlantique	16 504	1 359	29 463	2 321
Terre-Neuve-et-Labrador	3 818	586	6 750	1 052
Île-du-Prince-Édouard	836	127	1 533	241
Nouvelle-Écosse	5 778	743	10 350	1 317
Nouveau-Brunswick	6 073	764	10 830	1 195
Québec	50 477	5 283	81 818	8 503
Ontario	82 077	7 862	157 184	15 138
Provinces des Prairies	36 947	2 875	70 823	5 650
Manitoba	6 346	1 445	12 132	2 710
Saskatchewan	9 575	1 373	18 597	2 946
Alberta	21 026	2 005	40 093	3 966
Colombie-Britannique	31 810	2 774	66 263	6 532
Territoires¹	226	0	457	0
Yukon	166	0	330	0
Territoires du Nord-Ouest	59	0	128	0

1. Le Nunavut n'est pas inclus car il n'existe aucun logement faisant partie du champ de l'enquête dans le territoire.

Source : Statistique Canada, Enquête sur la classification des logements de 2021.

Tableau 6.3.1.1.4

Nombre de logements occupés classés par erreur comme étant inoccupés et nombre de personnes qui y vivaient, selon diverses régions géographiques, Recensement de 2016

Région géographique	Logements occupés classés par erreur comme étant inoccupés		Personnes vivant dans des logements occupés classés par erreur comme étant inoccupés	
	Nombre estimé	Erreur-type	Nombre estimé	Erreur-type
Canada	178 219	5 520	338 246	11 110
Urbaine	145 308	4 819	277 996	9 818
Rurale	32 911	2 692	60 250	5 199
Provinces de l'Atlantique	12 577	816	23 220	1 577
Terre-Neuve-et-Labrador	2 830	310	5 042	623
Île-du-Prince-Édouard	646	117	1 257	248
Nouvelle-Écosse	5 351	633	9 638	1 155
Nouveau-Brunswick	3 750	395	7 283	838
Québec	41 544	2 240	73 306	4 200
Ontario	62 536	4 309	120 951	8 429

Tableau 6.3.1.1.4

Nombre de logements occupés classés par erreur comme étant inoccupés et nombre de personnes qui y vivaient, selon diverses régions géographiques, Recensement de 2016

Région géographique	Logements occupés classés par erreur comme étant inoccupés		Personnes vivant dans des logements occupés classés par erreur comme étant inoccupés	
	Nombre estimé	Erreur-type	Nombre estimé	Erreur-type
Provinces des Prairies	30 779	1 637	58 289	3 684
Manitoba	4 680	460	8 502	906
Saskatchewan	6 839	1 022	12 395	2 134
Alberta	19 260	1 193	37 392	2 863
Colombie-Britannique	30 526	1 881	62 003	4 323
Territoires¹	256	0	477	0
Yukon	143	0	251	0
Territoires du Nord-Ouest	113	0	226	0

1. Le Nunavut n'est pas inclus car il n'existe aucun logement faisant partie du champ de l'enquête dans le territoire.

Source : Statistique Canada, Enquête sur la classification des logements de 2016.

6.3.1.2 Logements hors du parc immobilier classés par erreur comme étant inoccupés

Le [tableau 6.3.1.2](#) indique le nombre estimé de logements classés comme étant inoccupés ne faisant pas partie du parc immobilier ainsi que le taux d'erreur correspondant pour les logements inoccupés, selon diverses régions géographiques. Aucune correction n'a été apportée à la base de données du recensement pour tenir compte des logements hors du parc immobilier qui ont été classés par erreur comme étant inoccupés.

Le dénombrement des logements inoccupés ne faisant pas partie de l'univers des logements se traduit par un surdénombrement des logements. Les logements ne font pas partie de l'univers des logements s'ils sont utilisés à des fins commerciales, s'ils ne sont pas habitables à longueur d'année ou s'ils sont comptés en double au recensement. Cette dernière situation peut se produire lorsque deux adresses semblent associées à un logement ou lorsque deux questionnaires sont retournés par erreur pour un logement qui ne compte plus d'appartement distinct.

Les estimations de l'ECL du nombre de logements ne faisant pas partie du parc immobilier qui ont été classés par erreur comme étant des logements inoccupés n'ont pas été utilisées pour corriger la base de données, du fait du degré de subjectivité associé au fait de classer un logement comme étant habitable à longueur d'année. Pour être considéré comme habitable à longueur d'année, un logement doit être doté d'une source de chauffage ou d'énergie et permettre à ses occupants d'être complètement à l'abri des intempéries. Il est parfois difficile de juger si un logement est habitable; c'est le cas, notamment, lorsque le logement est un chalet, lorsque le logement est en voie de construction et qu'il est presque terminé ou lorsque la condition du logement s'est détériorée. Il convient de noter que les critères permettant de déterminer les logements qui ne font pas partie de l'univers des logements ont été mis à jour depuis le cycle de 2016 pour mieux refléter la définition fournie dans le présent paragraphe. Selon les critères mis à jour, les logements hors du parc immobilier auraient représenté 27,5 % de l'ensemble des logements classés comme étant inoccupés en 2016.

Ce pourcentage était similaire en 2021, alors que les logements hors du parc immobilier représentaient 26,3 % de l'ensemble des logements classés comme étant inoccupés. Parmi les provinces et les territoires, la fréquence des logements hors du parc immobilier qui étaient classés comme étant inoccupés variait de 15,1 % à Terre-Neuve-et-Labrador à 38,0 % au Nouveau-Brunswick. Le problème était plus fréquent dans les régions rurales (27,7 %) que dans les régions urbaines (24,8 %).

Les logements qui ne font pas partie du parc immobilier représentaient 2,1 % de tous les logements privés au Recensement de 2021. Ce taux était le même que le taux d'erreur en 2016 (2,1 %). Parmi les provinces et les territoires, le taux d'erreur variait de 0,8 % dans les Territoires du Nord-Ouest à 4,1 % à l'Île-du-Prince-Édouard.

Tableau 6.3.1.2

Logements hors du parc immobilier et classés par erreur comme étant des logements inoccupés selon diverses régions géographiques, Recensement de 2021

Région géographique	Nombre de logements inoccupés	Logements hors du parc immobilier et classés par erreur comme étant des logements inoccupés			
		Nombre estimé	Erreur-type	Taux estimé ¹ (%)	Erreur-type (%)
Canada	1 259 149	330 777	15 092	26,3	1,2
Urbaine	626 525	155 519	8 518	24,8	1,4
Rurale	632 624	175 258	11 166	27,7	1,8
Provinces de l'Atlantique	132 141	32 440	2 419	24,5	1,8
Terre-Neuve-et-Labrador	44 142	6 686	1 030	15,1	2,3
Île-du-Prince-Édouard	10 345	3 084	496	29,8	4,8
Nouvelle-Écosse	47 584	11 253	1 181	23,6	2,5
Nouveau-Brunswick	30 070	11 417	1 036	38,0	3,4
Québec	285 920	68 354	6 645	23,9	2,3
Ontario	433 952	129 372	10 506	29,8	2,4
Provinces des Prairies	237 232	60 729	5 986	25,6	2,5
Manitoba	49 464	14 586	3 371	29,5	6,8
Saskatchewan	60 916	17 947	2 362	29,5	3,9
Alberta	126 852	28 195	4 169	22,2	3,3
Colombie-Britannique	168 515	39 570	3 517	23,5	2,1
Territoires²	1 389	312	0	22,4	0,0
Yukon	1 016	247	0	24,3	0,0
Territoires du Nord-Ouest	373	65	0	17,4	0,0

1. Le taux est le nombre estimé de logements hors du parc immobilier et classés par erreur comme étant des logements inoccupés exprimé en pourcentage de l'ensemble des logements classés comme inoccupés.

2. Le Nunavut n'est pas inclus car il n'existe aucun logement faisant partie du champ de l'enquête dans le territoire.

Source : Statistique Canada, Enquête sur la classification des logements de 2021.

6.3.2 Logements non répondants

6.3.2.1 Taux d'occupation des logements non répondants

Le [tableau 6.3.2.1.1](#) présente le nombre estimatif et le taux de logements non répondants occupés lors du recensement, par région urbaine et rurale⁹ ainsi que par province et territoire. Le [tableau 6.3.2.1.2](#) présente ces renseignements pour le Recensement de 2016. Le tableau 6.3.2.1.1 indique que 61,5 % de tous les logements classés comme étant non répondants étaient occupés. Ce taux est en légère baisse par rapport au taux d'occupation de 63,1 % enregistré en 2016, quoique la différence ne soit pas significative sur le plan statistique. Le Recensement de 2021 était relativement cohérent en ce qui concerne le classement des logements non répondants dans les régions urbaines (61,2 %) et rurales (62,2 %). À l'échelle des provinces et des territoires en 2021, le taux d'occupation pour les logements non répondants était le plus élevé au Yukon (69,5 %), alors que c'est à Terre-Neuve-et-Labrador qu'il était le moins élevé (41,6 %).

9. Les termes « urbain » et « rural » sont définis par la strate de second niveau infraprovinciale utilisée pour l'estimation. Voir [Hong \(2023\)](#) pour obtenir plus de renseignements.

Rapport technique sur la couverture

Le [tableau 6.3.2.1.3](#) présente le nombre de logements occupés non répondants lors du Recensement de 2021 ainsi que le nombre de personnes prises en compte grâce à l'ECL pour ces logements. Le [tableau 6.3.2.1.4](#) fournit les données pour le Recensement de 2016. En 2021, 421 202 personnes ont été ajoutées au recensement pour 210 370 logements. Le chiffre comparable pour 2016 était de 357 666 personnes pour 179 823 logements.

Tableau 6.3.2.1.1
Logements non répondants occupés selon diverses régions géographiques, Recensement de 2021

Région géographique	Nombre de logements non répondants	Logements non répondants occupés			
		Nombre estimé	Erreur-type	Taux estimé ¹ (%)	Erreur-type (%)
Canada	342 162	210 370	4 198	61,5	1,2
Urbaine	242 381	148 282	3 699	61,2	1,5
Rurale	99 781	62 088	1 801	62,2	1,8
Provinces de l'Atlantique	22 470	13 288	804	59,1	3,6
Terre-Neuve-et-Labrador	5 751	2 392	470	41,6	8,2
Île-du-Prince-Édouard	1 044	671	44	64,3	4,2
Nouvelle-Écosse	8 787	6 094	266	69,4	3,0
Nouveau-Brunswick	6 888	4 131	265	60,0	3,8
Québec	79 340	51 664	1 709	65,1	2,2
Ontario	109 090	64 555	2 826	59,2	2,6
Provinces des Prairies	81 640	47 810	1 914	58,6	2,3
Manitoba	12 948	8 705	367	67,2	2,8
Saskatchewan	13 558	7 632	611	56,3	4,5
Alberta	55 134	31 472	1 761	57,1	3,2
Colombie-Britannique	48 986	32 621	1 030	66,6	2,1
Territoires²	636	434	0	68,2	0,0
Yukon	360	250	0	69,5	0,0
Territoires du Nord-Ouest	276	183	0	66,5	0,0

1. Le taux est le nombre estimé de logements non répondants occupés exprimé en pourcentage de l'ensemble des logements non répondants.

2. Le Nunavut n'est pas inclus car il n'existe aucun logement faisant partie du champ de l'enquête dans le territoire.

Source : Statistique Canada, Enquête sur la classification des logements de 2021.

Tableau 6.3.2.1.2
Logements non répondants occupés selon diverses régions géographiques, Recensement de 2016

Région géographique	Nombre de logements non répondants	Logements non répondants occupés			
		Nombre estimé	Erreur-type	Taux estimé ¹ (%)	Erreur-type (%)
Canada	284 966	179 823	2 525	63,1	0,9
Urbaine	227 692	144 678	2 219	63,5	1,0
Rurale	57 274	35 145	1 205	61,4	2,1
Provinces de l'Atlantique	20 914	11 371	487	54,4	2,3
Terre-Neuve-et-Labrador	4 779	2 637	313	55,2	6,5
Île-du-Prince-Édouard	1 302	802	64	61,6	4,9
Nouvelle-Écosse	7 839	4 199	272	53,6	3,5
Nouveau-Brunswick	6 994	3 733	248	53,4	3,6
Québec	58 039	39 376	1 110	67,8	1,9
Ontario	91 159	58 195	1 536	63,8	1,7

Rapport technique sur la couverture

Tableau 6.3.2.1.2

Logements non répondants occupés selon diverses régions géographiques, Recensement de 2016

Région géographique	Nombre de logements non répondants	Logements non répondants occupés			
		Nombre estimé	Erreur-type	Taux estimé ¹ (%)	Erreur-type (%)
Provinces des Prairies	64 103	38 020	1 264	59,3	2,0
Manitoba	10 811	6 375	586	59,0	5,4
Saskatchewan	11 143	6 453	333	57,9	3,0
Alberta	42 149	25 191	1 070	59,8	2,5
Colombie-Britannique	50 016	32 418	973	64,8	1,9
Territoires²	735	444	0	60,4	0,0
Yukon	405	227	0	56,1	0,0
Territoires du Nord-Ouest	330	217	0	65,7	0,0

1. Le taux est le nombre estimé de logements non répondants occupés exprimé en pourcentage de l'ensemble des logements non répondants.

2. Le Nunavut n'est pas inclus car il n'existe aucun logement faisant partie du champ de l'enquête dans le territoire.

Source : Statistique Canada, Enquête sur la classification des logements de 2016.

Tableau 6.3.2.1.3

Personnes vivant dans des logements non répondants occupés selon diverses régions géographiques, Recensement de 2021

Région géographique	Logements non répondants occupés		Personnes vivant dans des logements non répondants occupés	
	Nombre estimé	Erreur-type	Nombre estimé	Erreur-type
Canada	210 370	4 198	421 202	10 025
Urbaine	148 282	3 699	302 456	8 927
Rurale	62 088	1 801	118 745	4 406
Provinces de l'Atlantique	13 288	804	24 111	1 507
Terre-Neuve-et-Labrador	2 392	470	4 039	766
Île-du-Prince-Édouard	671	44	1 322	109
Nouvelle-Écosse	6 094	266	10 977	626
Nouveau-Brunswick	4 131	265	7 774	559
Québec	51 664	1 709	88 176	3 728
Ontario	64 555	2 826	135 062	6 867
Provinces des Prairies	47 810	1 914	106 973	5 000
Manitoba	8 705	367	21 176	1 179
Saskatchewan	7 632	611	18 042	1 649
Alberta	31 472	1 761	67 755	4 433
Colombie-Britannique	32 621	1 030	66 025	3 556
Territoires¹	434	0	855	0
Yukon	250	0	498	0
Territoires du Nord-Ouest	183	0	357	0

1. Le Nunavut n'est pas inclus car il n'existe aucun logement faisant partie du champ de l'enquête dans le territoire.

Source : Statistique Canada, Enquête sur la classification des logements de 2021.

Tableau 6.3.2.1.4

Personnes vivant dans des logements non répondants occupés selon diverses régions géographiques, Recensement de 2016

Région géographique	Logements non répondants occupés		Personnes vivant dans des logements non répondants occupés	
	Nombre estimé	Erreur-type	Nombre estimé	Erreur-type
Canada	179 823	2 525	357 666	7 800
Urbaine	144 678	2 219	286 286	7 204
Rurale	35 145	1 205	71 380	2 990
Provinces de l'Atlantique	11 371	487	22 686	1 163
Terre-Neuve-et-Labrador	2 637	313	5 054	686
Île-du-Prince-Édouard	802	64	1 459	130
Nouvelle-Écosse	4 199	272	8 651	783
Nouveau-Brunswick	3 733	248	7 523	503
Québec	39 376	1 110	68 744	2 761
Ontario	58 195	1 536	125 071	5 541
Provinces des Prairies	38 020	1 264	78 093	3 226
Manitoba	6 375	586	13 312	1 239
Saskatchewan	6 453	333	12 795	996
Alberta	25 191	1 070	51 986	2 807
Colombie-Britannique	32 418	973	62 089	3 278
Territoires¹	444	0	983	0
Yukon	227	0	452	0
Territoires du Nord-Ouest	217	0	531	0

1. Le Nunavut n'est pas inclus car il n'existe aucun logement faisant partie du champ de l'enquête dans le territoire.

Source : Statistique Canada, Enquête sur la classification des logements de 2016.

6.3.2.2 Logements hors du parc immobilier classés par erreur comme étant non répondants

Le [tableau 6.3.2.2](#) montre les erreurs de classification au Recensement de 2021 relativement aux logements classés par erreur comme étant des logements non répondants, alors qu'ils n'auraient pas dû être compris dans le parc immobilier. La [section 6.3.1.2](#) fournit la définition des logements ne faisant pas partie de l'univers des logements et un commentaire sur la difficulté d'établir si un logement devrait faire partie ou non du parc immobilier. Il convient de noter que les critères utilisés pour déterminer les logements qui ne font pas partie de l'univers des logements ont été mis à jour depuis le cycle de 2016 pour mieux refléter la définition fournie dans cette section. Grâce aux critères mis à jour, les logements non répondants hors du parc immobilier représentaient 7,0 % des logements non répondants en 2016.

À l'échelle nationale, les logements hors du parc immobilier représentaient 7,2 % de tous les logements non répondants en 2021. Le taux d'erreur était plus élevé dans les régions urbaines (7,6 %) que dans les régions rurales (6,2 %). Parmi les provinces et les territoires, la fréquence des logements hors du parc immobilier classés comme étant non répondants variait de 4,5 % à Terre-Neuve-et-Labrador à 15,4 % en Saskatchewan. Toujours à l'échelle nationale, les logements non répondants hors du parc immobilier représentaient 0,2 % de l'ensemble des logements privés. Ce taux était semblable au taux d'erreur de 0,1 % enregistré en 2016. Parmi les provinces et les territoires, le taux d'erreur variait de 0,1 % à 0,4 %.

Rapport technique sur la couverture

Tableau 6.3.2.2

Logements hors du parc immobilier et classés par erreur comme étant des logements non répondants selon diverses régions géographiques, Recensement de 2021

Région géographique	Nombre de logements non répondants	Logements hors du parc immobilier et classés comme étant des logements non répondants			
		Nombre estimé	Erreur-type	Taux estimé ¹ (%)	Erreur-type (%)
Canada	342 162	24 485	1 894	7,2	0,6
Urbaine	242 381	18 327	1 614	7,6	0,7
Rurale	99 781	6 159	875	6,2	0,9
Provinces de l'Atlantique	22 470	1 561	199	6,9	0,9
Terre-Neuve-et-Labrador	5 751	258	53	4,5	0,9
Île-du-Prince-Édouard	1 044	111	32	10,7	3,1
Nouvelle-Écosse	8 787	473	113	5,4	1,3
Nouveau-Brunswick	6 888	719	139	10,4	2,0
Québec	79 340	6 222	865	7,8	1,1
Ontario	109 090	7 051	1 000	6,5	0,9
Provinces des Prairies	81 640	5 929	1 086	7,3	1,3
Manitoba	12 948	757	140	5,8	1,1
Saskatchewan	13 558	2 083	725	15,4	5,3
Alberta	55 134	3 090	679	5,6	1,2
Colombie-Britannique	48 986	3 689	769	7,5	1,6
Territoires²	636	34	0	5,3	0,0
Yukon	360	20	0	5,6	0,0
Territoires du Nord-Ouest	276	13	0	4,9	0,0

1. Le taux est le nombre estimé de logements hors du parc immobilier et classés par erreur comme étant des logements non répondants exprimé en pourcentage de l'ensemble des logements non répondants.

2. Le Nunavut n'est pas inclus car il n'existe aucun logement faisant partie du champ de l'enquête dans le territoire.

Source : Statistique Canada, Enquête sur la classification des logements de 2021.

7. Étude sur le sous-dénombrement du recensement

L'objectif premier de l'Étude sur le sous-dénombrement du recensement (ESoR) est d'estimer le nombre de personnes faisant partie de la population cible du Recensement de 2021 qui n'ont pas été dénombrées lors de celui-ci aux échelons national, provincial et territorial. Un échantillon de personnes a été tiré de six bases de sondage indépendantes du Recensement de 2021. Les données des personnes choisies (PC) ont été appariées aux données fiscales et autres sources administratives afin d'obtenir des renseignements récents au sujet de leur résidence habituelle, de leurs adresses de contact, des membres de leur ménage et des groupes de personnes qui leur sont liés.

Un ensemble d'appariements automatisés complexes et de recherches manuelles ont été effectués pour trouver la PC dans la base de données des réponses (BDR) du Recensement de 2021. Les études sur la couverture du recensement (ECR), dont l'ESoR, ont été réalisées d'après une version de la BDR rendue disponible à la mi-octobre 2021 (c'est-à-dire avant la fin du traitement du recensement). On nomme cette version la BDR-ECR, qui est antérieure à la BDR finale de 2021. Quelques différences mineures existent entre la BDR-ECR et les versions ultérieures des bases de données du recensement. La BDR-ECR, une base de données de personnes, comprend tous les enregistrements des personnes dénombrées, à l'exception de trois groupes d'enregistrements. Le premier concerne les enregistrements du recensement imputés dans le cadre des IME. Le deuxième se rapporte à tous les enregistrements du recensement qui ont été ajoutés tardivement, soit après le début du traitement de l'ESoR, mais cette situation ne s'est pas présentée en 2021 contrairement aux cycles précédents. Le troisième groupe concerne les enregistrements du recensement communément appelés « dénombrements incomplets ». La [section 7.4.6](#) fournit de plus amples détails sur les dénombrements incomplets.

Lorsque la recherche est infructueuse, une collecte multimodale est entreprise en vue de déterminer si la PC fait partie de la population cible, et pour obtenir des renseignements supplémentaires (notamment des adresses) qui aideraient à trouver la PC dans la BDR-ECR. À la fin de la recherche, chaque PC est classée hors champ (décédée, émigrée ou temporairement à l'extérieur du Canada), dénombrée ou omise. Un petit nombre de cas de non-réponse, composé en grande partie de personnes qui n'ont pas pu être dépistées au moyen de la collecte, doivent être traités et servent à rajuster les poids des répondants à l'aide d'un modèle de rajustement de la non-réponse.

7.1 Échantillonnage

La base de sondage visant à couvrir la population cible de l'ESoR, qui inclut toutes les personnes qui auraient dû être dénombrées lors du Recensement de 2021, est formée à partir de six bases indépendantes du Recensement de 2021. Les cinq premières sont utilisées pour la sélection d'un échantillon servant à l'estimation du sous-dénombrement dans les dix provinces, tandis que les estimations pour les trois territoires sont calculées à l'aide des échantillons tirés de la dernière base de sondage seulement.

À l'échelon provincial, on commence par les personnes qui faisaient partie de la population cible du Recensement de 2016. Celles-ci comprennent toutes les personnes dénombrées au Recensement de 2016 et les personnes omises au Recensement de 2016, représentées par la portion de l'échantillon de PC de l'ESoR de 2016 qui ont été classées comme omises. Afin de tenir compte des personnes qui se sont ajoutées à la population cible depuis le recensement précédent, on ajoute les naissances et les immigrants intercensitaires (c.-à-d. entre les recensements de 2016 et de 2021) et les résidents non permanents en date du jour du Recensement de 2021. Les sources de données de ces bases de sondage sont les suivantes :

- Base du recensement : Personnes dénombrées lors du Recensement de 2016 et apparaissant dans la BDR-ECR de 2016.
- Base des personnes omises : Il n'existe pas de liste exhaustive des personnes omises. Toutefois, il existe un échantillon représentatif de ces personnes; il s'agit de l'échantillon de PC de l'ESoR de 2016 classées comme omises. Elles sont toutes incluses dans l'échantillon de 2016 avec leur poids de 2016.
- Base des naissances : Données des statistiques de l'état civil sur les naissances intercensitaires. Comme le fichier final des statistiques de l'état civil sur les naissances n'est disponible que tardivement, l'échantillon des naissances de l'ESoR est tiré d'un mélange de fichiers préliminaires et finaux ainsi que de données brutes des statistiques de l'état civil.

Rapport technique sur la couverture

- Base des immigrants : Données administratives provenant d'Immigration, Réfugiés et Citoyenneté Canada (IRCC) au sujet des immigrants arrivés au Canada pendant la période intercensitaire.
- Base des résidents non permanents : Données administratives provenant d'IRCC sur les personnes qui demandent le statut de réfugié en date du jour du recensement et sur celles qui ont obtenu un permis de travail ou d'études au Canada qui est valide le jour du recensement.

Pour chacun des territoires, les fichiers de l'assurance maladie comprenant les personnes admissibles aux soins de santé le jour du recensement forment la base de sondage principale. Bien que cette base ait une excellente couverture, cette dernière n'est pas complète et un ajustement doit par conséquent être effectué aux poids d'échantillonnage. Chaque base de sondage d'un territoire donné est indépendante des autres bases des territoires et ne sert qu'à faire l'estimation du sous-dénombrement de ce territoire en particulier. De même, les bases de sondage des territoires ne servent pas à l'estimation du sous-dénombrement dans les provinces. Dans le cadre de l'ESoR de 2021, les résidents non permanents dans les territoires qui avaient un permis de travail ou d'études et qui n'étaient pas déjà pris en compte dans les fichiers d'assurance maladie ont été ajoutés aux bases des territoires.

Aucune des cinq premières bases de sondage des provinces ne comprend les personnes émigrées ou à l'extérieur du Canada lors du Recensement de 2016 qui n'ont pas rempli de questionnaire du recensement en 2016 et qui sont revenues durant la période intercensitaire (« Canadiens de retour dans une province »). Selon le questionnaire détaillé du Recensement de 2021, on estime cette population à 252 089 personnes. De même, on estime à 13 426 le nombre de personnes de retour des territoires dans les provinces. À ce nombre s'ajoutent 120 personnes venant d'une réserve ou d'un établissement partiellement dénombré en 2016 et dénombré en 2021, puis 8 489 personnes venant des réserves ou des établissements de retour en 2016 et dénombrées en 2021, mais exclues de la base du Recensement de 2016. De plus, les personnes nées après le Recensement de 2016 à l'extérieur du pays ou dans les territoires qui ont la citoyenneté canadienne et qui sont revenues dans une des 10 provinces du Canada au jour du Recensement de 2021 ne sont pas prises en compte dans les cinq premières bases de sondage de l'ESoR. Selon le questionnaire détaillé du Recensement de 2021, on estime cette population à 16 925 personnes. Les estimations de l'erreur de couverture ne comprennent donc pas ces populations estimées à un total de 291 049 personnes.

Un problème qui se pose en raison de l'utilisation des bases de sondage multiples est la possibilité qu'une même personne soit incluse dans plus d'une base. Par exemple, une personne comprise dans la base des immigrants pouvait être titulaire d'un permis de travail et se trouver au Canada en mai 2016, et ainsi être dénombrable dans le cadre du Recensement de 2016. Elle serait donc à la fois dans la base des immigrants et dans celle du recensement si elle a été dénombrée, ou dans celle des personnes omises si elle a été omise. Ainsi, il est important de déterminer tous les cas de chevauchement des bases, sinon les estimations produites risquent d'être trop élevées puisque les personnes peuvent être comptées deux fois dans les bases de sondage. Autant que possible, ce chevauchement est décelé au moment de la création des bases de sondage, mais une partie est également cernée plus tard à partir des renseignements fournis par les répondants.

Le plan d'échantillonnage varie d'une base à l'autre selon la nature de la liste utilisée. Pour la base du Recensement de 2016, on a utilisé un plan stratifié à un seul degré. La méthodologie de stratification a été modifiée considérablement lors de l'ESoR de 2021. Avant d'effectuer la stratification, plusieurs couplages déterministes ont été effectués. D'abord, il y a eu un couplage de la base avec les données fiscales, et plus de 96 % des personnes ont été liées. Ensuite, il y a eu un couplage avec les fichiers de mortalité des statistiques de l'état civil. Il y a également eu un couplage avec les fichiers d'IRCC afin de trouver les résidents non permanents dans la base. Enfin, il y a eu un couplage avec la BDR de 2021 à l'aide du processus de méga-appariement qui est aussi utilisé pour le traitement de l'échantillon de l'ESoR. Ce processus fournit des suggestions de dénombrement potentiel, et un indicateur de la force de cette suggestion. Certaines suggestions sont assez fortes pour qu'on considère la personne dénombrée sans avoir besoin de vérifier la suggestion. On appelle ces cas des autodénombrements. À la suite de ces couplages, la stratification de la base a été effectuée. Deux strates à tirage complet ont été créées : la strate des personnes décédées et la strate des personnes autodénombrées. Ensuite, six strates à tirage partiel ont été créées en tenant compte de la probabilité de dénombrement des personnes (force de la suggestion dans la BDR de 2021), de la situation fiscale et de la probabilité d'être hors cible au recensement. Cependant, les personnes dénombrées dans les réserves et les établissements lors du Recensement de 2016 forment des strates séparées en utilisant les mêmes critères, mais en regroupant certaines strates étant donné que la population est plus petite et plus homogène.

Ensuite, les strates à tirage partiel ont été stratifiées par province. Pour les personnes demeurant dans les six plus petites provinces en 2016, la province de stratification correspondait à la province de résidence en 2016 (dans la BDR de 2016). Pour les personnes dans les quatre plus grandes provinces en 2016, la dérivation de la province de stratification variait selon la strate. Dans les strates à fortes probabilités de dénombrement dans la BDR en 2021, la province de dénombrement potentiel dans la BDR de 2021 a été utilisée. Sinon, dans le cas où la personne était liée aux données fiscales, c'est la province de résidence la plus récente selon ces données qui a été utilisée. En dernier recours, c'est la province figurant dans la BDR de 2016 qui a été utilisée.

La base des personnes omises est une base échantillonnale, car il n'existe pas de liste de toutes les personnes omises lors du Recensement de 2016. L'échantillon de cette base est constitué de tous les cas classés comme « omis » dans le cadre de l'ESoR de 2016. Quoique l'échantillon ne soit pas stratifié comme tel, une stratification implicite est inévitable puisque les cas omis en 2016 proviennent de bases et de strates différentes.

Pour créer la base de sondage des naissances, on a obtenu une copie des enregistrements de naissances intercensitaires des statistiques de l'état civil par l'intermédiaire du Système national d'acheminement, qui permet d'avoir accès à ces données plus rapidement. La base contient toutes les naissances entre le 10 mai 2016 et le 10 mai 2021 inclusivement. La base a ensuite été stratifiée selon la province de résidence de la mère, ou selon la province de naissance si cette donnée n'était pas disponible.

On a créé la base de sondage des immigrants à partir des dossiers d'IRCC. La base des immigrants contient toutes les personnes immigrées au pays entre le 10 mai 2016 et le 10 mai 2021 inclusivement. On a soustrait de la base des immigrants de 2016 ceux qui étaient résidents non permanents au jour du Recensement de 2016, ceux-ci étant déjà compris dans la base du Recensement de 2016 ou la base des personnes omises de 2016. La base des immigrants a été stratifiée par province. La province a été dérivée selon les renseignements disponibles dans un fichier d'adresses fourni par IRCC et dans le fichier d'immigration d'IRCC. La province de résidence la plus probable au jour du Recensement de 2021 a été retenue. Ensuite, les immigrants de toutes les provinces ont été séparés en deux strates selon leur date d'immigration. Les immigrants arrivés entre le 10 mai 2016 et le 30 avril 2020 ont formé la première strate, tandis que ceux arrivés entre le 1^{er} mai 2020 et le 10 mai 2021 ont formé la seconde strate, ce qui s'explique par le fait que les immigrants arrivés depuis moins longtemps ont habituellement un taux d'omission plus élevé au recensement.

La base des résidents non permanents (titulaires d'un permis de travail ou d'études et demandeurs du statut de réfugié) a été créée à partir des dossiers d'IRCC. Les résidents non permanents au jour du Recensement de 2016 et les immigrants intercensitaires ont été enlevés de la base des résidents non permanents de 2021. La base a été stratifiée par province, selon la province de résidence la plus probable au jour du Recensement de 2021. Pour ce faire, un couplage déterministe de la base a été effectué aux données fiscales. Le fichier d'adresses d'IRCC et les différents fichiers de résidents non permanents d'IRCC ont aussi été utilisés. À la fin du processus, un certain nombre de résidents non permanents n'avaient aucune province de résidence qui leur était associée (résidents ayant un permis ouvert); ceux-ci ont donc été placés dans une strate nationale.

Dans les provinces, la taille totale de l'échantillon de 2021 a été déterminée de façon à atteindre deux objectifs principaux. Premièrement, le budget de collecte de l'ESoR de 2021 devait demeurer le même que celui de l'ESoR de 2016 (mais ajusté en fonction des augmentations de coût à l'unité entre 2016 et 2021). Ce n'est qu'une partie des personnes dans l'échantillon qui nécessite une collecte, et les proportions varient selon les bases et les strates. Deuxièmement, l'ESoR visait à obtenir des erreurs-types du taux de sous-dénombrement semblables entre les provinces de taille comparable. On visait des erreurs-types plus petites pour les plus grandes provinces que pour les petites provinces, car cela permettait d'obtenir une petite erreur-type à l'échelon national. Autant que possible, les erreurs-types ne devaient pas être plus élevées que celles obtenues en 2016.

Ainsi, à partir de 2020, en mettant constamment à jour les paramètres entrant dans le calcul de l'erreur-type du sous-dénombrement et le calcul du nombre de personnes nécessitant une collecte, on a fait des simulations de tailles d'échantillons par base de sondage et strate afin d'obtenir les erreurs-types voulues à tous les niveaux (national, provincial, âge et genre). Les bases et les résultats de l'ESoR de 2016 ont servi à faire ces simulations. Étant donné que certaines bases de sondage étaient prêtes avant d'autres, les tailles d'échantillons ont été déterminées pour ces bases avant d'établir les tailles pour les autres bases et strates. Entre autres, la taille d'échantillon de la strate de la base des personnes omises en 2016 était déjà fixée, toutes les personnes considérées comme des personnes omises dans le cadre de l'ESoR de 2016 ayant été choisies. Ensuite, la taille

Rapport technique sur la couverture

de la première strate de la base des immigrants a été déterminée à l'été 2020, et ainsi de suite pour les autres strates et bases (naissances et résidents non permanents). La répartition de l'échantillon a été achevée en novembre 2021 avec la stratification de la base du Recensement de 2016 comme décrite ci-dessus.

Dans plusieurs strates, une taille totale était déterminée pour l'ensemble des dix provinces, puis une répartition par la méthode de puissance a été utilisée afin de répartir l'échantillon total dans les provinces. Des tailles d'échantillons minimales étaient aussi fixées dans les plus petites provinces.

De plus, pour certaines strates de la base du recensement, une sous-stratification par sexe et groupe d'âge a été effectuée afin de s'assurer d'avoir un nombre suffisant de personnes omises à l'échelle de ces domaines. De même, la répartition de l'échantillon aux strates des réserves de la base du recensement a été effectuée de façon à obtenir une précision du sous-dénombrement dans les réserves au moins aussi bonne qu'à l'ESoR de 2016. L'échantillon total final réparti était de 32 534 PC parmi les bases dans les provinces. Le [tableau 7.1.1](#) présente la répartition finale de l'échantillon selon la strate pour toutes les provinces. Selon cette répartition de l'échantillon, à l'échelle provinciale, les erreurs-types visées pour le taux de sous-dénombrement variaient entre 0,16 % et 0,42 %, et pour l'ensemble des provinces, de 0,09 %. Il convient de souligner que la répartition résultante ne garantit pas que ce niveau de précision sera nécessairement obtenu étant donné que des hypothèses ont été avancées à propos de plusieurs paramètres entrant dans le calcul de l'erreur-type du sous-dénombrement (tailles des strates et bases, taux de personnes omises, taux de réponse à la collecte de l'ESoR, etc.). De plus, les effets de la pandémie de COVID-19 ont pu affecter la justesse de ces hypothèses, notamment en ce qui concerne le nombre d'immigrants et de résidents non permanents, la migration interprovinciale et les taux d'omission au Recensement de 2021.

Tableau 7.1.1

Répartition de l'échantillon, bases de sondage et strates pour toutes les provinces

Base de sondage	Strate pour chaque province	Nombre de personnes
Total du tirage complet	...	26 944 027
Recensement de 2016	Décédé	1 239 662
	Auto-dénombré dans une province	25 704 365
Total du tirage partiel	...	32 534
Recensement de 2016	Hors réserve TP_1 : fortes suggestions de dénombrement	5 559
	Hors réserve TP_2 : fortes suggestions de dénombrement incomplet	369
	Hors réserve TP_3 : probabilité élevée d'être hors champ	510
	Hors réserve TP_4 : suggestions modérées de dénombrement	757
	Hors réserve TP_5 : probabilité élevée d'être omise	5 041
	Hors réserve TP_6 : autres	1 712
	Dans une réserve TP_7 : suggestions fortes ou modérées de dénombrement	270
	Dans une réserve TP_8 : probabilité élevée d'être omise	505
	Dans une réserve TP_9 : autres	200
	Dans une réserve TP_10 : Terre-Neuve-et-Labrador et Île-du-Prince-Édouard	60
Personnes omises de 2016	Aucune autre stratification	4 821
Naissances	Aucune autre stratification	5 978
Immigrants	Du 10 mai 2016 au 30 avril 2020	2 593
	Du 1 ^{er} mai 2020 au 10 mai 2021	588
Résidents non permanents	Aucune autre stratification	3 571

... n'ayant pas lieu de figurer

TP = tirage partiel

Source : Statistique Canada, Étude sur le sous-dénombrement du Recensement de 2021.

Rapport technique sur la couverture

La méthodologie d'échantillonnage pour les territoires était semblable à celle de la base du recensement pour les provinces. Les personnes comprises dans la base de sondage de chacun des territoires ont été appariées aux données fiscales et ensuite à la BDR de 2021, à l'aide du processus de méga-appariement qui est aussi utilisé pour le traitement de l'échantillon de l'ESoR (voir la [section 7.2.1](#)). À la suite de ces étapes, la stratification de la base a été effectuée en tenant compte de la force du couplage avec la BDR de 2021, de l'endroit du dénombrement et de l'activité fiscale récente. Une strate à tirage complet d'autodénombrements dans le territoire a été formée, et six strates à tirage partiel ont été formées (voir le [tableau 7.1.2](#)). Pour la première et la sixième strate, une sous-stratification par sexe et trois groupes d'âge (0 à 17 ans, 18 à 29 ans, 30 ans et plus) a été effectuée.

En ce qui concerne la répartition de l'échantillon pour les territoires, la première étape consistait à déterminer l'échantillon total à allouer à chaque territoire afin d'atteindre une précision semblable et adéquate du sous-dénombrement. En 2021, l'erreur-type visée pour le taux de sous-dénombrement était d'environ 0,40 % au Yukon et dans les Territoires du Nord-Ouest (une amélioration par rapport à 2016), et de 0,60 % au Nunavut (semblable à 2016). En utilisant les résultats de l'ESoR de 2016, des hypothèses de taux d'omission, de taux de dénombrement et autres ont été calculées pour chaque strate. Pour la première strate à tirage partiel, la taille d'échantillon a été fixée manuellement dans chaque territoire étant donné que cette strate avait très peu d'effet sur la précision du taux de sous-dénombrement, mais plus d'incidence sur la précision du taux de dénombrement. Cela a une importance pour le calcul d'un facteur de calage au moment de la pondération. De plus, il fallait tenir compte du fardeau de travail des employés qui devaient vérifier l'échantillon de cette strate. De même, un échantillon a été fixé manuellement pour la quatrième strate, car elle représentait des personnes qui sont presque assurément hors cible, mais qui font l'objet d'un certain travail de recherche par les employés de l'ESoR. Ensuite, de façon itérative, on a effectué une répartition optimale de l'échantillon total entre les autres strates à tirage partiel, y compris les six sous-strates de la dernière strate. Une taille totale approximative a été fixée au départ, puis la précision de la répartition optimale était calculée, et cela a été répété en augmentant ou en diminuant la taille totale jusqu'à l'obtention de la précision voulue pour le taux de sous-dénombrement dans chaque territoire. L'échantillon total final réparti était de 4 285 PC choisies parmi les bases dans les territoires.

Le tableau 7.1.2 présente la répartition selon la strate pour tous les territoires.

Tableau 7.1.2
Répartition de l'échantillon, strates selon le territoire

Strate	Territoires du Nord-Ouest			Total
	Yukon	Nunavut		
Tirage complet : auto-dénombré dans son territoire	27 881	26 696	16 981	71 558
Total du tirage partiel	1 156	1 331	1 798	4 285
TP_1 : fortes suggestions de dénombrement	530	440	468	1 438
TP_2 : suggestions modérées de dénombrement	57	196	356	609
TP_3 : fortes suggestions de dénombrement incomplet	30	30	44	104
TP_4 : fortes suggestions de dénombrement hors de son territoire	53	78	70	201
TP_5 : probabilité élevée d'être hors champ	97	83	96	276
TP_6 : probabilité élevée d'être omise (sous-stratification)				
Femmes, 0 à 17 ans	30	59	158	247
Femmes, 18 à 29 ans	48	44	69	161
Femmes, 30 ans et plus	109	117	157	383
Hommes, 0 à 17 ans	33	61	132	226
Hommes, 18 à 29 ans	54	61	65	180
Hommes, 30 ans et plus	115	162	183	460

TP = tirage partiel

Source : Statistique Canada, Étude sur le sous-dénombrement du Recensement de 2021.

Rapport technique sur la couverture

Le tableau 7.1.3 présente la répartition de l'échantillon pour le Canada, les provinces et les territoires.

Tableau 7.1.3
Taille de l'échantillon pour le Canada, les provinces et les territoires

Province ou territoire	Strate à tirage complet (nombre de personnes)	Strate à tirage partiel (nombre de personnes)
Canada	27 015 585	36 819
Toutes les provinces	26 944 027	32 534
Terre-Neuve-et-Labrador	393 554	1 551
Île-du-Prince-Édouard	106 063	1 437
Nouvelle-Écosse	696 275	1 943
Nouveau-Brunswick	579 964	1 680
Québec	6 668 208	4 298
Ontario	10 415 555	7 126
Manitoba	947 750	2 579
Saskatchewan	794 538	2 540
Alberta	2 940 437	4 215
Colombie-Britannique	3 401 683	5 015
RNP-CA	0	150
Tous les territoires	71 558	4 285
Yukon	27 881	1 156
Territoires du Nord-Ouest	26 696	1 331
Nunavut	16 981	1 798

RNP-CA = résidents non permanents dont la province est inconnue

Source : Statistique Canada, Étude sur le sous-dénombrement du Recensement de 2021.

Une méthode d'échantillonnage systématique à l'intérieur des strates a été utilisée pour sélectionner les échantillons. Voici la liste des variables de tri utilisées pour obtenir un échantillon efficace (stratification implicite), classées selon la base de sondage :

- Base du Recensement de 2016 : sexe, âge, code M¹⁰, géographie de 2016, situation fiscale, raison pour être potentiellement hors cible et province probable en 2021 (si stratifiée dans les six plus petites provinces);
- Base des naissances : âge au jour du recensement, sexe, groupe d'âge de la mère et code postal;
- Base des immigrants : groupe d'âge, sexe et pays de naissance;
- Base des résidents non permanents : type de permis, groupe d'âge, sexe et pays de naissance;
- Base des territoires : sexe, âge, code M, situation fiscale et municipalité de résidence.

Aucun échantillonnage n'était requis pour la base des personnes omises de 2016, car toutes les personnes omises de l'ESoR de 2016 ont été sélectionnées dans l'échantillon de l'ESoR de 2021.

Après la sélection des échantillons provinciaux et territoriaux, ceux-ci doivent être préparés en vérifiant la qualité des renseignements pour les différentes variables d'intérêt, soit les variables géographiques et démographiques. On procède, par exemple, à la vérification de la qualité des noms et à la validation des dates de naissance.

10. Le code M est un code associé à chaque suggestion de dénombrement obtenue à l'aide du méga-appariement, en fonction de la force du lien obtenu.

Les adresses sont normalisées afin de faciliter les activités de traitement ultérieures. Afin de mettre à jour les renseignements géographiques, en particulier pour les échantillons du recensement et des personnes omises dont l'information date de 2016, ceux-ci sont appariés aux dossiers de l'Agence du revenu du Canada (ARC), dont ceux de l'impôt sur le revenu des particuliers de 2015 à 2021 et des prestations canadiennes pour enfants de 2016 à 2022. À l'aide des fichiers de l'ARC et des données des statistiques de l'état civil, on vérifie également si des personnes sélectionnées sont décédées. Cette étape de préparation est très importante, car elle aide à déterminer les personnes qui sont dénombrées dans les bases du recensement et à communiquer avec les personnes non trouvées en vue de les interviewer.

7.2 Traitement et classification

7.2.1 Traitement

L'objectif du traitement est de fournir des renseignements sur la classification des PC en vue de l'estimation et de l'ajustement pour la non-réponse. Ainsi, le traitement vise à :

- déterminer si les PC sont dénombrées au sein de la BDR du recensement;
- déterminer si les PC font partie de la population cible du recensement;
- fournir davantage de renseignements sur l'ajustement pour la non-réponse.

Les résultats du traitement ont été enregistrés dans une classification attribuée à chaque PC aux fins d'estimation et de totalisation (voir la [section 7.4](#) et la [section 9](#)).

La majeure partie du travail de traitement a fait intervenir une recherche informatisée automatisée dans la version des études sur la couverture du recensement de la BDR du Recensement de 2021 (BDR-ECR), afin de déterminer si la PC était dénombrée ou non.

Divers renseignements ont été utilisés pour cette recherche, notamment les noms et prénoms ainsi que les dates de naissance. Les numéros de téléphone et les adresses associés à la PC ou à des membres de son ménage ont également été utilisés. Des questionnaires où la personne aurait pu être dénombrée ont été relevés à partir de sources variées, notamment :

- des appariements avec la BDR-ECR, en utilisant la date de naissance et le sexe de la PC et des membres de son ménage, ou le nom, le code postal ou le numéro de téléphone de la PC;
- des adresses de sélection provenant de la base de sondage;
- des mises à jour des adresses provenant de dossiers fiscaux;
- des renseignements de l'interview téléphonique assistée par ordinateur (ITAO) [voir la [section 7.3](#)].

La première étape après la préparation de l'échantillon a été d'effectuer une recherche de chaque PC dans la BDR-ECR en traitant toutes les PC dont l'adresse figurait dans la base de sondage et les données fiscales. Deux résultats ont été obtenus. Lorsque les PC ont été trouvées, elles ont généralement été classées comme « dénombrées » et aucun autre traitement n'a été requis, à l'exception des PC identifiées ensuite, au moyen des données des statistiques de l'état civil, comme décédées avant le recensement. Les cas des PC non trouvées ont été envoyés pour faire l'objet d'une collecte. Pendant la collecte, la recherche dans la BDR-ECR s'est poursuivie. Lorsque les données de l'ITAO étaient disponibles, les chercheurs ont pu déterminer si chaque PC faisait ou non partie de la population cible du recensement. Dans l'affirmative, les données de l'ITAO permettaient parfois d'approfondir les recherches.

La recherche des PC a été menée à la fois automatiquement et manuellement par du personnel de codage dirigé par des spécialistes du domaine. Un manuel de procédures très détaillées indiquant les étapes précises de codage des résultats de recherche a été fourni au personnel de codage pour assurer l'uniformité du codage. Les recherches automatisées ont d'abord eu lieu. Un questionnaire du recensement existait pour les adresses obtenues grâce à une correspondance avec la BDR-ECR. On a calculé une mesure de similarité entre le

questionnaire du recensement et les données disponibles pour l'enquête. Lorsque cette mesure dépassait un seuil préétabli, on concluait automatiquement que la PC avait été dénombrée à cette adresse. Dans ces cas, ni cette adresse ni les autres adresses de la PC n'avaient à être traitées par le personnel de codage. Des programmes informatiques ont également déterminé lorsqu'une adresse était un double d'une autre. Ces adresses en double n'avaient pas non plus à être traitées.

Pour d'autres cas, on a eu recours à un couplage manuel à l'aide de l'Application de vérification interactive de DoLink (AVID), application créée précisément pour cette opération. Le personnel de codage a utilisé, en outre, plusieurs outils dans le cadre de ce traitement, tels que des fichiers de référence géographique, des annuaires téléphoniques électroniques et le fichier des attributs des rues. Des questionnaires du recensement ou des unités de collecte du recensement suggérés correspondaient souvent à l'adresse utilisée à la première étape de la recherche. Le personnel pouvait également effectuer une recherche dans la BDR-ECR à l'aide de paramètres souples aux étapes ultérieures du traitement (recherche par nom, date de naissance, etc.). Les résultats de la recherche manuelle étaient alors modifiés manuellement par des corrections intégrées à l'AVID afin de limiter les erreurs. Un fichier comprenant les résultats de recherche a ensuite été produit. Les données de ce fichier ont été utilisées pour classer les PC.

7.2.2 Classification

Le traitement fournit les renseignements nécessaires en vue de déterminer si les PC sont :

- incluses dans la « population cible du recensement » ou « hors champ » (non incluses);
- « classifiées » ou « non classifiées »;
- « listées » ou « non listées »;
- « identifiables » ou « non identifiables »;
- « dénombrées »;
- « omises ».

Certaines PC figurent dans plus d'une catégorie. On en expliquera les raisons plus en détail dans cette section.

7.2.2.1 Classifications « population cible » ou « hors champ »

La « population cible du recensement » comprend le groupe de personnes mentionné à la [section 2.2](#). Une PC est considérée « hors champ » si elle ne fait pas partie de la population cible du recensement. Chaque PC classée comme « hors champ » se voit attribuer l'un des états suivants pour cette classification : décédée, émigrée ou représentée par une autre base de sondage. Pour qu'une personne soit classée comme étant décédée, elle doit apparaître comme décédée dans au moins deux sources administratives (les fichiers de mortalité des statistiques de l'état civil, les fichiers d'impôts, les fichiers nécrologiques), ou sinon lors de l'interview de collecte de l'ESoR. Les émigrants permanents ou temporaires ont également été déterminés au moyen d'une interview de collecte en fonction de certains critères et des réponses fournies concernant leur lieu de résidence le jour du recensement, la période de temps passé à l'extérieur du Canada, leur intention de revenir vivre au pays et la raison pour laquelle ils étaient à l'extérieur du pays le jour du recensement. D'autres PC sont également classées comme « émigrées listées », qu'elles soient ou non des répondants lors de la collecte. Il s'agit des résidents non permanents (de la base des personnes omises et de la base du Recensement de 2016) qui n'avaient plus de permis de travail ou d'études en 2021 ni le statut d'immigrant depuis 2016.

Les PC classées comme « représentées par une autre base de sondage » comprennent notamment les cas sélectionnés dans une province, mais classés dans l'un des trois territoires. De même, les cas sélectionnés dans un territoire, mais classés dans une province ou un autre territoire, sont eux aussi classés comme « représentés par une autre base de sondage ».

Les PC classées dans la population cible du recensement sont soit « dénombrées », soit « omises », soit « non classifiées » (voir la [section 7.2.2.2](#)). Une PC est considérée comme « dénombrée » si elle figure dans la BDR-ECR. La classification « omises » est attribuée aux PC dans la population cible du recensement qui ne sont pas dénombrées ni « non classifiées ».

7.2.2.2 Classification pour la non-réponse et le rajustement de la non-réponse

L'attribution des classifications « listées » et « non classifiées » dépend de l'utilité des adresses fournies et des renseignements tirés de l'ITAO. Dans plusieurs cas, la collecte fournit des renseignements ainsi qu'une ou plusieurs adresses qui sont introuvables à partir d'autres sources. Dans d'autres cas, il est possible de trouver toutes les adresses et tous les renseignements obtenus au moyen de la collecte à partir d'autres sources.

Une PC est « listée » si elle a été classée sans le recours aux données provenant de l'ITAO; même si la collecte des données a été effectuée, les adresses et les renseignements obtenus au moyen de l'interview n'ont pas été requis.

Une personne est considérée comme « non classifiée » si l'on peut déterminer qu'elle fait partie de la population cible, mais qu'on ne peut déterminer si elle a été omise ou non. Cette situation survient lorsque le lieu de résidence le jour du recensement, tel que défini à la [section 2.4](#), est connu, mais non déterminé dans la BDR-ECR. Les personnes dont le lieu de résidence le jour du recensement n'est pas assez précis (p. ex. si seul le nom d'une grande ville est fourni comme lieu de résidence le jour du recensement) et les personnes sans domicile fixe ont été incluses dans cette catégorie.

Les PC pour lesquelles il est impossible de déterminer une ou plusieurs des caractéristiques de la liste présentée plus haut sont considérées comme des non-répondants. Il existe trois types de non-répondants :

- Une PC est « non identifiée » lorsqu'il est impossible de déterminer si elle est listée ou non. Autrement dit, comme les renseignements disponibles sur la PC sont incomplets, il est impossible de procéder ni à son appariement avec la BDR-ECR ni à la collecte de ses renseignements au moyen d'une interview.
- Une PC est « non dépistée » lorsqu'il est impossible de déterminer si elle est incluse ou non dans la population cible du recensement.
- Une PC « non classifiée » est considérée comme un cas de non-réponse partielle. En effet, on sait que la personne fait partie de la population cible, mais on ignore si elle est omise ou dénombrée.

7.2.2.3 Répartition de l'échantillon selon la classification

Le [tableau 7.2](#) montre la répartition de l'échantillon selon la classification et la base de sondage. Ce tableau exclut les personnes dans les strates à tirage complet étant donné que ces personnes ont été classées (dénombrées ou décédées) avant la sélection de l'échantillon. La classification est déterminée à partir de combinaisons précises des caractéristiques de la liste présentée plus haut. Au départ, un échantillon total de 36 819 PC a été sélectionné dans les provinces et les territoires. De ce nombre, 22 083 PC ont été classées comme « dénombrées », 7 453, comme « omises » et 5 171, comme des non-répondants, dont 169 ayant été classées comme « non classifiées ». Les 2 112 autres PC ont été classées comme « hors champ », dont 583 personnes « décédées », 938 « émigrants » (permanents ou temporaires), 405 personnes hors de l'univers des territoires ou hors de l'univers des provinces, et 186 personnes, pour d'autres raisons. Un rajustement de la non-réponse a été effectué lors du processus d'estimation (voir la [section 7.4](#)). Il est important de noter qu'aux fins de la classification et, par conséquent, de l'estimation, la définition d'un non-répondant diffère de la définition habituelle d'un non-répondant, selon laquelle la collecte de données est tentée, mais non terminée. Il en est ainsi parce que pour procéder à la classification, on utilise des données provenant de plusieurs sources, dont la collecte. Afin d'éviter toute confusion, la [section 7.3](#) portant sur la collecte fait référence à la « collecte terminée » plutôt qu'à la « réponse ».

7.2.2.4 Implications de la classification

Les PC « dépistées » sont les PC pour lesquelles il est possible de déterminer si elles sont incluses dans la population cible du recensement ou non. Aux fins d'estimation et de totalisation, les PC dépistées qui sont également classifiées sont les répondants. Puisque les noms, y compris ceux des membres du ménage, et les adresses sont disponibles dans la BDR-ECR, et que les outils pour consulter la base de données sont suffisamment puissants, il est possible de vérifier si une PC est dénombrée à une adresse, même si l'adresse fournie est vague.

Rapport technique sur la couverture

L'utilité de savoir si une PC est dénombrée va de soi. Les PC qui font partie de la population cible du recensement et qui ne sont pas dénombrées, et donc classées comme « omises », sont la base de l'estimation du sous-dénombrement. On voulait également classer les PC en se fondant sur les caractéristiques mentionnées précédemment afin de choisir les répondants les plus appropriés pour représenter les non-répondants.

Enfin, sauf pour les PC non classifiées, on a aussi déterminé l'adresse au jour du recensement (lieu habituel de résidence) de chaque PC dans la population cible du recensement. Cette adresse représente celle où, selon les directives relatives au recensement, la PC devrait avoir été dénombrée. Si la PC a été dénombrée, l'adresse du dénombrement sera considérée comme l'adresse au jour du recensement, en dépit d'autres renseignements fournis qui pourraient laisser croire que les directives relatives au recensement n'ont pas été bien comprises.

Pour obtenir de plus amples renseignements sur le traitement et la classification, veuillez consulter [Parenteau \(2023\)](#).

Tableau 7.2
Classification des personnes choisies, bases de sondage pour le Canada

Classification	Strate provinciale									
	Recensement de 2016 ¹		Personnes omises de 2016		Naissances		Immigrants		Résidents non permanents	
	nombre	%	nombre	%	nombre	%	nombre	%	nombre	%
Total	14 983	100,0	4 821	100,0	5 978	100,0	3 181	100,0	3 571	100,0
Dénombrées	7 354	49,1	3 122	64,8	5 210	87,2	2 610	82,0	2 015	56,4
Listées	7 201	48,1	3 112	64,6	5 206	87,1	2 604	81,9	1 997	55,9
Non listées	153	1,0	10	0,2	4	0,1	6	0,2	18	0,5
Omises	4 156	27,7	710	14,7	432	7,2	284	8,9	630	17,6
Listées	821	5,5	86	1,8	68	1,1	22	0,7	49	1,4
Non listées	3 335	22,3	624	12,9	364	6,1	262	8,2	581	16,3
Hors champ	882	5,9	433	9,0	102	1,7	100	3,1	188	5,3
Listées	505	3,4	327	6,8	79	1,3	10	0,3	104	2,9
Non listées	377	2,5	106	2,2	23	0,4	90	2,8	84	2,4
Non-réponse	2 591	17,3	556	11,5	234	3,9	187	5,9	738	20,7
Dépistées, non classifiées	87	0,6	17	0,4	17	0,3	2	0,1	10	0,3
Identifiées, non dépistées	2 492	16,6	539	11,2	217	3,6	185	5,8	728	20,4
Non identifiées	12	0,1	0	0,0	0	0,0	0	0,0	0	0,0

Tableau 7.2
Classification des personnes choisies, bases de sondage pour le Canada

Classification	Strate territoriale		Total	
	Bases de sondage territoriales ¹		nombre	%
	nombre	%		
Total	4 285	100,0	36 819	100,0
Dénombrées	1 772	41,4	22 083	60,0
Listées	1 760	41,1	21 880	59,4
Non listées	12	0,3	203	0,6
Omises	1 241	29,0	7 453	20,2
Listées	238	5,6	1 284	3,5
Non listées	1 003	23,4	6 169	16,8
Hors champ	407	9,5	2 112	5,7
Listées	293	6,8	1 318	3,6
Non listées	114	2,7	794	2,2
Non-réponse	865	20,2	5 171	14,0
Dépistées, non classifiées	36	0,8	169	0,5
Identifiées, non dépistées	829	19,3	4 990	13,6
Non identifiées	0	0,0	12	0,0

1. Excluant les strates à tirage complet.

Source : Statistique Canada, Étude sur le sous-dénombrement du Recensement de 2021.

7.3 Collecte

7.3.1 Aperçu

Le personnel du bureau central d'Ottawa a travaillé étroitement avec le personnel des bureaux régionaux (BR) de Statistique Canada pour recueillir les données à l'étape de l'Étude sur le sous-dénombrement du recensement (ESoR) de l'enquête. Les suggestions et recommandations formulées par les BR après l'ESoR de 2016 ont été prises en considération aux fins de la conception et du déroulement de l'enquête de 2021.

Le but principal de l'ESoR est de trouver (retracer) les personnes choisies (PC) pertinentes et d'en recueillir les données démographiques et les coordonnées, afin de pouvoir les classer comme dénombrées, omises ou hors du champ d'observation pour le recensement. Les résultats de cette classification sont utilisés pour estimer le nombre de personnes ayant été omises, ou sous-dénombrées, dans le recensement. Pour aider à trouver et classer les PC, l'adresse le jour du recensement et la composition du ménage ont été recueillies, ainsi que toute autre adresse à laquelle la PC peut avoir été dénombrée. D'autres renseignements, comme la langue maternelle de la PC, ont également été recueillis pour les tableaux de l'étude sur la couverture.

La taille de l'échantillon du tirage partiel de l'ESoR était de 36 819 (la [section 7.1](#) décrit le plan de sondage). Au cours du traitement de précollecte, on a tenté de trouver ces cas dans la BDR-ECR, dans les statistiques de l'état civil et dans d'autres fichiers administratifs. Les cas correspondants ou trouvés dans ces fichiers, et pouvant donc être classés comme dénombrés ou décédés avant le jour du recensement, n'ont pas été envoyés à la collecte. Tous les autres cas non classés ont été envoyés à la collecte. Le nombre total de cas envoyés à la collecte (la taille de l'échantillon de la collecte) a été de 13 096. Au cours de la période de collecte, l'équipe de traitement a continué d'essayer de faire correspondre certains des cas; ceux pouvant être classés ont été retirés de la collecte (voir le [tableau 7.3.2](#) pour connaître ces chiffres).

Par définition, la collecte avait lieu par répondant substitut pour les PC âgées de moins de 18 ans. On a aussi eu recours à des répondants substituts lorsque la PC n'était pas disponible pendant la période de collecte ou lorsqu'elle était difficile à joindre. Globalement, 34 % des cas traités l'ont été par répondant substitut et le pourcentage de cas par répondant substitut complétés par des intervieweurs a été supérieur à celui des cas effectués par autodéclaration.

Dans le cas des PC décédées, il était important d'établir si cette personne était décédée avant, après ou le jour du recensement, puisque différents cheminements de questionnaire étaient alors utilisés, selon la date du décès. Dans certains cas, il a été déterminé avant la collecte que la PC était décédée avant le jour du recensement (p. ex. en faisant correspondre des dossiers fiscaux et des statistiques de l'état civil). Ces cas n'ont alors pas été envoyés à la collecte. Cependant, en cas de doute, le cas était envoyé avec une note signalant que la PC était peut-être décédée.

Il était impératif que la PC pertinente (ou un répondant substitut pour cette personne) soit interrogée. En cas de collecte de données sur la mauvaise personne, la classification correspondante obtenue serait incorrecte. Le système d'interview téléphonique assistée par ordinateur (ITAO) a été conçu pour indiquer aux intervieweurs de vérifier que la personne interrogée était la PC pertinente au début de l'interview. En cas d'interview effectuée avec une autre personne que la PC (p. ex. personne au nom et à la date de naissance similaires), le cas était retourné à la collecte pour être effectué auprès de la personne pertinente.

L'ESoR est une enquête plurimodale obligatoire. Le mode principal de collecte de données est l'ITAO; le mode secondaire est l'autodénombrement. Pour 2021, l'ESoR a eu recours à des questionnaires électroniques en ligne pour les deux modes, puisqu'elle est passée au Système intégré de collecte et des opérations (SICO), qui est une application de collecte normalisée développée à Statistique Canada. Auparavant, le mode d'autodéclaration de l'ESoR avait recours à des questionnaires papier. La transition vers un questionnaire électronique a constitué une grande amélioration, puisqu'elle a réduit le fardeau du répondant ainsi que le temps d'opération et les coûts associés à l'envoi des questionnaires papier par la poste et la saisie manuelle des données retournées.

Le troisième mode de collecte a été des visites en personne par des intervieweurs sur le terrain. Le plan pour l'ESoR de 2021 était de continuer à avoir recours à des interviews sur le terrain pour une portée limitée, comme pour les cycles précédents (pour l'ESoR de 2016, seul 0,5 % des cas ont été effectués par des intervieweurs sur le terrain), mais plutôt que d'utiliser des questionnaires papier comme par le passé, les intervieweurs sur le terrain utilisaient un ordinateur portable et la même application que les intervieweurs travaillant par téléphone. Toutefois, toutes les interviews en personne ont été annulées à l'étape de la planification de la collecte, du fait de la pandémie de COVID-19.

7.3.2 Opérations

La collecte des données pour l'ESoR a commencé le 28 mars 2022 dans tous les bureaux régionaux. Le dernier jour de collecte active a été le 4 novembre 2022. Le [tableau 7.3.2](#) présente la répartition des cas chargés dans le système ITAO à partir du bureau central au fil du temps. La majorité des cas ont été envoyés au début de la collecte, le 28 mars, et comprenaient les cas relatifs à des adultes de toutes les bases de sondage à l'exception du Nunavut. Le total corrigé représente le nombre de cas envoyés à la collecte, à l'exception des cas supprimés de la collecte.

Rapport technique sur la couverture

Tableau 7.3.2
Nombre total de cas en collecte

Description	Nombre
Cas commencés le 28 mars 2022 : adultes dans toutes les bases de sondage sauf le Nunavut	9 922
Cas commencés le 27 avril 2022 : mineurs dans toutes les bases de sondage (y compris la majorité de la base de sondage des naissances) sauf le Nunavut	1 822
Cas commencés le 6 juin 2022 : base de sondage du Nunavut et cas restants de la base de sondage des naissances	1 352
Nombre total de cas envoyés	13 096
Cas abandonnés par le bureau central : la collecte n'est plus requise (classés lors du traitement comme étant dénombrés ou hors champ)	309
Nombre total ajusté	12 787

Source : Statistique Canada, Étude sur le sous-dénombrement du Recensement de 2021.

Des lettres de présentation expliquant l'ESoR et avisant la PC (ou le répondant substitut) qu'elle avait été sélectionnée pour répondre à l'enquête ont été envoyées pour tous les cas dont la collecte commençait en mars et avril et pour lesquels une adresse postale valide existait. Un numéro de téléphone a été fourni pour toute question, dans les cas où le répondant souhaitait communiquer avec le BR pour répondre à l'enquête par téléphone. Un code d'accès sécurisé ainsi qu'un lien vers le questionnaire d'autodéclaration ont également été fournis aux cas sans numéro de téléphone à joindre (nécessitant un dépistage). Aucune lettre de présentation n'a été envoyée pour les cas dont la collecte a commencé en juin; ces cas ont reçu une lettre de rappel en juillet. Ces lettres de rappel ont été envoyées à tous les cas n'étant pas encore traités vers la moitié de la collecte. Une deuxième lettre de rappel a été envoyée un mois plus tard. Toutes les lettres de rappel comprenaient un code d'accès sécurisé et un lien vers le questionnaire d'autodéclaration. Une nouveauté pour l'ESoR de 2021 est qu'à l'approche de la fin de la collecte, des courriels de rappel ont été envoyés pour tous les cas incomplets pour lesquels on disposait d'une adresse de courriel valide.

Près de la fin de la collecte, dans le cadre d'un effort visant à accroître les taux de réponse, les BR de Toronto et de l'Ouest ont commencé un processus similaire aux visites d'interview sur le terrain effectuées par le passé. Si l'adresse d'une PC se trouvait proche du lieu d'une visite effectuée par un intervieweur pour une autre enquête, cet intervieweur se rendait alors à cette adresse pour essayer d'y rencontrer la PC. S'il trouvait la PC ou confirmait qu'il s'agissait bien du lieu de résidence de la PC, il demandait un numéro de téléphone pour que le BR rappelle afin de terminer l'interview. S'il parvenait à parler avec la PC, il pouvait également lui fournir un code d'accès sécurisé pour remplir le questionnaire en ligne. En cas d'absence de la PC, l'intervieweur essayait de recueillir des coordonnées pouvant permettre le dépistage de la PC.

Une analyse de la qualité des données a été effectuée pour vérifier l'exhaustivité et l'exactitude de chaque cas. Les cas où des champs clés étaient restés vides ou renfermaient des données ambiguës, ou encore lorsque les données recueillies correspondaient à une autre personne que la PC, étaient réactivés et retournés à la collecte aux fins de suivi. Au total, 41 cas ont été réactivés dans le cadre de l'ESoR de 2021. Les cas passant avec succès l'analyse de qualité des données ont été compilés en lots aux fins de traitement, de la façon décrite à la [section 7.2.1](#).

La gestion de la qualité des opérations de collecte comprenait une séance de formation virtuelle bidirectionnelle pour les gestionnaires de la collecte des données régionaux, qui, à leur tour, formaient ensuite les intervieweurs. Des réunions hebdomadaires entre le bureau central et les BR ont eu lieu pendant la collecte pour discuter de la progression et répondre aux problèmes qui survenaient. Un outil de communication fondé sur un système de billets a été utilisé pour centraliser et faciliter les communications entre le bureau central et les BR. Il permettait d'effectuer le suivi de toutes les questions ainsi que des problèmes et de veiller à une résolution rapide. Les gestionnaires des BR ont attribué des ressources à l'enquête tout en respectant les besoins des autres enquêtes réalisées dans leur région. Les efforts soutenus afin d'interviewer les personnes ayant initialement refusé de participer à l'enquête ont permis d'améliorer les taux de réponse.

Le bureau central a créé des rapports de gestion détaillés chaque jour et chaque semaine pour documenter les progrès de la collecte. Ces rapports présentaient le nombre de cas recueillis et les taux de réponse par province sélectionnée et base de sondage.

7.3.3 Dépistage

Dans le cadre de la préparation de l'échantillon, les cas ont été couplés aux données fiscales et à d'autres données administratives pour fournir des coordonnées à jour des PC et des membres de leur ménage. Dans certains cas, les données initiales de l'ITAO étaient désuètes ou incomplètes, si bien qu'un dépistage était nécessaire. Le dépistage est le processus de recherche des coordonnées d'une PC ou d'un répondant substitut; il s'agit d'un élément majeur de l'ESoR.

Ces indices ont été chargés dans l'application d'ITAO comme coordonnées de remplacement avant la collecte; d'autres indices ont été envoyés aux BR lorsqu'on en trouvait lors du traitement au cours de la période de collecte. Un plus grand nombre de fichiers de sources de dépistage ont été envoyés à la collecte pour l'ESoR de 2021 (29 fichiers, par rapport à 13 en 2016); l'amélioration du traitement signifiait que seuls de nouveaux numéros de téléphone et de nouvelles adresses étaient envoyés aux BR, sans duplication de sources antérieures.

Dans le cadre de l'ESoR, des ententes ont été conclues avec 11 provinces et territoires (dont 9 ont eu recours à des personnes réputées être employées) qui ont envoyé des informations de dépistage. Le bureau central a envoyé des fichiers contenant les noms de PC, qui ont été comparés à des dossiers de soins de santé et renvoyés avec les coordonnées mises à jour. Le recours à des personnes réputées être employées signifiait que le nom et la date de naissance de la PC pouvaient être fournis; facilitant la mise en correspondance entre les fichiers.

Au début de la collecte des données, seuls 2,1 % des cas présentaient des coordonnées insuffisantes et devaient faire l'objet d'un dépistage. Du fait de la qualité et de la quantité des sources de dépistage mis à disposition par le bureau central, des numéros de téléphone fournis par le bureau central ont été utilisés pour 90,6 % des cas traités. Un pourcentage supplémentaire de 8,6 % des cas traités a pu être contacté grâce à un nouveau numéro de téléphone fourni grâce aux efforts de dépistage des BR; enfin un pourcentage final de 0,8 % supplémentaire des cas ont été traités lorsque les répondants ont contacté le BR.

7.3.4 Statistiques de collecte

De nombreuses statistiques ont fait l'objet d'une surveillance tout au long de la période de collecte de données et ont été analysées après la fin de la collecte.

Rapport technique sur la couverture

Le tableau 7.3.4.1 présente les taux d'achèvement provinciaux et territoriaux selon le mode de collecte. Sur les 7 702 cas traités, 87,6 % l'ont été par ITAO et 12,4 % par autodéclaration en ligne.

Tableau 7.3.4.1

Nombre de cas terminés et taux d'achèvement selon la méthode de collecte, Canada, provinces et territoires de sélection

Province ou territoire	Nombre de cas envoyés	Intervieweur		Autodéclaration		Total	
		Nombre de cas terminés	Taux d'achèvement (%)	Nombre de cas terminés	Taux d'achèvement (%)	Nombre de cas terminés	Taux d'achèvement (%)
Canada	12 787	6 745	52,7	957	7,5	7 702	60,2
Terre-Neuve-et-Labrador	503	291	57,9	33	6,6	324	64,4
Île-du-Prince-Édouard	487	278	57,1	52	10,7	330	67,8
Nouvelle-Écosse	620	374	60,3	37	6,0	411	66,3
Nouveau-Brunswick	522	286	54,8	34	6,5	320	61,3
Québec	1 315	769	58,5	93	7,1	862	65,6
Ontario	2 406	1 214	50,5	235	9,8	1 449	60,2
Manitoba	852	451	52,9	47	5,5	498	58,5
Saskatchewan	832	425	51,1	51	6,1	476	57,2
Alberta	1 375	703	51,1	100	7,3	803	58,4
Colombie-Britannique	1 746	828	47,4	152	8,7	980	56,1
Yukon	460	239	52,0	29	6,3	268	58,3
Territoires du Nord-Ouest	632	345	54,6	33	5,2	378	59,8
Nunavut	950	529	55,7	56	5,9	585	61,6
RNP-CA	87	13	14,9	5	5,7	18	20,7

RNP-CA = résidents non permanents dont la province est inconnue

Source : Statistique Canada, Étude sur le sous-dénombrement du Recensement de 2021.

Le tableau 7.3.4.2 présente les taux d'achèvement par base de sondage et mode de collecte. Conformément aux tendances historiques, la base des résidents non permanents a enregistré le taux d'achèvement le plus bas (49,4 %), puisque les PC de cette base de sondage tendent à être plus mobiles et à fournir moins des coordonnées, rendant le dépistage difficile.

Tableau 7.3.4.2

Nombre de cas terminés et taux d'achèvement selon la base de sondage et la méthode de collecte pour le Canada

Base de sondage	Nombre de cas envoyés	Intervieweur		Autodéclaration		Total	
		Nombre de cas terminés	Taux d'achèvement (%)	Nombre de cas terminés	Taux d'achèvement (%)	Nombre de cas terminés	Taux d'achèvement (%)
Total	12 787	6 745	52,7	957	7,5	7 702	60,2
Recensement de 2016	6 773	3 720	54,9	482	7,1	4 202	62,0

Rapport technique sur la couverture

Tableau 7.3.4.2

Nombre de cas terminés et taux d'achèvement selon la base de sondage et la méthode de collecte pour le Canada

Base de sondage	Nombre de cas envoyés	Intervieweur		Autodéclaration		Total	
		Nombre de cas terminés	Taux d'achèvement (%)	Nombre de cas terminés	Taux d'achèvement (%)	Nombre de cas terminés	Taux d'achèvement (%)
Personnes omises de 2016	1 310	691	52,7	84	6,4	775	59,2
Naissances	671	377	56,2	40	6,0	417	62,1
Immigrants	553	280	50,6	87	15,7	367	66,4
Résidents non permanents	1 438	564	39,2	146	10,2	710	49,4
Yukon	460	239	52,0	29	6,3	268	58,3
Territoire du Nord-Ouest	632	345	54,6	33	5,2	378	59,8
Nunavut	950	529	55,7	56	5,9	585	61,6

Source : Statistique Canada, Étude sur le sous-dénombrement du Recensement de 2021.

Le tableau 7.3.4.3 présente les taux d'achèvement par sexe et groupe d'âge. Les taux d'achèvement les plus bas ont été pour les personnes des deux sexes âgées de 20 à 44 ans; le taux le plus élevé correspondant à celui des femmes âgées de 45 ans et plus.

Tableau 7.3.4.3

Nombre de cas terminés et taux d'achèvement selon la méthode de collecte, le sexe et le groupe d'âge pour le Canada

Sexe et groupe d'âge	Nombre de cas envoyés	Intervieweur		Autodéclaration		Total	
		Nombre de cas terminés	Taux d'achèvement (%)	Nombre de cas terminés	Taux d'achèvement (%)	Nombre de cas terminés	Taux d'achèvement (%)
Les deux sexes	12 783	6 745	52,8	957	7,5	7 702	60,3
0 à 19 ans	1 930	1 062	55,0	140	7,3	1 202	62,3
20 à 29 ans	2 420	1 198	49,5	169	7,0	1 367	56,5
30 à 44 ans	4 697	2 303	49,0	389	8,3	2 692	57,3
45 ans et plus	3 736	2 182	58,4	259	6,9	2 441	65,3
Hommes	6 952	3 609	51,9	496	7,1	4 105	59,0
0 à 19 ans	963	530	55,0	74	7,7	604	62,7
20 à 29 ans	1 273	623	48,9	99	7,8	722	56,7
30 à 44 ans	2 678	1 305	48,7	199	7,4	1 504	56,2
45 ans et plus	2 038	1 151	56,5	124	6,1	1 275	62,6
Femmes	5 831	3 136	53,8	461	7,9	3 597	61,7
0 à 19 ans	967	532	55,0	66	6,8	598	61,8
20 à 29 ans	1 147	575	50,1	70	6,1	645	56,2
30 à 44 ans	2 019	998	49,4	190	9,4	1 188	58,8
45 ans et plus	1 698	1 031	60,7	135	8,0	1 166	68,7

Note : Ce tableau exclut quatre cas pour lesquels le sexe était inconnu.

Source : Statistique Canada, Étude sur le sous-dénombrement du Recensement de 2021.

7.4 Estimation

L'estimation de l'ESoR s'est divisée en deux parties. D'abord, on a effectué la pondération des PC, puis on a calculé le sous-dénombrement du recensement. La pondération est le processus comprenant la détermination des poids d'échantillonnage initiaux des PC ainsi que tous les ajustements subséquents effectués à ces poids initiaux menant à la création des poids finaux des PC. La pondération comprend plusieurs étapes, qui sont décrites aux sections [7.4.1](#) à [7.4.4](#). La méthodologie du calcul du sous-dénombrement du recensement est décrite à la [section 7.4.6](#).

7.4.1 Calcul du poids initial

Pour les PC de toutes les bases de sondage sauf la base des personnes omises de 2016, les poids initiaux ont été attribués en fonction de l'inverse des probabilités de sélection dans l'échantillon. Cependant, le poids initial d'une PC de la base des personnes omises de 2016 correspond au poids final qui lui a été attribué lors de l'ESoR de 2016 au moment où la PC a été classée comme « omise ».

7.4.2 Ajustements aux poids initiaux

Les poids des PC de la base du Recensement de 2016 qui ont été dénombrées plus d'une fois en 2016 ont pour leur part été rajustés à la baisse pour tenir compte du fait que ces personnes avaient plus d'une chance d'être sélectionnées.

Ensuite, un ajustement aux poids initiaux influents dans la base des personnes omises de 2016 a été effectué. L'objectif était de réduire l'effet des poids élevés et influents sur les estimations et les erreurs-types à l'aide d'un élagage de leurs poids initiaux. Quelques personnes parmi les 4 821 personnes de la base des personnes omises de 2016 avaient un poids initial très élevé. La méthode utilisée consistait à tronquer des poids à un multiplicateur de la médiane des poids dans chaque groupe d'élagage formé. Les groupes d'élagage ont été formés selon la province de sélection et cinq groupes d'âge. Une personne ayant un poids supérieur au seuil voyait son poids être réduit à cette valeur. Les poids tronqués ont été redistribués uniformément aux autres personnes dans le groupe d'élagage.

7.4.3 Rajustement de la non-réponse

Pour réduire les biais statistiques, les poids initiaux des répondants devaient être rajustés pour tenir compte de la non-réponse. Le poids des personnes pour lesquelles une classification n'a pu être établie (les non-répondants) a été redistribué parmi les personnes pour lesquelles une classification a été établie (les répondants). Il existe trois types de non-réponse. Il y a d'abord les personnes non identifiées (seulement 12 PC). Les poids initiaux de ces personnes ont été transférés aux personnes identifiées dans chaque strate d'échantillonnage.

Le deuxième type de non-réponse concerne les personnes non dépistées (4 990 PC). L'ajustement consistait à former des groupes homogènes de réponse (GHR) parmi les personnes non listées (les personnes listées étant les personnes classifiées sans l'aide de la collecte de l'ESoR) et à transférer les poids des personnes non dépistées aux personnes dépistées non listées à l'intérieur des GHR.

La première étape de la création des GHR consistait à rassembler les personnes non listées (12 337 PC) en groupes principaux d'après leur propension estimée à être dans la population cible. Les groupes ont été formés en fonction de l'analyse de la corrélation entre plusieurs indicateurs fiscaux, notamment ceux pour 2020 et 2021, et la classification finale des personnes dépistées non listées. Jusqu'à sept groupes principaux ont été créés selon la base de sondage. Ces groupes principaux étaient aussi fortement corrélés à la propension à répondre. La deuxième étape de la création des GHR consistait à grouper les personnes non listées selon leur propension à répondre dans chacun des domaines, où un domaine est défini par le croisement d'une base de sondage avec un groupe principal. Dans chaque domaine, l'analyse de la propension à répondre a été faite au moyen d'un modèle de régression logistique à l'échelle nationale (ainsi qu'à l'échelle régionale, si les données le permettaient) et de l'analyse de tableaux de fréquences croisées à plusieurs niveaux. Dans le cas des modèles, plusieurs variables auxiliaires disponibles tant pour les personnes dépistées que non dépistées ont été utilisées, à savoir les variables disponibles dans les bases de sondage (p. ex. âge, sexe, relation avec les autres membres du ménage, pays

d'origine, identité autochtone et type de résident non permanent), les variables disponibles dans les données fiscales pour les personnes liées (p. ex. présence ou non dans certains fichiers, fréquence des changements d'adresses depuis 2016 et type d'adresse), les variables reliées aux renseignements relatifs aux coordonnées (p. ex. nombre et sources de numéros de téléphone, disponibilité d'une adresse et lien de la dernière adresse connue avec le Recensement de 2021) et quelques autres variables. Ainsi, les variables auxiliaires significativement corrélées avec la propension à répondre ont été déterminées et utilisées pour former les GHR. Dans la plupart des domaines, les GHR ont été formés à l'intérieur de la province ou du territoire de sélection. L'ajustement consistait donc à transférer le poids des personnes non dépistées aux personnes dépistées non listées à l'intérieur de chaque GHR.

Le troisième ajustement pour la non-réponse est celui des personnes non classifiées (169 PC). Une personne non classifiée est une personne qui avait sa résidence principale dans une province ou un territoire donné le jour du recensement (qui faisait donc partie de la population cible du recensement), mais dont on ne sait pas avec certitude si elle avait été omise ou dénombrée. Suivant le même principe que pour les personnes non dépistées, des groupes homogènes de personnes classifiées ont été formés à l'intérieur de chaque base de sondage et province de classification. L'ajustement consistait à transférer le poids des personnes non classifiées aux personnes classifiées non listées à l'intérieur de chaque groupe homogène.

7.4.4 Ajustements finaux aux poids des personnes classées

7.4.4.1 Ajustement pour les poids influents

À cette étape, certaines PC ont un poids qui est élevé et considéré comme influent dans leur province de classification. Afin de réduire l'effet des poids élevés et influents sur les estimations provinciales et leurs erreurs-types, un ajustement aux poids influents a été effectué dans les cinq bases des provinces. La méthode utilisée est l'élagage des poids à un multiplicateur de la médiane des poids dans chaque groupe d'élagage formé. Il y a deux types de poids influents à cette étape.

D'abord, il y a les PC dont la province de classification est différente de la province de sélection. Donc, le poids est très élevé par rapport aux autres PC dans cette province de classification. Prenons, par exemple, une PC sélectionnée en Ontario ayant un grand poids, qui est classée à l'Île-du-Prince-Édouard. Dans cette situation, le poids est tronqué selon le seuil établi par groupe d'élagage. Un facteur entre quatre et six fois la médiane pour chacun des groupes a été utilisé comme seuil d'élagage. Les groupes d'élagage ont été formés selon la province de classification et cinq groupes d'âge. Les poids tronqués d'une PC ont été redistribués uniformément aux autres PC dans la même province de sélection, la même base de sondage, la même classification (personne dénombrée, omise ou hors cible), le même état (listée ou non listée) et par groupe d'âge. Ainsi, le poids influent d'une PC omise dans une province de classification donnée était réparti à d'autres personnes omises, mais dans la province de sélection de la PC. Pour ce premier type de poids influents, il y a eu 49 PC dont le poids a été tronqué, soit 33 personnes dénombrées et 16 personnes omises.

Le deuxième type de poids influents concerne les PC de la base des personnes omises de 2016 seulement, qui avaient encore un poids élevé et influent à l'intérieur de leur province de classification même si celle-ci était identique à la province de sélection (qui est, en fait, la province de classification en 2016). Pour ce type de poids influents, le seuil a été fixé à quatre fois la médiane des poids dans le groupe d'élagage. Les poids tronqués des PC ont été redistribués uniformément aux autres PC dans la même province de classification et la même classification, ce qui n'a donc pas d'effet sur l'estimation du sous-dénombrement provincial. Pour ce deuxième type de poids influents, il y a eu 95 PC dont le poids a été tronqué, soit 10 personnes dénombrées, 55 personnes omises et 30 personnes hors cible.

7.4.4.2 Calage des poids pour la base des naissances

Pour l'échantillon de la base des naissances, un calage aux personnes dénombrées a été effectué afin de tenir compte des cas où un échantillon provincial contiendrait trop ou trop peu de personnes dénombrées. Un couplage automatisé déterministe appliqué à la BDR-ECR de 2021 a permis de déterminer les totaux de contrôle par province pour le groupe de calage des personnes dénombrées. Puis, pour les autres personnes figurant dans la base, un couplage aux données fiscales a permis de déterminer leur province de résidence au jour du recensement (sinon, la province de sélection a été utilisée) afin de déterminer les totaux de contrôle par province pour le groupe de calage des personnes non dénombrées. En outre, des totaux de contrôle par année d'âge (0 à 4 ans) ont été calculés. Le calage a été effectué à l'aide d'une méthode de ratissage aux marges en utilisant les 20 totaux de contrôle décrits ci-dessus pour la première marge, de même que 5 groupes de calage selon l'âge comme deuxième marge. Pour ce faire, le Système généralisé d'estimation (G-Est) de Statistique Canada a été utilisé.

7.4.4.3 Calage des poids pour la base des immigrants

Pour l'échantillon de la base des immigrants, un calage au nombre de personnes dans certains groupes de calage a été effectué afin de tenir compte des cas où un échantillon provincial contiendrait trop ou trop peu de personnes dénombrées ou de personnes faisant partie des autres groupes. Un couplage automatisé déterministe appliqué à la BDR-ECR de 2021 a permis de déterminer les totaux de contrôle par province pour le groupe de calage des personnes dénombrées. Puis, pour les autres personnes figurant dans la base, un couplage aux données fiscales a permis de déterminer leur situation fiscale (active ou non) et leur province de résidence au jour du recensement (sinon, la province de sélection a été utilisée) afin de déterminer les totaux de contrôle par province pour les autres groupes de calage des personnes non dénombrées. Dans les quatre plus grandes provinces, trois totaux de contrôle ont été déterminés : pour les personnes dénombrées, pour les personnes ayant des activités fiscales récentes, et pour les autres personnes. Cependant, dans les six autres provinces, seulement deux totaux de contrôle ont été déterminés : pour les personnes dénombrées, et les autres personnes. Ainsi, 24 totaux de contrôle ont été formés. Une simple méthode de poststratification a alors été effectuée pour faire le calage de la base des immigrants.

7.4.4.4 Rajustement par poststratification pour les territoires

Après avoir rajusté les poids initiaux, on a observé que l'estimation de personnes dénombrées dans les territoires était habituellement inférieure au chiffre comparable du recensement. Cette situation est attribuable à un sous-dénombrement de la population cible du recensement dans les fichiers d'assurance maladie. Pour pallier ce sous-dénombrement, les poids des PC dans les territoires ont été rajustés afin que l'estimation du nombre de personnes dénombrées égale le chiffre comparable du recensement pour ce territoire. Les ajustements ont été effectués pour six groupes de calage (selon l'âge et le genre) dans chaque territoire.

7.4.4.5 Ajustement pour le chevauchement de bases ou de strates

Pour un petit nombre de PC dans les cinq bases des provinces, le poids n'est pas le poids final, car un autre ajustement doit être effectué pour tenir compte du chevauchement entre les bases de sondage ou, dans certains cas, du chevauchement entre les strates de la base du recensement (donc surdénombrement en 2016), mais qui a été relevé seulement à la suite de la collecte de l'ESoR de 2021. Pour ce qui est des quelques PC en chevauchement de bases, ce sont en majorité des PC de la base des immigrants ou de la base des résidents non permanents qui étaient finalement prises en compte dans la base du Recensement de 2016 (donc dénombrées en 2016). Cette information n'était pas connue au moment de la préparation de ces bases de sondage. Ainsi, un facteur d'ajustement a été calculé en tenant compte de la probabilité de sélection dans les deux bases de sondage.

7.4.5 Distribution pondérée selon la classification

Le tableau 7.4.5 présente la distribution pondérée des PC selon la classification et la base de sondage. Pour un rappel des définitions, veuillez consulter la [section 7.2](#). Seules les PC trouvées dans la BDR-ECR ont été classées comme « dénombrées ». Les personnes de la population cible qui ne sont pas dans la BDR-ECR ont été classées comme « omises ». Les PC restantes ont été classées comme « hors champ » (p. ex. décédées ou émigrées).

Tableau 7.4.5
Classification pondérée des personnes choisies, bases de sondage pour le Canada

Classification	Strate provinciale									
	Recensement de 2016		Personnes omises de 2016		Naissances		Immigrants		Résidents non permanents	
	nombre	%	nombre	%	nombre	%	nombre	%	nombre	%
Total	32 933 387	100,0	2 830 944	100,0	1 855 111	100,0	1 072 833	100,0	1 140 539	100,0
Dénombrées	29 127 257	88,4	1 784 797	63,0	1 646 438	88,8	874 651	81,5	639 236	56,0
Listées	29 023 031	88,1	1 773 009	62,6	1 643 876	88,6	871 537	81,2	626 405	54,9
Non listées	104 226	0,3	11 788	0,4	2 562	0,1	3 114	0,3	12 831	1,1
Omises	2 083 885	6,3	662 494	23,4	164 767	8,9	130 942	12,2	387 586	34,0
Listées	243 914	0,7	41 300	1,5	16 954	0,9	7 987	0,7	14 693	1,3
Non listées	1 839 971	5,6	621 194	21,9	147 813	8,0	122 955	11,5	372 893	32,7
Hors champ	1 722 245	5,2	383 653	13,6	43 906	2,4	67 240	6,3	113 717	10,0
Listées	1 402 710	4,3	206 632	7,3	25 675	1,4	3 964	0,4	32 613	2,9
Non listées	319 535	1,0	177 021	6,3	18 231	1,0	63 276	5,9	81 104	7,1

Classification	Strate territoriale			
	Bases de sondage territoriales		Total	
	nombre	%	nombre	%
Total	137 867	100,0	39 970 681	100,0
Dénombrées	94 583	68,6	34 166 962	85,5
Listées	94 272	68,4	34 032 130	85,1
Non listées	311	0,2	134 832	0,3
Omises	32 760	23,8	3 462 434	8,7
Listées	5 567	4,0	330 415	0,8
Non listées	27 193	19,7	3 132 019	7,8
Hors champ	10 524	7,6	2 341 285	5,9
Listées	7 768	5,6	1 679 362	4,2
Non listées	2 756	2,0	661 923	1,7

Source : Statistique Canada, Étude sur le sous-dénombrement du Recensement de 2021.

7.4.6 Calcul du sous-dénombrement du recensement

On définit maintenant que :

C	=	chiffre publié du recensement du nombre de personnes dans la population cible
\hat{S}	=	estimation du sous-dénombrement
	=	estimation du nombre de personnes non incluses en C , mais qui auraient dû l'être
\hat{M}	=	estimation du nombre de personnes de la population visée par l'ESoR qui n'ont pas été dénombrées
	=	somme du poids final des personnes considérées comme omises
X	=	nombre de personnes incluses en C qui ne peuvent pas, dans le cadre de l'ESoR, être déterminées avec certitude comme des personnes dénombrées.

On estime le sous-dénombrement de la population du recensement par le nombre (pondéré) de personnes omises moins le nombre de personnes comptées par le recensement (le terme C) mais exclues de la BDR-ECR :

$$\hat{S} = \hat{M} - X$$

X a trois composantes : les imputations, les dénombrements incomplets et les dénombrements tardifs.

L'adresse de la PC le jour du recensement désigne un logement pour lequel un dénombrement a été imputé. C'est le cas, notamment, des logements non répondants pour lesquels les données d'un autre logement ont été utilisées dans le cadre de l'IME.

Certains dénombrements figurant dans la base de données du recensement ont été jugés trop incomplets pour être utilisés aux fins de l'ESoR afin de déterminer si une PC était une personne dénombrée. Le dénombrement est généralement incomplet en pareil cas en raison de données manquantes ou non valides en ce qui a trait à la date de naissance ou au nom (p. ex. « ? », « M. », « Inconnu » ou « Personne 1 »). Lorsque la PC a fait l'objet d'un tel dénombrement, elle est considérée comme une personne « omise ». C'est ce qu'on appelle un « dénombrement incomplet de l'ESoR ». Cette catégorie de dénombrement inclut aussi certains types de logements collectifs pour lesquels seul le nombre de résidents habituels a été recueilli au recensement (pas de noms ni de dates de naissance). Il y a eu imputation dans la BDR des données des personnes vivant dans ces logements collectifs.

À l'échelle nationale, X représente un peu moins de la moitié de \hat{M} . La valeur de X a augmenté par rapport à 2016 en raison d'une augmentation du nombre de personnes imputées dans le cadre de l'IME et de l'augmentation des imputations dans certains types de logements collectifs (dénombrements incomplets).

Rapport technique sur la couverture

Le tableau 7.4.6 présente les chiffres nationaux des différentes composantes de l'estimation du sous-dénombrement de la population, notamment les chiffres pour les trois composantes du terme X .

Tableau 7.4.6
Composantes de l'estimation du sous-dénombrement de la population pour le Canada

Composante	Nombre de personnes
Estimation de M	3 462 434
X total	1 564 558
X de personnes imputées	931 346
X de dénombremments tardifs	0
X de dénombremments incomplets selon l'ESoR	633 212
Estimation de S	1 897 876

ESoR = Étude sur le sous-dénombrement du recensement

M = nombre de personnes de la population visée par l'Étude sur le sous-dénombrement du recensement (ESoR) qui n'ont pas été dénombrées
X = nombre de personnes comprises dans le chiffre publié du recensement, mais qui n'ont pas pu être déterminées avec certitude comme étant dénombrées dans le cadre de l'ESoR

S = sous-dénombrement

Source : Statistique Canada, Étude sur le sous-dénombrement du Recensement de 2021.

Enfin, le calcul de la variance des estimations du sous-dénombrement est le suivant :

$$v(\hat{S}) = v(\hat{M} - X) = v(\hat{M})$$

$v(\hat{M})$ = estimation de la variance de \hat{M} selon le plan de l'ESoR.

Le calcul de la variance a été fait à l'aide de la méthode classique du rééchantillonnage bootstrap. Pour ce faire, des poids de 500 répliques bootstrap ont été produits.

8. Étude sur le surdénombrement du recensement

8.1 Aperçu

Une erreur de surdénombrement se produit lorsque des personnes dans le champ de l'enquête sont dénombrées plus d'une fois ou que des personnes qui n'auraient pas dû être dénombrées sont incluses dans la population cible d'une enquête ou d'un recensement. L'Étude sur le surdénombrement du recensement (ESuR) vise à estimer le nombre de personnes dénombrées plusieurs fois dans le cadre du Recensement de la population du Canada¹¹.

L'ESuR de 2021 a compris deux types de couplages : déterministe et probabiliste. Le couplage déterministe (CD) a relevé des paires définitives de doublons de personnes; ce qui signifie que ces personnes ont été dénombrées plusieurs fois et représentent donc un surdénombrement. Cette méthodologie s'est fondée sur une modification de l'Étude par appariement automatisé (EAA), utilisée dans le cadre de précédents cycles de recensement pour évaluer l'ESuR. Le couplage probabiliste (CP) a relevé des paires possibles de doublons de personnes en se fondant sur des méthodes utilisées dans des cycles antérieurs de l'ESuR. Les données utilisées par l'ESuR proviennent de la base de données des réponses du Recensement de 2021 et de données administratives provenant de la Base statistique de données démographiques, fournies par la Section de recherche en recensement de la Division des méthodes d'intégration statistique. La base de sondage de l'ESuR a été créée en plusieurs étapes et comprend des paires définitives et possibles de doublons de personnes relevés lors du CD et du CP, ainsi qu'une extension de la base de sondage fondée sur les ménages. Un échantillon de paires possibles de doublons de personnes a été tiré de la base de sondage de l'ESuR et envoyé aux fins de vérification manuelle pour déterminer si les paires échantillonnées étaient effectivement des doublons de personnes. Au moyen du résultat de la vérification manuelle des paires échantillonnées et des paires définitives de doublons de personnes détectés par le CD, une estimation du surdénombrement a été obtenue.

8.2 Étapes du couplage

8.2.1 Données utilisées pour les couplages

Deux sources de données ont été utilisées pour ces couplages.

Tout d'abord, la version des études sur la couverture du recensement de la Base de données des réponses du recensement (BDR-ECR; appelée la BDR dans le présent chapitre) était une version de la Base de données des réponses du recensement qui n'incluait pas les dénombrements tardifs ou incomplets ni les personnes ajoutées dans le cadre du processus d'imputation des ménages complets. La BDR contenait un peu plus de 35 millions d'enregistrements et comprenait les réponses des personnes vivant dans des logements privés et collectifs. Elle contenait des noms (notamment des prénoms et des noms de famille), des données démographiques (notamment la date de naissance et le sexe) et des données géographiques (notamment la province ou le territoire et le code postal, ainsi que des variables géographiques propres au recensement, comme l'unité de collecte [UC], la subdivision de recensement [SDR] et la région métropolitaine de recensement [RMR]).

Deuxièmement, des données administratives (ADM) ont été utilisées, provenant de la Base statistique de données démographiques canadiennes, fournies par la Section de recherche en recensement de la Division des méthodes d'intégration statistique. Elles comprenaient des enregistrements de plusieurs sources de données ADM et visaient à représenter des personnes dans le champ d'enquête du recensement. Ces données ADM consistaient en 53 millions d'enregistrements environ. Elles incluaient des noms (prénoms et noms de famille), des données démographiques (y compris la date de naissance et le sexe) et des données géographiques (y compris la province ou le territoire et le code postal).

11. Avant 2016, les études menées pour déterminer l'incidence de l'inclusion de personnes hors du champ de l'enquête ont démontré que cette erreur était négligeable; elle n'est donc pas prise en compte par l'ESuR. Toutefois, une étude pilote est à l'examen après le cycle de production de 2021 pour réévaluer cette hypothèse.

Les variables d'appariement suivantes ont été utilisées dans les couplages (le cas échéant) :

- noms : variables de prénom(s) et de nom(s) de famille;
- données démographiques : variables de date de naissance et de sexe;
- données géographiques : variables de province ou territoire et de code postal, ainsi que variables géographiques de recensement.

8.2.2 Couplage déterministe

Le CD vise à déterminer les paires de grande qualité de doublons de personnes, consistant en deux enregistrements de la BDR, classés comme des paires définitives de surdénombrement. Les programmes d'appariement déterministe habituellement utilisés pour l'EAA ont été modifiés pour inclure, dans le cadre des critères de couplage, une comparaison des noms en considérant les appariements entre un ménage vivant dans un logement privé et un ménage vivant dans un logement collectif.

Ce CD se basait sur la série d'opérations suivantes :

- Des programmes d'appariement déterministe ont été utilisés pour relever les paires de ménages « similaires ». La similarité était décrite du point de vue de leur proximité géographique relative (ménages situés dans la même UC, ménages vivant dans des UC différentes, mais dans la même SDR, etc.) et du nombre de personnes appariées entre ces ménages. Les personnes étaient appariées selon les variables de nom, de sexe et de date de naissance. Deux personnes étaient jugées présenter une correspondance exacte si elles avaient le même sexe, les mêmes jour, mois et années de naissance ainsi que le même nom. Deux personnes étaient jugées être une proche correspondance si leurs noms correspondaient et que trois des quatre autres composantes (sexe et jour, mois et année de naissance) correspondaient ou si seulement le jour et le mois de naissance étaient inversés. Les paires de ménages consistaient en l'un ou les deux ménages vivant dans un logement privé.
- Une liste initiale de paires possibles de doublons de personnes a été créée à partir des paires de ménages.
- Un échantillon de vérification a été tiré de la liste initiale de paires possibles de doublons de personnes aux fins de vérification manuelle, afin de confirmer leur grande qualité avant de les classer comme paires définitives de doublons de personnes (c.-à-d. surdénombrement).
- Une liste finale de paires de doublons de personnes a été déterminée, puis elles ont été classées comme paires définitives de doublons de personnes provenant du CD.

Ce couplage a donné 460 572 paires définitives de doublons de personnes provenant du CD.

8.2.3 Couplage probabiliste

Le CP visait à relever les paires possibles de doublons de personnes. Ce couplage consistait à un couplage d'enregistrements probabiliste interne de la BDR entière avec elle-même, appelé couplage BDR–BDR, et d'un couplage d'enregistrements probabiliste externe de la BDR avec des données ADM, appelé couplage BDR–ADM. Le couplage BDR–BDR a donné des paires d'enregistrements de BDR, alors que le couplage BDR–ADM a donné des paires au sein desquelles un enregistrement provenait de la BDR et l'autre de données ADM; des paires d'enregistrements de la BDR ont été ensuite dérivées.

Le couplage probabiliste s'est effectué avec G-Coup, système de couplage d'enregistrements probabiliste conçu à Statistique Canada utilisant la méthodologie Fellegi-Sunter pour résoudre des problèmes de couplage de fichiers volumineux lorsqu'aucun identificateur direct n'est commun aux deux sources ([Fellegi et Sunter, 1969](#)). Comme dans les cycles antérieurs, G-Coup a été utilisé en 2021 et les séries suivantes d'opérations ont été effectuées séparément pour les couplages BDR–BDR et BDR–ADM.

La première tâche d'un couplage probabiliste consiste à construire un ensemble de paires potentielles (aussi appelé **ensemble couplé**), qui sert à estimer les caractéristiques de l'ensemble de paires de **réelles correspondances**. Pour ce faire, un ensemble de critères de sélection a été appliqué, réduisant le produit cartésien de toutes les correspondances possibles à un espace de comparaison plus gérable. Des améliorations ont été apportées aux critères de sélection de 2021, afin de relever des défis relatifs aux critères de sélection de 2016. De plus, plutôt que d'utiliser les mêmes critères de sélection pour les deux couplages interne et externe, des critères ont été élaborés, testés et optimisés séparément pour ces deux opérations. De nombreuses paires de BDR dérivées de l'ensemble couplé BDR-ADM avaient des paires de BDR-ADM correspondantes qui ont été recueillies à partir de différents critères; ce qui suggère qu'une comparaison directe ne serait pas en mesure de recueillir ces données. Les critères de sélection du couplage BDR-BDR ont fourni un ensemble couplé de 86 429 651 paires BDR-BDR. Les critères de sélection du couplage BDR-ADM externe ont fourni 70 274 756 paires. Seules 41 474 581 de ces paires faisaient intervenir plusieurs enregistrements de la BDR couplés au même enregistrement ADM. Par conséquent, l'ensemble couplé BDR-ADM comprenait ces 41 474 581 paires.

Une fois un ensemble couplé de paires obtenu, les enregistrements des paires ont été comparés en appliquant des règles de couplage dans G-Coup, qui a calculé les poids des résultats des règles de couplage. Des règles de couplage de qualité, portant sur tous les ensembles de caractéristiques pour lesquelles deux enregistrements concordent, ont été nécessaires pour assurer l'exhaustivité de la base de sondage de l'ESuR résultant du CP. Si certains ensembles de caractéristiques ne sont pas visés par les règles de couplage, les paires présentant de telles caractéristiques sont susceptibles de se voir attribuer un poids de couplage inférieur et d'être rejetées lorsque des seuils sont appliqués. De nombreuses améliorations ont été apportées aux règles de couplage de 2021 pour veiller à ce que les poids de couplage estimés soient bien corrélés avec la probabilité qu'une paire soit une véritable correspondance. En 2021, davantage de variables de couplage ont été ajoutées aux règles et les résultats des règles existantes utilisées en 2016 ont été modifiés, p. ex. les règles sur les noms. Les résultats basés sur les variables géographiques propres au recensement (comme l'identifiant unique d'un logement [appelé FRAME_ID], l'UC et la SDR) ont été ajoutés en 2021 et pouvaient uniquement être appliqués au couplage BDR-BDR.

Un seuil de poids de couplage pour chaque province et territoire a ensuite été établi pour les couplages BDR-BDR et BDR-ADM. L'objectif du choix de seuils de poids de couplage était de répartir de façon optimale les paires de l'ensemble couplé en deux catégories : paires non appariées et paires appariées potentielles. Comme dans les cycles antérieurs, les seuils provinciaux et territoriaux ont été choisis hors de G-Coup, car les outils intégrés de détermination de seuils ne fonctionnaient pas bien avec les nombreuses règles de couplage définies par l'utilisateur utilisées par le CP de l'ESuR. Ces seuils ont été choisis en deux étapes. Tout d'abord, un ensemble de seuils préliminaires a été choisi. En général, le choix d'un seuil inférieur est quelque peu subjectif. L'ESuR a utilisé les lignes directrices d'examen de profil élaborées par les experts de l'Environnement de couplage de données sociales (ECDS) de Statistique Canada, afin d'aider à la sélection d'un seuil préliminaire plus bas. Pour éviter de manquer un surdénombrement potentiel, un seuil relativement bas a d'abord été choisi. Un échantillon de paires au-dessus et en dessous de ce seuil préliminaire a ensuite été sélectionné, et le seuil a été ajusté en fonction des besoins. Le seuil final a été choisi afin de réduire au minimum le taux de correspondance manqué visé à 0,01. Puisqu'un ensemble de paires définitives de doublons de personnes a été obtenu à l'étape du CD, aucun seuil supérieur n'a été sélectionné. Toutes les paires des ensembles de couplage BDR-BDR et BDR-ADM dont le poids était supérieur au seuil ont été sélectionnées et jugées être des paires potentielles.

8.3 Création de la base de sondage de l'Étude sur le surdénombrement du recensement

La base de sondage de l'ESuR a été créée en plusieurs étapes et comprenait les paires couplées déterminées lors du CD et du CP, ainsi qu'une extension de la base de sondage fondée sur les ménages. Des unités d'échantillonnage ont ensuite été créées.

Comme cela a été précédemment décrit, le CD a été utilisé pour déterminer un ensemble de paires BDR-BDR qui ont été classées comme paires définitives de doublons de personnes. Le couplage probabiliste a servi à déterminer un ensemble de paires potentielles. Le couplage interne de la BDR à elle-même a permis de déterminer des paires BDR-BDR potentielles qui ont été classées comme paires possibles de doublons de personnes. Les paires BDR-BDR devaient être dérivées de l'ensemble couplé BDR-ADM pour les inclure à

la base de sondage de l'ESuR. Les paires potentielles relevées dans le cadre du couplage BDR-ADM ont été converties en paires BDR-BDR. Lorsque deux enregistrements de la BDR étaient liés au même enregistrement ADM, ces deux enregistrements de la BDR devenaient une paire BDR-BDR. Les paires BDR-ADM individuelles n'avaient pas d'intérêt, puisque l'objectif était de mesurer le surdénombrement (c.-à-d. les doublons de personnes) relatif à la BDR. L'ensemble final de paires BDR-BDR dérivé du couplage BDR-ADM contenait 4 301 512 paires, qui ont été classées comme paires possibles de doublons de personnes.

Comme dans les cycles précédents, la base de sondage a ensuite été enrichie au moyen de paires supplémentaires n'ayant pas encore été relevées dans le cadre du couplage probabiliste, mais ayant été créées à partir de paires de ménages couplées lors des étapes de couplage interne et externe. Cette étape visait à déterminer des paires supplémentaires possibles de doublons de personnes au sein des ménages des paires capturées pouvant ne pas avoir été relevées dans le cadre du CP, car ce dernier était fondé sur des comparaisons de personnes plutôt que de ménages. Les paires potentielles obtenues à cette étape ont été appelées des paires d'extension et classées comme paires possibles de doublons de personnes. Pour créer l'ensemble de paires d'extension, une paire de ménages a d'abord été produite pour chaque paire BDR-BDR classée comme paire possible de doublons de personnes provenant du CP en y ajoutant les autres membres du ménage. Ensuite, à l'aide des variables de sexe et de date de naissance, de nouvelles paires BDR-BDR ont été relevées en comparant les personnes présentes dans la paire de ménages. Des règles de comparaison ont été appliquées pour relever les paires pouvant représenter des cas de surdénombrement. Les paires d'extension comprenaient des paires provenant de deux ménages privés ou des paires au sein desquelles une personne d'un ménage privé était couplée à une personne d'un logement collectif. On a exclu les paires pour lesquelles les deux enregistrements provenaient de logements collectifs.

L'ensemble couplé final comprenait des paires du CD, des paires d'extension, des paires du CP de BDR-ADM et des paires du CP de BDR-BDR.

Tableau 8.3.1
Répartition des paires potentielles de personnes dénombrées en double selon le type de couplage

Type de couplage	Fréquence	Pourcentage
CD	460 572	3,62
Extension	471 688	3,71
BDR-ADM	4 301 351	33,80
BDR-BDR	7 491 998	58,87

CD = couplage déterministe

BDR-ADM = couplage probabiliste de la base de données des réponses du recensement aux données administratives

BDR-BDR = couplage probabiliste de la base de données des réponses du recensement avec elle-même

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

Lorsque des paires de l'ensemble du CP étaient également relevées par le CD, ce type de couplage a été défini comme CD. Ensuite, les paires possibles de doublons de personnes obtenues du CD, du CP et de l'extension ont été combinées et les doublons ont été éliminés.

Depuis 2011, l'ESuR a utilisé des groupes d'enregistrements interconnectés plutôt que des paires d'enregistrements pour estimer le surdénombrement du recensement. En effet, le surdénombrement estimé par paires d'enregistrements serait biaisé positivement en présence de dénombrements triples ou plus élevés. Ainsi, des groupes mutuellement exclusifs d'enregistrements de la BDR reliés ont été établis; la majeure partie des groupes d'enregistrements de la base de sondage résultant en une ou deux paires (faisant intervenir deux ou trois enregistrements). Pour les cas où des groupes d'enregistrements contenaient plus de 10 liens, une méthode de théorie graphique a été appliquée pour réduire le groupe en petits sous-groupes appelés « voisinages » ([Dasylva et al., 2015](#)), afin de faciliter la vérification manuelle.

Enfin, la base de sondage de l'ESuR consistait en trois types d'unités d'échantillonnage : des paires, des groupes et des voisinages. Les unités d'échantillonnage ont été réparties en catégories selon trois types de processus :

- 1) CD seul, composé de paires et de groupes ou de voisinages d'enregistrements de la BDR provenant du CD;
- 2) CP seul, composé de paires et de groupes ou voisinages d'enregistrements de la BDR provenant du couplage

Rapport technique sur la couverture

BDR–BDR, du couplage BDR–ADM et des paires d’extension; 3) CP–CD, composé de groupes ou de voisinages d’enregistrements de la BDR provenant du CP et du CD (y compris des paires d’extension).

Tableau 8.3.2

Répartition des paires, des groupes et des voisinages issus du couplage déterministe seul, du couplage probabiliste seul et du couplage déterministe-couplage probabiliste dans la base de sondage de l’Étude sur le surdénombrement du Recensement de 2021

Type d’unité d’échantillonnage	Type de processus			Total
	CD seul	CP seul	CP-CD	
Groupe	4 822	1 635 296	86 930	1 727 048
Voisinage	64	161 641	6 493	168 198
Paire	345 243	5 931 084	0	6 276 327
Total	350 129	7 728 021	93 423	8 171 573

CD = couplage déterministe

CP = couplage probabiliste

CP-CD = couplage probabiliste-couplage déterministe (certaines paires du groupe ont été déterminées par le couplage probabiliste seul, et d’autres ont été déterminées par le couplage déterministe)

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

8.4 Plan d’échantillonnage

Le premier niveau de stratification a été obtenu par type de processus de couplage, ce qui a donné trois strates.

- La strate 1 comprenait des paires et des groupes ou voisinages du CD constitués uniquement de paires du CD. Cette strate a été traitée comme une strate à tirage complet et les unités d’échantillonnage de cette strate ont été classées comme des paires définitives de doublons de personnes.
- La strate 2 comprenait des paires et des groupes ou voisinages du CP constitués uniquement de paires du CP. Un échantillon probabiliste a été extrait de cette strate et les paires ont été envoyées en vérification manuelle.
- La strate 3 comprenait des groupes ou voisinages constitués d’une combinaison de paires du CP et du CD. Cette strate a ensuite été répartie en deux sous-strates. La première sous-strate était composée de groupes et voisinages comprenant au moins une paire du CD, échantillonnée dans le cadre de l’échantillon de vérification du CD utilisé pour confirmer la qualité de ces paires. Cette sous-strate a été traitée comme strate à tirage complet. La deuxième sous-strate était composée de groupes et voisinages ne comprenant pas de paire du CD, faisant partie de l’échantillon de vérification du CD. Elle comprenait un échantillon probabiliste de groupes ou voisinages CP–CD en ayant été extraits. Les paires du CP en groupes avec les paires du CD appartenant à la première sous-strate ont été envoyées pour vérification manuelle, ainsi que les paires du CP et du CD sélectionnées dans la deuxième sous-strate.

La taille d’échantillons ciblée a été d’environ 55 000 paires de la strate du CP seul et environ 4 500 paires de la strate CP–CD. Dans la présente section, intraprovincial signifie que tous les enregistrements de la BDR d’une unité d’échantillonnage proviennent de la même province ou du même territoire, et interprovincial signifie que les enregistrements de la BDR d’une unité d’échantillonnage proviennent de plusieurs provinces ou territoires. Les tableaux de cette section présentent les nombres de paires que l’unité d’échantillonnage soit une paire, un groupe ou un voisinage. Les groupes et voisinages sont décomposés selon leurs paires constituantes, afin de dériver le nombre de paires. À des fins de simplicité, des paires échantillonnées ont été envoyées en vérification manuelle plutôt que des groupes d’enregistrements.

Pour la strate du CP seul, les sous-strates de types d’unités d’échantillonnage ont été encore stratifiées en 14 strates : 13 strates provinciales comprenant des unités d’échantillonnage (paires ou groupes ou voisinages d’enregistrements interconnectés) au sein desquelles tous les enregistrements appartiennent à la même province ou au même territoire, et une strate interprovinciale au sein de laquelle les unités d’échantillonnage comprennent des enregistrements provenant des différentes provinces et différents territoires. Comme en 2016, les unités interprovinciales peuvent être des groupes comprenant également des paires intraprovinciales. Cela ne peut

être évité lors de l'utilisation de groupes d'enregistrements interconnectés. Pour mieux contrôler la taille de l'échantillon, les unités d'échantillonnage de groupe et de voisinage ont été encore stratifiées selon le nombre de paires dans le groupe. Enfin, les unités d'échantillonnage ont été triées selon la propension de surdénombrement estimé dans le cas de groupes ou de voisinages et leurs probabilités de correspondance conditionnelle dans le cas de paires; un échantillon systématique a alors été extrait¹².

Pour la première sous-strate CP-CD, les paires du CD faisant partie de l'échantillon de vérification avaient déjà été vérifiées et n'ont donc pas été envoyées en vérification manuelle. Cela a été avantageux et a permis un échantillon plus important de la sous-strate CP-CD avec au moins une paire du CD dans l'échantillon de vérification. Les groupes CP-CD pour lesquels aucune des paires du CD faisaient partie de l'échantillon de vérification ont été encore stratifiés en 14 strates : 13 strates intraprovinciales et une strate interprovinciale. Comme pour la strate du CP seul, ces 14 sous-strates ont en outre été stratifiées selon le nombre de liens, afin de mieux contrôler la taille de l'échantillon. Les unités d'échantillonnage ont ensuite été triées selon la propension estimée au surdénombrement avant qu'un échantillon systématique en soit tiré.

8.4.1 Strate du couplage déterministe seul

Comme cela est mentionné ci-dessus, les paires du CD seul et les groupes ou voisinages ont été jugés être des correspondances définitives et n'ont pas été envoyés en vérification manuelle. Comme le montre le [tableau 8.3.2](#), le nombre de groupes d'enregistrements interconnectés parmi les paires du CD était inférieur à celui des paires du CP. Au tableau 8.4.1.1, présentant la répartition des paires du CD seul, par province ou territoire, et les paires interprovinciales, on observe également moins de paires interprovinciales du CD seul que de paires interprovinciales du CP seul (1,39 % pour le tableau 8.4.1.1 contre 18,34 % pour le [tableau 8.4.2.2](#)). C'est ce à quoi l'on s'attend pour des paires étant des correspondances réelles.

Tableau 8.4.1.1
Fréquence des paires issues du couplage déterministe seul selon la province ou le territoire et la strate interprovinciale

Province ou territoire	Fréquence	Pourcentage
Terre-Neuve-et-Labrador	5 133	1,42
Île-du-Prince-Édouard	1 678	0,47
Nouvelle-Écosse	9 145	2,54
Nouveau-Brunswick	8 230	2,28
Québec	77 599	21,54
Ontario	122 913	34,12
Manitoba	11 908	3,31
Saskatchewan	13 159	3,65
Alberta	36 290	10,07
Colombie-Britannique	67 876	18,84
Yukon	463	0,13
Territoires du Nord-Ouest	494	0,14
Nunavut	391	0,11
Interprovinciale	5 001	1,39

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

12. Pour obtenir de plus amples détails sur le calcul de la propension de surdénombrement, veuillez demander le document interne « *Rapport méthodologique de l'Étude sur le surdénombrement du recensement de 2021* ». (2023) à Statistique Canada en écrivant à l'adresse infostats@statcan.gc.ca.

8.4.2 Strate du couplage probabiliste seul

Le tableau 8.4.2.1 présente le nombre de paires pour chaque type d'unité d'échantillonnage et une estimation du nombre d'unités d'échantillonnage nécessaires pour obtenir approximativement le même nombre de paires dans l'échantillon. La répartition en paires et groupes ou voisinages était proportionnelle à la taille.

Tableau 8.4.2.1
Fréquence des paires, des unités d'échantillonnage et des tailles d'échantillon selon le type d'unité d'échantillonnage

Type d'unité d'échantillonnage	Nombre de paires	Nombre d'unités d'échantillonnage	Taille d'échantillon (en nombre de paires)	Pourcentage d'échantillon (en nombre de paires)	Taille d'échantillon (en nombre d'unités d'échantillonnage)
Groupe ou voisinage	6 411 761	1 796 937	28 110	52	9 599
Paire	5 931 084	5 931 084	25 920	48	25 920
Total	12 342 845	7 728 021	54 030	100	35 519

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

Paires du couplage probabiliste seul

Les paires du CP seul ont d'abord été stratifiées par paires intraprovinciales et interprovinciales. Le tableau 8.4.2.2 ci-dessous donne la répartition des paires intraprovinciales et interprovinciales parmi les paires issues du CP seul. La répartition de l'échantillon en sous-strates intraprovinciales et interprovinciales était proportionnelle à la taille.

Tableau 8.4.2.2
Fréquence des paires intraprovinciales et interprovinciales parmi les paires issues du couplage probabiliste seul et tailles d'échantillon

Type de paires	Fréquence des paires	Pourcentage	Nombre de paires échantillonnées
Intraprovinciales	4 843 438	81,66	22 004
Interprovinciales	1 087 646	18,34	4 753

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

Dans la strate des paires intraprovinciales, une répartition par la méthode de puissance est utilisée pour répartir les paires issues du CP seul entre les provinces; la taille étant considérée comme équivalant au nombre de paires dans chaque province et $q = \frac{1}{2}$. Les paires ont ensuite été triées selon leurs probabilités de correspondance conditionnelles avant qu'un échantillon systématique n'en soit tiré. Prenez note que les trois territoires sont à tirage complet. Le tableau 8.4.2.3 présente la répartition des paires intraprovinciales issues du CP seul selon la province ou le territoire.

Tableau 8.4.2.3

Répartition de l'échantillon des paires intraprovinciales issues du couplage probabiliste seul selon la province ou le territoire

Province ou territoire	Fréquence
Terre-Neuve-et-Labrador	568
Île-du-Prince-Édouard	300
Nouvelle-Écosse	760
Nouveau-Brunswick	748
Québec	6 915
Ontario	5 600
Manitoba	730
Saskatchewan	701
Alberta	1 740
Colombie-Britannique	2 655
Yukon	383
Territoires du Nord-Ouest	448
Nunavut	456
Taille d'échantillon totale	22 004

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

Les paires interprovinciales issues du CP seul ont ensuite été stratifiées selon la combinaison distincte de provinces et réparties proportionnellement à leur taille. Parmi les paires interprovinciales, on compte 78 combinaisons uniques de provinces. Dans les sous-strates de combinaisons de provinces, les paires ont été triées selon leurs probabilités de correspondance conditionnelles et une méthode d'échantillonnage systématique a été utilisée pour en extraire l'échantillon.

Groupes et voisinages du couplage probabiliste seul

Pour les groupes et voisinages, les paires ont d'abord été stratifiées par groupes intraprovinciaux et interprovinciaux. Un groupe a été jugé être interprovincial s'il contenait au moins une paire interprovinciale. Le tableau 8.4.2.4 présente la répartition des groupes intraprovinciaux et interprovinciaux dans la strate issu du CP seul. L'échantillon a été réparti proportionnellement à la taille entre les strates intraprovinciale et interprovinciale.

Tableau 8.4.2.4

Fréquence des groupes ou des voisinages intraprovinciaux et interprovinciaux et des tailles d'échantillon

Type de groupe	Fréquence des paires	Pourcentage	Nombre de paires échantillonnées	Nombre de groupes échantillonnés
Intraprovinciaux	4 070 750	63,49	17 970	6 656
Interprovinciaux	2 341 011	36,51	10 140	3 022

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

Au sein de la strate intraprovinciale, les groupes ont été répartis par provinces à l'aide d'une répartition par la méthode de puissance. Le tableau 8.4.2.5 présente la répartition des unités d'échantillonnage intraprovinciales issues du CP seul selon la province ou le territoire. Étant donné le faible nombre d'unités d'échantillonnage dans les territoires, ces sous-strates étaient à tirage complet. Pour mieux contrôler la taille finale de l'échantillon, les strates provinciales ont en outre été stratifiées selon la taille du groupe en fonction du nombre de paires dans le groupe. L'échantillon de chaque strate provinciale a été réparti entre les groupes de manière proportionnelle à leur taille. Au moins une unité d'échantillonnage a été tirée de chaque strate.

Rapport technique sur la couverture

Tableau 8.4.2.5
Répartition des unités d'échantillonnage intraprovinciales issues du couplage probabiliste seul selon la province et le territoire

Niveau du groupe (province ou territoire)	Nombre de paires échantillonnées	Nombre de groupes échantillonnés
Terre-Neuve-et-Labrador	270	112
Île-du-Prince-Édouard	100	43
Nouvelle-Écosse	439	187
Nouveau-Brunswick	450	184
Québec	7 594	2 514
Ontario	4 949	1 899
Manitoba	385	162
Saskatchewan	342	143
Alberta	1 119	467
Colombie-Britannique	2 123	866
Yukon	54	24
Territoires du Nord-Ouest	68	27
Nunavut	64	28
Taille d'échantillon totale	17 957	6 656

Note : Les trois territoires sont des strates à tirage complet.

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

La strate des groupes et des voisinages interprovinciaux a été divisée en deux sous-strates : ceux comprenant une province ou un territoire majoritaire (c.-à-d. la plupart des paires du groupe appartiennent à une province ou un territoire) et ceux sans province ou territoire majoritaire (c.-à-d. les paires au sein du groupe sont réparties également entre les provinces ou les territoires concernés). La répartition des paires selon la présence ou l'absence d'un groupe ou d'un voisinage majoritaire est présentée au tableau 8.4.2.6.

Tableau 8.4.2.6
Fréquence des paires au sein des groupes et des voisinages interprovinciaux issues du couplage probabiliste seul selon les groupes avec une province ou un territoire majoritaire et les groupes sans province ou territoire majoritaire et taille des échantillons

Type de groupe	Fréquence des paires	Pourcentage	Nombre de paires échantillonnées	Nombre de groupes échantillonnés
Avec province ou territoire majoritaire	515 949	90	9 099	2 702
Sans province ou territoire majoritaire	57 923	10	1 041	422

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

Les groupes et voisinages interprovinciaux avec province ou territoire majoritaire ont été ensuite encore stratifiés selon la province ou le territoire majoritaire dans le groupe par répartition par la méthode de puissance. Les unités d'échantillonnage au sein des sous-strates provinciales ont ensuite été stratifiées selon le nombre de paires dans les groupes. La répartition aux groupes était proportionnelle à la taille. Les unités d'échantillonnage ont ensuite été triées selon le surdénombrement attendu dans le groupe et la proportion de paires intraprovinciale au sein du groupe, avant qu'un échantillon systématique en soit tiré. Étant donné que seuls 102 groupes comptaient un territoire majoritaire, ces strates ont été à tirage complet. Au moins quatre unités d'échantillonnage ont été tirées des autres strates. Le tableau 8.4.2.7 présente la répartition des unités d'échantillonnage interprovinciales du CP seul ayant une province ou un territoire majoritaire selon la province ou le territoire majoritaire.

Tableau 8.4.2.7

Répartition des unités d'échantillonnage interprovinciales issues du couplage probabiliste seul ayant une province ou un territoire majoritaire selon la province ou le territoire majoritaire

Niveau du groupe (province ou territoire)	Nombre de paires échantillonnées	Nombre de groupes échantillonnés
Terre-Neuve-et-Labrador	282	91
Île-du-Prince-Édouard	127	37
Nouvelle-Écosse	481	156
Nouveau-Brunswick	433	144
Québec	2 000	540
Ontario	2 593	685
Manitoba	346	108
Saskatchewan	264	99
Alberta	898	292
Colombie-Britannique	1 399	448
Yukon	129	49
Territoires du Nord-Ouest	102	39
Nunavut	36	14
Taille d'échantillon totale	9 090	2 702

Note : Les trois territoires sont des strates à tirage complet.

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

Les groupes sans province ou territoire majoritaire ont été stratifiés selon la taille du groupe, et l'échantillon a été réparti proportionnellement à la taille. Le tableau 8.4.2.8 montre la répartition des groupes interprovinciaux issus du CP seul sans province ou territoire majoritaire selon la taille du groupe.

Tableau 8.4.2.8

Répartition des groupes interprovinciaux issus du couplage probabiliste seul sans province ou territoire dominant selon le nombre de paires dans le groupe

Nombre de paires	Nombre de paires échantillonnées	Nombre de groupes échantillonnés
2	548	274
3	372	124
4	36	9
5	55	11
6	6	1
7	7	1
8	8	1
9	9	1
Taille d'échantillon totale	1 041	422

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

8.4.3 Strate du couplage probabiliste-couplage déterministe

La répartition des paires issues du CP et du CD dans les groupes et les voisinages du CP-CD est présentée au tableau 8.4.3.1.

Tableau 8.4.3.1

Paires issues du couplage probabiliste et paires issues du couplage déterministe dans les groupes et les voisinages du couplage probabiliste-couplage déterministe

Couplées par	Fréquence	Pourcentage
CP	250 270	70,64
CD	104 020	29,36

CP = couplage probabiliste

CD = couplage déterministe

Note : Le terme « couplage probabiliste-couplage déterministe » signifie que certaines paires dans le groupe ont été déterminées par le couplage probabiliste seul, et d'autres ont été déterminées par le couplage déterministe.

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

Comme cela a été mentionné précédemment, un échantillon de paires du CD a été extrait au cours de l'étape du CD et envoyé en vérification manuelle pour évaluer la qualité des paires du CD et veiller à ce que toutes les paires du CD puissent être classées comme paires définitives de doublons de personnes. On a appelé cet échantillon, l'échantillon de vérification du CD. Pour utiliser cet échantillon de vérification du CD, les groupes auxquels appartenaient ces paires échantillonnées ont été traités comme strates à tirage complet et les paires correspondantes du CP, ainsi que toutes les paires correspondantes du CD ne faisant pas partie de l'échantillon de vérification du CD, ont été envoyées en vérification manuelle.

On a compté 1 010 paires du CD échantillonnées parmi les paires des groupes d'enregistrements interconnectés du CP-CD. Ces paires appartenaient à 929 groupes. La répartition des paires issues du CP et du CD parmi ces 929 groupes est présentée au tableau 8.4.3.2.

Tableau 8.4.3.2

Paires issues du couplage probabiliste et paires issues du couplage déterministe parmi les 929 groupes du couplage probabiliste-couplage déterministe qui contenaient des paires issues du couplage déterministe faisant partie de l'échantillon de vérification du couplage déterministe

Couplées par	Fréquence	Pourcentage
CP	2 553	71,31
CD	1 027	28,69

CP = couplage probabiliste

CD = couplage déterministe

Note : Le terme « couplage probabiliste-couplage déterministe » signifie que certaines paires dans le groupe ont été déterminées par le couplage probabiliste seul, et d'autres ont été déterminées par le couplage déterministe.

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

Au total, 17 paires issues du CD et 2 553 paires issues du CP ont été envoyées en vérification manuelle. Les 1 010 paires issues du CD qui faisaient partie de l'échantillon de vérification du CD avaient déjà été vérifiées. Elles n'ont, par conséquent, pas été envoyées en vérification manuelle.

Un échantillon supplémentaire de 533 groupes (1 930 paires) a été sélectionné dans la strate CP-CD. La strate CP-CD a été stratifiée par province ou territoire au niveau du groupe et taille du groupe, et l'échantillon a été sélectionné de sorte que l'échantillon CP-CD complet soit approximativement proportionnel à la taille. La répartition provinciale au niveau de la paire de l'échantillon CP-CD complet est présentée au [tableau 8.4.3.3](#).

Tableau 8.4.3.3

Répartition des paires dans l'échantillon du couplage probabiliste-couplage déterministe selon la province ou le territoire au niveau de la paire

Province ou territoire	Fréquence	Pourcentage
Terre-Neuve-et-Labrador	62	1,13
Île-du-Prince-Édouard	32	0,58
Nouvelle-Écosse	87	1,58
Nouveau-Brunswick	105	1,91
Québec	1 583	28,76
Ontario	1 683	30,57
Manitoba	91	1,65
Saskatchewan	83	1,51
Alberta	311	5,65
Colombie-Britannique	714	12,97
Yukon	15	0,27
Territoires du Nord-Ouest	15	0,27
Nunavut	9	0,16
Interprovinciaux	715	12,99
Total	5 505	100,00

Note : Le terme « couplage probabiliste-couplage déterministe » signifie que certaines paires dans le groupe ont été déterminées par le couplage probabiliste seul, et d'autres ont été déterminées par le couplage déterministe.

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

Les strates provinciales ont ensuite été stratifiées selon le nombre de liens et un échantillon systématique en a été tiré.

8.4.4 Tailles d'échantillon finales (par paire)

Le tableau 8.4.4.1 ci-dessous présente les tailles d'échantillon finales pour les strates du CP seul et du CP-CD ayant été envoyées en vérification manuelle. La strate du CD seul comprenait 360 280 paires classées comme paires définitives de doublons de personnes.

Tableau 8.4.4.1

Tailles d'échantillon finales pour les strates du couplage probabiliste seul et du couplage probabiliste-couplage déterministe qui sont envoyées pour vérification manuelle

Strate	Nombre de paires selon la strate	Nombre de paires échantillonnées envoyées pour vérification manuelle (après suppression des paires en double dans les voisinages qui se chevauchent)
CP-CD (sans 1 010 paires du CD faisant partie de l'échantillon de vérification du CD)	92 494	4 495
Paires interprovinciales du CP seul	1 087 646	4 753
Paires intraprovinciales du CP seul	4 844 635	22 004
Groupes intraterritoriaux du CP seul (tirage complet)	4 070 750	186
Groupes intraprovinciaux du CP seul	484	18 153
Groupes interterritoriaux du CP seul avec territoire majoritaire (tirage complet)	266	266

Tableau 8.4.4.1

Tailles d'échantillon finales pour les strates du couplage probabiliste seul et du couplage probabiliste-couplage déterministe qui sont envoyées pour vérification manuelle

Strate	Nombre de paires selon la strate	Nombre de paires échantillonnées envoyées pour vérification manuelle (après suppression des paires en double dans les voisinages qui se chevauchent)
Groupes interprovinciaux du CP seul avec province majoritaire	2 184 665	8 815
Groupes interprovinciaux du CP seul sans province ou territoire majoritaire	156 079	1 040
Taille totale	12 437 019	59 712

CP = couplage probabiliste

CD = couplage déterministe

CP-CD = couplage probabiliste-couplage déterministe (certaines paires dans le groupe ont été déterminées par le couplage probabiliste seul, et d'autres ont été déterminées par le couplage déterministe)

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

8.5 Opération de vérification manuelle

L'opération de vérification manuelle a été une tâche de bureau répondant à plusieurs objectifs :

- effectuer une vérification indépendante des paires échantillonnées, afin de déterminer si elles constituaient des cas de surdénombrement;
- examiner les membres de ménage associés aux paires échantillonnées pour relever des cas potentiels de surdénombrement supplémentaire absents de la base de sondage de l'ESuR;
- coder la cause potentielle du cas de surdénombrement (c.-à-d. le scénario de surdénombrement).

La vérification manuelle a été effectuée paire par paire. Lorsqu'un groupe ou un voisinage était échantillonné, toutes les paires qui le constituaient étaient examinées manuellement. Toutefois, les codeurs ne disposaient pas des renseignements relatifs au regroupement pour les paires au sein des groupes et voisinages. Chaque paire a été vérifiée seule. Les paires n'étaient examinées qu'une fois, même si elles appartenaient à plus d'un voisinage échantillonné.

Le processus de vérification manuelle consistait en un examen exhaustif de toute l'information disponible dans la BDR. Comme en 2016, il comprenait les étapes suivantes :

1. comparer les personnes échantillonnées de la BDR en fonction du nom, du sexe, de la date de naissance et des relations, ainsi que d'autres renseignements supplémentaires ajoutés en 2021;
2. comparer les membres des ménages de la BDR en fonction des mêmes critères;
3. évaluer les résultats permettant de conclure ou non au surdénombrement entre deux personnes de la BDR et entre deux ménages de la BDR;
4. déterminer le scénario de surdénombrement en cas de surdénombrement (le tableau 8.5.1 fournit une liste des codes de scénarios de surdénombrement et leur description).

Tableau 8.5.1

Codes de scénario de surdénombrement

Code	Description
1.1	Deux FRAME_ID différents pour le même ménage; même adresse ou adresse similaire
1.2	Deux FRAME_ID différents pour le même ménage; adresses différentes
2.1	Enfant de parents dans des ménages séparés
2.2	Enfant (de 0 à 17 ans) avec autre(s) membre(s) de la famille
2.3	Enfant (de 0 à 17 ans) avec autre(s) adulte(s) non apparenté(s)

Tableau 8.5.1
Codes de scénario de surdénombrement

Code	Description
3.1	Étudiant ou jeune adulte (de 18 à 24 ans) ayant récemment quitté le domicile familial
3.2	Jeune adulte (de 18 à 24 ans) entrant/quittant une relation de mariage/conjoint de fait
3.3	Jeune adulte (de 18 à 24 ans) avec autre(s) membre(s) de la famille
3.4	Jeune adulte (de 18 à 24 ans) avec autre(s) adulte(s) non apparenté(s)
4.1	Adulte (de 25 ans ou plus) ayant récemment quitté le domicile familial
4.2	Adulte (de 25 ans ou plus) entrant/quittant une relation de mariage/conjoint de fait
4.3	Adulte (de 25 ans ou plus) avec autre(s) membre(s) de la famille
4.4	Adulte (de 25 ans ou plus) avec autre(s) adulte(s) non apparenté(s)
5.1	Un des ménages n'est pas une résidence privée
6.1	Surdénombrement intraménage (même FRAME_ID)
7.1	Autre

FRAME_ID = identificateur unique de ménage

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

L'échantillon a été divisé en lots de 500 paires de ménages (ménage A et ménage B). Chaque lot a été attribué à un agent (vérificateur), qui a procédé à l'examen et décidé si la personne choisie du ménage A était un doublon de la personne choisie (surdénombree) du ménage B pour chaque paire de ménages du lot. Une paire d'enregistrements sélectionnée était la paire échantillonnée d'intérêt. De plus, le vérificateur relevait des paires supplémentaires de doublons de personnes (le cas échéant) contenues dans chaque paire de ménages et dans chaque ménage.

En cas d'incertitude quant au codage d'un cas, les vérificateurs avaient reçu la directive de le transmettre à un superviseur, qui à son tour consultait l'équipe de la qualité des données (équipe de spécialistes de domaine de la Section de la mesure de la couverture de la Division des méthodes d'intégration statistique) ou lui transmettait le cas. En 2021, certaines paires échantillonnées complexes ont été envoyées directement à l'équipe de la qualité des données aux fins de vérification. Les cas échantillonnés complexes comprenaient :

- des cas intraménage, c'est-à-dire, lorsque la paire fait partie d'un même ménage (p. ex. même personne deux fois dans la liste);
- des ménages d'une seule personne (lorsque la paire est issue de deux ménages différents; chacun de taille 1).

Les cycles antérieurs ont révélé que ces cas complexes échantillonnés nécessitaient l'expertise de l'équipe de la qualité des données pour les coder correctement. Cette équipe pouvait également consulter des sources supplémentaires de renseignements lui permettant de prendre une décision exacte, comme consulter les données du questionnaire du cycle du recensement actuel et/ou antérieur et utiliser les couplages effectués par l'équipe de l'ECDS à Statistique Canada. Tous les cas échantillonnés devaient être codés avec certitude, puisqu'aucune non-réponse n'était autorisée.

La confiance dans les résultats codés était requise pour l'opération manuelle, puisque les résultats contribuaient directement à l'estimation du surdénombrement. Ainsi, une vérification à 100 % a été mise en œuvre. Cela signifie que deux vérificateurs différents codaient le même lot. Une fois un lot codé par deux vérificateurs différents, leurs résultats étaient comparés. Tous les champs codés étaient comparés. Si l'un des codes ne correspondait pas, le cas était envoyé à l'équipe de la qualité des données pour qu'une décision éclairée soit prise. Cette stratégie de vérification à 100 % visait à assurer des résultats codés de grande qualité et une rétroaction continue était en outre fournie aux agents tout au long de l'opération de vérification manuelle.

8.6 Pondération et estimation

8.6.1 Pondération

Le poids de départ d'une unité d'échantillonnage était simplement l'inverse de sa probabilité de sélection. Les unités d'échantillonnage qui étaient des groupes ou des voisinages variaient en termes de nombre de paires qu'elles contenaient. Ces unités étaient stratifiées selon le nombre de paires lors de l'échantillonnage, afin de mieux contrôler la taille finale de l'échantillon. Toutefois, pour les groupes et voisinages interprovinciaux, les chiffres provinciaux ou territoriaux pondérés pouvaient différer de la base de sondage. Ainsi, une étape de calage a été ajoutée afin d'assurer une représentation adéquate du nombre de paires dans chaque province et territoire. Les poids de sondage des groupes et des voisinages interprovinciaux ont été calés, afin que le nombre estimé de paires intraprovinciales et interprovinciales dans chaque province ou territoire soit égal au total correspondant sur la base. Le Système généralisé d'estimation (G-Est) de Statistique Canada a été utilisé pour effectuer le calage. Le tableau 8.6.1.1 montre les facteurs de calage pour chaque province et territoire.

Tableau 8.6.1.1

Facteur de calage moyen (rapport entre le total sur la base et l'estimation pondérée) selon la strate et le type de paire pour les groupes et les voisinages intraprovinciaux et interprovinciaux

Province ou territoire	Intraprovinciaux	Interprovinciaux
Terre-Neuve-et-Labrador	0,76	0,74
Île-du-Prince-Édouard	0,69	1,30
Nouvelle-Écosse	1,01	0,96
Nouveau-Brunswick	0,99	0,88
Québec	0,98	1,08
Ontario	0,99	0,99
Manitoba	1,36	1,07
Saskatchewan	1,13	0,90
Alberta	1,08	1,08
Colombie-Britannique	1,07	0,97
Yukon	1,47	1,21
Territoires du Nord-Ouest	1,40	0,42
Nunavut	1,22	2,80

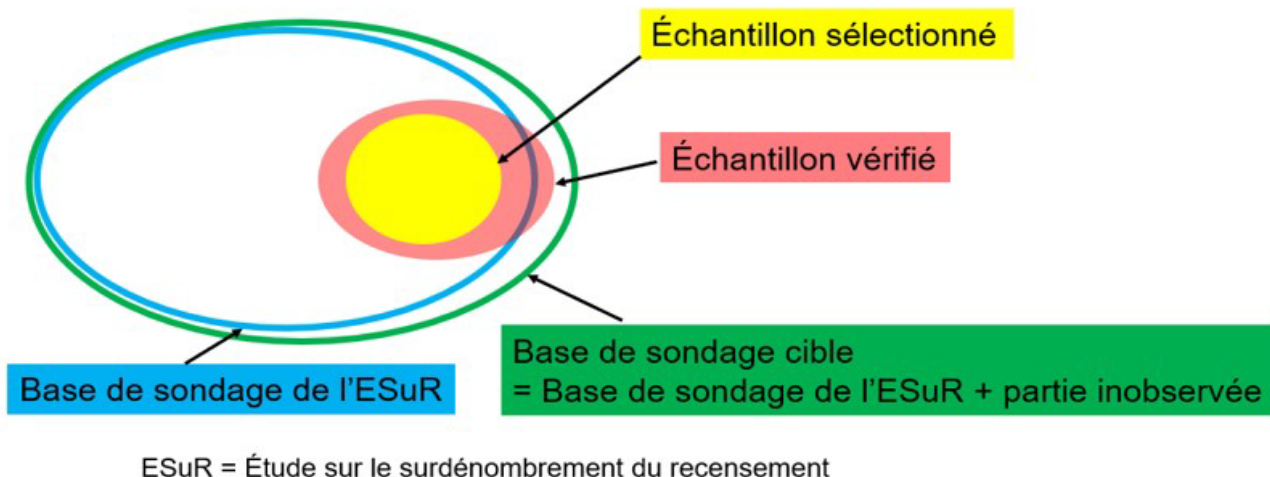
Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

Au cours de l'opération de vérification manuelle, les vérificateurs ont relevé les cas de surdénombrement dans les ménages de paires échantillonnées non couvertes par la base de sondage de l'ESuR; ces paires de doublons de personnes ont été appelées des paires supplémentaires de surdénombrement trouvées au cours de la vérification manuelle. Cela se produisait lorsque les différences entre les deux enregistrements étaient trop importantes pour que la paire ait été capturée par les processus de couplage. Par exemple, en cas de plusieurs coquilles, erreurs ou trop de différences dans les champs de couplage utilisés par le processus de couplage, la paire de surdénombrement ne figurait pas dans la base de sondage de l'ESuR.

Cette situation est illustrée à la figure 1 ci-dessous. L'ovale avec un contour bleu représente la base de sondage de l'ESuR tandis que l'ovale avec un contour vert représente la base de sondage cible, qui comprend un petit nombre de paires n'ayant pas pu être capturées par les processus de couplage (c.-à-d. la partie non observée de la base de sondage cible). L'ovale jaune plein représente l'échantillon sélectionné, qui comprend les paires de personnes échantillonnées, tandis que l'ovale rouge plein représente l'échantillon vérifié, qui comprend les paires de personnes échantillonnées et les membres de leur ménage. Aucun poids n'est directement associé aux paires de l'ovale rouge plein qui se trouvent en dehors de la base de sondage de l'ESuR (c.-à-d. une petite partie de l'ovale rouge se trouve dans la partie inobservée de la base de sondage cible). La Méthode généralisée de partage des poids (MGPP) (Lavallée, P. 2007) a été utilisée pour attribuer des poids à partir des poids des paires

échantillonnées, par lesquelles ces paires ont été indirectement échantillonnées. Ainsi, un poids a été dérivé pour toutes les paires supplémentaires du surdénombrement trouvées au cours de la vérification manuelle et ces paires ont été ajoutées à l'échantillon aux fins d'estimation. Cela a remplacé l'ajustement fondé sur l'EAA, qui tenait compte du surdénombrement mesuré par l'EAA hors de la base de sondage de l'ESuR. Ce procédé était utilisé depuis l'ESuR de 2006.

Figure 1
Illustration de l'échantillon sélectionné, de l'échantillon vérifié, de la base de sondage de l'Étude sur le surdénombrement du recensement et de la base de sondage cible



Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

Des limitations étaient associées à la manière de détecter des paires supplémentaires de surdénombrement. Les doublons de ménages comptant une seule personne ou les doublons de personnes pour lesquelles les autres membres du ménage n'ont rien en commun au sein de la partie inobservée de la base de sondage cible ne seraient pas relevés lors de la vérification manuelle de tous les membres du ménage d'une paire échantillonnée. Ainsi, on reconnaît que l'ESuR de 2021 peut ne pas toujours représenter l'ensemble de la base de sondage cible des doublons de personnes dans le recensement. Cela aurait également été le cas lors de l'utilisation de l'EAA pour ajuster l'ESuR dans les cycles antérieurs. Cependant, la partie inobservée de la base de sondage cible devrait être extrêmement petite.

8.6.2 Estimation

Les résultats de l'opération de vérification manuelle ont été traités, afin de créer des groupes de surdénombrement servant à l'estimation. Les groupes de surdénombrement étaient formés de tous les enregistrements de la BDR couplés ensemble par un surdénombrement vérifié. Les estimations de l'ESuR se sont basées sur la somme des estimations de surdénombrement comptabilisées dans chaque groupe de surdénombrement. Pour un groupe de surdénombrement qui était une paire, le compte de surdénombrement était simplement de 1. Si le groupe de surdénombrement figurait dans un petit groupe d'enregistrements (c.-à-d. groupe non réparti en voisinages), alors :

Le compte de surdénombrement était égal au nombre d'enregistrements dans le groupe de surdénombrement – 1.

Rapport technique sur la couverture

Pour les groupes de surdénombrement répartis en voisinages, le surdénombrement était compté en suivant les deux étapes ci-après :

1. Calculer le surdénombrement dans chaque voisinage dont l'ancrage (c.-à-d. l'enregistrement de la BDR tenant lieu de centre du voisinage) intervenait dans le surdénombrement vérifié pour ce groupe de surdénombrement, de la façon suivante :

$$\text{Surdénombrement pour le voisinage} = \frac{(\text{nombre d'enregistrements appartenant au groupe de surdénombrement} - 1)}{\text{nombre d'enregistrements appartenant au groupe de surdénombrement}}$$

2. Additionner le surdénombrement de chaque voisinage pour obtenir le surdénombrement total du groupe de surdénombrement.

Le surdénombrement pour un domaine était obtenu en calculant au prorata le surdénombrement total de la paire, du groupe ou du voisinage par la proportion d'enregistrements de la BDR faisant partie du domaine parmi ceux appartenant au groupe de surdénombrement.

Pour les groupes et les voisinages interprovinciaux, le surdénombrement calculé pour une unité était multiplié par le poids après calage pour obtenir l'estimation pondérée. Les paires supplémentaires de surdénombrement relevées au cours de la vérification manuelle étaient multipliées par leur poids d'échantillonnage dérivé par la MGPP, afin d'obtenir l'estimation pondérée. Sinon, le surdénombrement calculé pour une unité était multiplié par son poids de sondage initial pour obtenir l'estimation pondérée. La variance de l'estimation était calculée à l'aide de G-Est.

8.7 Résultats

L'ESuR de 2021 a estimé que 755 635 personnes ont été dénombrées plusieurs fois dans le cadre du Recensement de la population de 2021. Ces résultats ont été examinés selon chacune des composantes ayant mené à la création de la base de sondage et sa contribution à l'estimation générale du surdénombrement du recensement. Les raisons potentielles de ces dénombrements multiples de personnes dans le cadre du recensement ont également été examinées.

8.7.1 Surdénombrement par composante

Les composantes de l'ESuR ont caractérisé chaque cas de surdénombrement (définitif ou vérifié manuellement) en déterminant les paires présentes dans l'unité d'échantillonnage. Il en existe quatre types :

- CD seul : toutes les paires du groupe de surdénombrement identifiées par le CD;
- CP seul : toutes les paires du groupe de surdénombrement identifiées uniquement par le CP;
- CP-CD : certaines paires du groupe ont été identifiées uniquement par le CP, alors que le CD en a détecté d'autres;
- vérification manuelle du surdénombrement (VMSD) : toutes les paires dans le groupe de surdénombrement étaient des paires supplémentaires de doublons de personnes identifiées au cours de la vérification manuelle qui ne se retrouvaient pas dans la base de sondage de l'ESuR et pour lesquelles un poids d'échantillonnage indirect a été dérivé à l'aide de la MGPP.

Il est important de ne pas oublier que les paires relevées par les deux étapes du CP et du CD ont été classées dans la catégorie CD; par conséquent la catégorie « CD seul » comprend tous les groupes constitués uniquement de paires détectées par le CD, même si certaines de ces mêmes paires ont également pu être relevées par le CP.

Rapport technique sur la couverture

Le tableau 8.7.1.1 présente le nombre de cas de surdénombrement estimés par chacune des composantes de l'ESuR, ainsi que le pourcentage du surdénombrement total estimé représenté, pour le Canada ainsi que par province ou territoire.

Tableau 8.7.1.1
Contribution de chaque composante de l'Étude sur le surdénombrement du Recensement de 2021 au surdénombrement total estimé pour chaque province et territoire

Province ou territoire	CD seul		CP seul		CP-CD		VMSD		Total	
	Nombre estimé	% du total	Nombre estimé	% du total	Nombre estimé	% du total	Nombre estimé	% du total	Nombre estimé	Erreur- type
Canada	352 059	46,6	318 459	42,1	81 172	10,7	3 946	0,5	755 635	9 648
Terre-Neuve-et- Labrador	5 148	50,5	4 664	45,8	382	3,7	0	0,0	10 194	439
Île-du-Prince- Édouard	1 678	51,0	1 284	39,0	311	9,5	16	0,5	3 289	191
Nouvelle-Écosse	9 200	47,6	8 412	43,5	1 639	8,5	94	0,5	19 344	736
Nouveau- Brunswick	8 079	48,8	6 890	41,7	1 440	8,7	132	0,8	16 541	641
Québec	76 760	42,1	80 126	43,9	25 242	13,8	385	0,2	182 513	5 915
Ontario	120 765	44,7	118 619	43,9	29 382	10,9	1 334	0,5	270 100	6 888
Manitoba	11 930	51,5	10 231	44,2	970	4,2	29	0,1	23 160	757
Saskatchewan	13 210	54,7	9 501	39,3	1 194	4,9	258	1,1	24 163	689
Alberta	36 902	47,3	35 085	44,9	5 527	7,1	570	0,7	78 084	2 736
Colombie- Britannique	66 976	53,2	42 762	34,0	15 017	11,9	1 078	0,9	125 832	2 778
Yukon	479	57,9	315	38,1	21	2,5	12	1,4	827	38
Territoires du Nord-Ouest	508	60,6	293	35,0	25	3,0	12	1,4	837	15
Nunavut	423	56,3	277	36,9	23	3,1	28	3,7	751	17

CD seul = toutes les paires du groupe de surdénombrement identifiées par le couplage déterministe

CP seul = toutes les paires du groupe de surdénombrement identifiées uniquement par le couplage probabiliste

CP-CD = certaines paires du groupe ont été identifiées uniquement par le couplage probabiliste, alors que le couplage déterministe en a détecté d'autres

VMSD = toutes les paires dans le groupe de surdénombrement étaient des paires supplémentaires de doublons de personnes identifiées au cours de la vérification manuelle qui ne se retrouvaient pas dans la base de sondage de l'ESuR et pour lesquelles un poids d'échantillonnage indirect a été dérivé à l'aide de la méthode généralisée du partage des poids

Note : En raison de l'arrondissement, la somme des estimations de la couverture pourrait ne pas correspondre au total.

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

Au niveau national, les composantes CD seul et CP seul représentaient respectivement 46,6 % et 42,1 % du total du surdénombrement estimé, alors que la composante CP-CD représentait 10,7 %, et la composante VMSD représentait 0,5 %.

La contribution de la composante CD seul à l'estimation provinciale ou territoriale totale était supérieure pour les Territoires du Nord-Ouest (60,6 %) et le Yukon (57,9 %) et inférieure pour l'Ontario (44,7 %) et le Québec (42,1 %). La contribution de la composante CP seule à l'estimation provinciale ou territoriale totale était supérieure pour Terre-Neuve-et-Labrador (45,8 %) et l'Alberta (44,9 %) et inférieure pour les Territoires du Nord-Ouest (35,0 %) et la Colombie-Britannique (34,0 %). La contribution de la composante CP-CD était supérieure au Québec (13,8 %) et en Colombie-Britannique (11,9 %) et inférieure pour les trois territoires (de 2,5 % à 3,1 %). Enfin, la contribution de la composante de VMSD était supérieure pour les trois territoires (de 1,4 % à 3,7 %) et inférieure pour le Manitoba (0,1 %) et Terre-Neuve-et-Labrador (0,0 %), où aucune paire supplémentaire de doublons de personnes n'a été relevée au cours de l'opération de vérification manuelle sans qu'elle ne figure déjà dans la base de sondage de l'ESuR.

8.7.2 Surdénombrement par scénario

Le [tableau 8.7.2.1](#) présente le surdénombrement estimé selon la raison possible de ce surdénombrement (appelé scénario de surdénombrement) au niveau national et au niveau des provinces et des territoires pour 2021. Il est important de mentionner que ces résultats ne sont pas comparables aux résultats du surdénombrement de 2016 par scénario pour deux raisons.

- Le scénario de surdénombrement a été codé au cours de l'opération de vérification manuelle. Puisque les paires issues du CD seul ont été jugées être des paires définitives de doublons de personnes sans vérification manuelle, un scénario de surdénombrement n'est pas disponible pour ces paires.
- Les codes utilisés pour les scénarios ont été modifiés pour le cycle de 2021, afin d'améliorer la cohérence du codage et l'utilité des résultats.

À l'exclusion des cas issus du CD seul, près de 25 % de tout le surdénombrement au niveau national se situe entre deux ménages identiques. Cette proportion est un peu inférieure pour Terre-Neuve-et-Labrador et supérieure pour la Colombie-Britannique.

Lorsque l'on ne considère que le surdénombrement au sein de ménages non identiques et que l'on exclut à nouveau les cas issus du CD seul, le scénario de surdénombrement le plus fréquent est un enfant dénombré par les deux parents de ménages distincts, comme c'était le cas en 2016 et au cours des cycles précédents. Cela s'observe pour toutes les provinces et tous les territoires, à l'exception de la Nouvelle-Écosse et du Nunavut. En Nouvelle-Écosse, le scénario le plus fréquent était un étudiant ou jeune adulte (âgé de 18 à 24 ans) ayant récemment quitté le domicile, alors qu'au Nunavut, il s'agissait d'un enfant (âgé de 0 à 17 ans) avec d'autres personnes apparentées.

Tableau 8.7.2.1
Répartition du surdénombrement du Recensement de 2021 selon le scénario pour chaque province et territoire

Province ou territoire	Scénario de surdénombrement				
	1.1	2.1	2.2	2.3	3.1
	Ménages identiques	Enfant de parents dans des ménages séparés	Enfant (de 0 à 17 ans) avec autre(s) membre(s) de la famille	Enfant (de 0 à 17 ans) avec autre(s) adulte(s) non apparenté(s)	Étudiant ou jeune adulte (de 18 à 24 ans) ayant récemment quitté le domicile familial
	pourcentage				
Canada	12,5	11,3	0,8	0,3	5,8
Terre-Neuve-et-Labrador	9,5	11,5	1,2	0,0	7,7
Île-du-Prince-Édouard	11,2	11,5	0,3	0,4	9,0
Nouvelle-Écosse	10,8	9,9	2,0	0,0	15,5
Nouveau-Brunswick	11,0	10,5	1,7	0,5	6,4
Québec	12,1	16,6	0,7	0,2	6,0
Ontario	13,0	10,6	0,5	0,1	5,4
Manitoba	10,5	8,8	2,4	0,5	5,1
Saskatchewan	9,3	11,1	1,5	0,8	4,0
Alberta	11,5	9,3	0,7	0,8	6,6
Colombie-Britannique	14,6	7,0	0,7	0,5	4,7
Yukon	9,2	13,1	0,4	0,2	4,0
Territoires du Nord-Ouest	10,8	8,2	2,1	1,0	1,4
Nunavut	13,6	4,5	8,8	1,2	1,6

Rapport technique sur la couverture

Tableau 8.7.2.1

Répartition du surdénombrement du Recensement de 2021 selon le scénario pour chaque province et territoire

Province ou territoire	Scénario de surdénombrement					
	3.2	3.3	3.4	4.1	4.2	4.3
	Jeune adulte (de 18 à 24 ans) entrant/quittant une relation de mariage/conjoint de fait	Jeune adulte (de 18 à 24 ans) avec autre(s) membre(s) de la famille	Jeune adulte (de 18 à 24 ans) avec autre(s) adulte(s) non apparenté(s)	Adulte (de 25 ans ou plus) ayant récemment quitté le domicile familial	Adulte (de 25 ans ou plus) entrant/quittant une relation de mariage/conjoint de fait	Adulte (de 25 ans ou plus) avec autre(s) membre(s) de la famille
	pourcentage					
Canada	1,2	0,5	0,7	3,6	3,5	3,9
Terre-Neuve-et-Labrador	2,5	0,3	0,0	1,6	4,9	3,6
Île-du-Prince-Édouard	1,1	0,8	0,3	3,7	1,8	1,2
Nouvelle-Écosse	1,9	0,0	0,6	1,9	3,0	2,1
Nouveau-Brunswick	3,4	0,4	0,4	2,8	4,5	2,7
Québec	1,7	0,6	0,5	4,1	4,7	3,9
Ontario	0,6	0,2	0,9	4,1	3,0	4,9
Manitoba	1,7	1,1	0,9	3,2	3,5	2,4
Saskatchewan	0,6	1,5	1,1	2,4	1,5	3,7
Alberta	2,1	0,5	0,5	3,3	3,8	3,9
Colombie-Britannique	0,5	0,5	0,9	3,0	2,8	2,5
Yukon	0,8	0,8	0,3	0,8	3,1	1,9
Territoires du Nord-Ouest	0,7	0,6	0,5	0,9	2,7	2,9
Nunavut	1,1	2,7	0,7	1,2	0,8	4,8

Rapport technique sur la couverture

Tableau 8.7.2.1

Répartition du surdénombrement du Recensement de 2021 selon le scénario pour chaque province et territoire

Province ou territoire	Scénario de surdénombrement				
	4.4	5.1	6.1	7.1	8.1
	Adulte (de 25 ans ou plus) avec autre(s) adulte(s) non apparenté(s)	Un des ménages n'est pas une résidence privée	Surdénombrement intraménage (même FRAME_ID)	Autre	Couplage déterministe
	pourcentage				
Canada	1,6	2,4	0,6	3,2	48,2
Terre-Neuve-et-Labrador	1,0	3,0	0,3	3,3	49,8
Île-du-Prince-Édouard	1,3	2,5	0,0	0,8	54,0
Nouvelle-Écosse	1,2	1,0	0,4	1,8	47,9
Nouveau-Brunswick	0,4	2,2	0,0	3,1	50,0
Québec	0,9	2,3	0,7	2,9	41,9
Ontario	1,6	2,3	0,7	3,0	49,2
Manitoba	1,8	4,1	0,3	2,5	51,2
Saskatchewan	1,7	3,1	0,2	2,8	54,6
Alberta	2,5	2,7	0,5	4,4	47,0
Colombie-Britannique	2,1	2,3	0,4	4,0	53,6
Yukon	1,4	1,4	0,1	3,0	59,5
Territoires du Nord-Ouest	2,7	1,8	0,5	1,7	61,4
Nunavut	1,2	3,2	0,4	2,3	51,7

FRAME_ID = identificateur unique de ménage

Note : Le surdénombrement selon le scénario est estimé au niveau de la paire plutôt qu'au niveau du groupe. C'est pourquoi il y a une légère différence entre les pourcentages lorsqu'on les compare à ceux du [tableau 8.7.1.1](#).

Source : Statistique Canada, Étude sur le surdénombrement du Recensement de 2021.

9. Estimation

Les sections [6.2](#), [7.4](#) et [8.6](#) portent respectivement sur l'estimation pour l'Enquête sur la classification des logements, l'Étude sur le sous-dénombrement du recensement (ESoR) et l'Étude sur le surdénombrement du recensement (ESuR). La présente section décrit la façon dont les résultats des études sur la couverture du recensement sont combinés pour obtenir une estimation du sous-dénombrement de la population (S), du surdénombrement de la population (R) et du sous-dénombrement net de la population (N) dans différents domaines. L'incidence des erreurs d'échantillonnage sur la qualité des estimations est également mesurée au moyen d'une estimation des erreurs-types pour chaque estimation. Les résultats de l'ESoR et les données du recensement sont utilisés pour établir une estimation du sous-dénombrement, alors que les résultats de l'ESuR fournissent une estimation du surdénombrement. Le sous-dénombrement net est la différence entre le sous-dénombrement et le surdénombrement. La présente section détaille le calcul de ces estimations et des erreurs-types connexes.

Les définitions suivantes sont utilisées :

C	=	chiffre publié du recensement du nombre de personnes dans la population cible
\hat{S}	=	estimation du sous-dénombrement
	=	estimation du nombre de personnes non incluses en C , mais qui auraient dû l'être
\hat{R}	=	estimation du surdénombrement
	=	estimation du nombre de dénombremments inclus en C , mais qui n'auraient pas dû l'être
\hat{N}	=	estimation du sous-dénombrement net
	=	estimation du nombre de personnes non incluses dans C , mais qui auraient dû l'être, moins le nombre de dénombremments inclus dans C , alors qu'ils n'auraient pas dû être
	=	$\hat{S} - \hat{R}$
\hat{P}	=	estimation du nombre de personnes faisant partie de la population cible du recensement calculé à partir des dénombremments du recensement et de l'estimation du sous-dénombrement net de la population
	=	$C + \hat{N}$
\hat{T}_S	=	estimation du taux de sous-dénombrement
	=	$100 * \frac{\hat{S}}{\hat{P}} = 100 * \frac{\hat{S}}{C + \hat{N}}$
\hat{T}_R	=	estimation du taux de surdénombrement
	=	$100 * \frac{\hat{R}}{\hat{P}} = 100 * \frac{\hat{R}}{C + \hat{N}}$
\hat{T}_N	=	estimation du taux de sous-dénombrement net
	=	$100 * \frac{\hat{N}}{\hat{P}} = 100 * \frac{\hat{S} - \hat{R}}{C + \hat{N}}$

\hat{S} découle des résultats de l'ESoR et des données du recensement et \hat{R} est obtenu de l'ESuR, comme suit :

Tableau 9.1
Composantes de l'estimation des erreurs de couverture de la population pour le Canada

Composante	Nombre de personnes
Estimation de S	1 897 876
Estimation de R	755 635
Estimation de N	1 142 241
C	36 991 981
C + estimation de N	38 134 222

S = sous-dénombrement

R = surdénombrement

N = sous-dénombrement net

C = chiffre publié du recensement

Sources : Statistique Canada, études sur la couverture du Recensement de 2021 et Recensement de 2021.

Les erreurs-types estimées sont définies comme suit.

Par définition, nous avons $v(\hat{S}) = v(\hat{M})$ (voir la [section 7.4.6](#)).

$v(\hat{M})$ = estimation de la variance de \hat{M} selon le plan de l'ESoR

$v(\hat{R})$ = estimation de la variance de \hat{R} selon le plan de l'ESuR

Alors,

$$et(\hat{S}) = \sqrt{v(\hat{M})}$$

$$et(\hat{T}_S) = \sqrt{\left(\frac{\hat{S}^2 + \hat{P}^2 - 2\hat{S}\hat{P}}{\hat{P}^4}\right)v(\hat{M}) + \frac{\hat{S}^2}{\hat{P}^4}v(\hat{R})}$$

$$et(\hat{R}) = \sqrt{v(\hat{R})}$$

$$et(\hat{T}_R) = \sqrt{\left(\frac{\hat{R}^2}{\hat{P}^4}\right)v(\hat{M}) + \left(\frac{\hat{R}^2 + \hat{P}^2 - 2\hat{R}\hat{P}}{\hat{P}^4}\right)v(\hat{R})}$$

$$et(\hat{N}) = \sqrt{v(\hat{M}) + v(\hat{R})}$$

$$et(\hat{T}_N) = \sqrt{\left(\frac{(\hat{S} - \hat{R})^2 + \hat{P}^2 - 2(\hat{S} - \hat{R})\hat{P}^2}{\hat{P}^4}\right)[v(\hat{M}) + v(\hat{R})]}$$

10. Évaluation des études sur la couverture

10.1 Étude sur le sous-dénombrement du recensement

10.1.1 Introduction

Les résultats de la plus vaste étude sur la couverture, soit l'Étude sur le sous-dénombrement du recensement (ESoR), peuvent être évalués en comparant ses estimations avec des données portant sur les mêmes caractéristiques, mais tirées d'autres sources, comme la base de données du Recensement de 2021 et des données administratives utilisées par le Programme des estimations démographiques (PED). Les comparaisons avec les estimations de l'ESoR servent à évaluer ces dernières et à quantifier les écarts en matière de concepts et de mesures.

Malgré quelques différences conceptuelles entre l'ESoR et le Recensement de 2021, les estimations du nombre de personnes dénombrées au Recensement de 2021 par l'ESoR peuvent être comparées aux chiffres du recensement. Afin de rendre ces chiffres comparables, certains ajustements ont d'abord été apportés aux chiffres du recensement.

Les estimations des composantes de l'accroissement démographique intercensitaire peuvent être comparées aux estimations provenant d'autres sources. Les estimations de l'ESoR en ce qui concerne le nombre de personnes décédées entre le Recensement de 2016 et celui de 2021 peuvent être comparées aux chiffres tirés des statistiques de l'état civil. Les chiffres du solde de migration interprovinciale calculés par le PED à partir des données provenant de l'ARC peuvent être comparés aux estimations de l'ESoR. Finalement, les estimations des composantes de l'accroissement démographique de l'ESoR peuvent être comparées aux estimations similaires des données administratives.

10.1.2 Comparaisons avec les chiffres du recensement

Puisque le plan d'échantillonnage stratifié à un seul degré de l'ESoR produit des estimations sans biais, les différences entre les estimations de l'ESoR et les chiffres du recensement sont principalement attribuables à une erreur d'échantillonnage dans les estimations de l'ESoR, aux écarts conceptuels entre les deux sources ou à des biais systématiques influençant les deux sources, ce qui entraîne une sous-estimation ou une surestimation des caractéristiques étudiées.

Personnes dénombrées

Les comparaisons établies pour les provinces et l'ensemble du pays sont présentées au [tableau 10.1.2.1](#), de même que l'erreur-type de l'estimation de l'ESoR et la valeur t servant à tester l'hypothèse selon laquelle il n'existe aucun écart entre l'estimation de l'ESoR et les chiffres comparables du recensement. Les ajustements suivants sont apportés aux chiffres du recensement publiés, afin de tenir compte des différences conceptuelles entre les deux sources.

- Les ajustements apportés à partir de l'imputation des ménages complets ont été exclus, car, même s'ils ont été inclus dans les chiffres du recensement, ils ne faisaient pas partie de l'estimation de l'ESoR du nombre de personnes dénombrées.
- L'estimation du surdénombrement du Recensement de 2021 a été soustraite, car la base de données du recensement contenait des personnes surdénombrées, alors que l'estimation de l'ESoR était établie d'après le nombre de personnes uniques ayant été dénombrées (et non d'après le nombre de dénombremments).
- L'estimation obtenue du nombre de personnes résidant hors du Canada cinq ans auparavant (à l'exclusion des immigrants intercensitaires et des résidents non permanents) à partir du questionnaire détaillé du Recensement de 2021 a également été soustraite, car les estimations de l'ESoR n'incluaient pas la majorité de ces personnes. Pour les mêmes raisons, on a également soustrait l'estimation du nombre d'enfants de 0 à 4 ans nés hors du Canada, mais ayant la citoyenneté canadienne.

Rapport technique sur la couverture

- De même, pour les provinces, on a soustrait l'estimation du nombre de personnes résidant dans un territoire cinq ans auparavant, car elles n'étaient pas couvertes par les bases de sondage provinciales de l'ESoR.
- Le nombre de personnes provenant des réserves (participantes au Recensement de 2021, mais non participantes au Recensement de 2016) a également été soustrait, car les estimations de l'ESoR n'incluaient pas la majorité de ces personnes.

Tableau 10.1.2.1

Comparaison entre le nombre estimé de personnes dénombrées dans l'Étude sur le sous-dénombrement du recensement et les chiffres du recensement comparables pour le Canada, les provinces et les territoires

Province ou territoire	Personnes dénombrées		Chiffre du recensement comparable	Écart	Valeur t ¹
	ESoR				
	Nombre estimé	Erreur-type			
Canada	34 166 964	40 467	34 380 739	-213 775	-5,28
Terre-Neuve-et-Labrador	478 988	2 432	480 370	-1 382	-0,57
Île-du-Prince-Édouard	142 075	1 033	144 380	-2 305	-2,23
Nouvelle-Écosse	899 614	4 153	907 675	-8 061	-1,94
Nouveau-Brunswick	724 776	3 361	725 234	-458	-0,14
Québec	7 912 911	18 532	7 976 614	-63 703	-3,44
Ontario	13 243 488	29 802	13 306 050	-62 562	-2,10
Manitoba	1 217 061	5 621	1 231 102	-14 041	-2,50
Saskatchewan	1 033 774	5 457	1 031 302	2 472	0,45
Alberta	3 894 128	13 778	3 932 868	-38 740	-2,81
Colombie-Britannique	4 525 565	15 461	4 550 560	-24 995	-1,62
Yukon	34 815	0	34 815	0	...
Territoires du Nord-Ouest	34 360	0	34 360	0	...
Nunavut	25 409	0	25 409	0	...

... n'ayant pas lieu de figurer

ESoR = Étude sur le sous-dénombrement du recensement

1. Une valeur t supérieure à 1,96 ou inférieure à -1,96 signifie que l'écart est significatif au seuil de 95 %.

Sources : Statistique Canada, études sur la couverture du Recensement de 2021 et Recensement de 2021.

À l'échelle nationale, l'estimation de l'ESoR du nombre de personnes dénombrées au Recensement de 2021 était inférieure au chiffre comparable du recensement (-0,62 %). Pour les recensements de 1996 à 2016, l'écart au niveau national entre l'estimation de l'ESoR et les chiffres comparables du recensement se situait entre -0,09 % et 0,12 %. C'est la première fois depuis que cette comparaison est effectuée que la différence est statistiquement significative à l'échelle nationale. À l'échelle provinciale, l'estimation de l'ESoR est inférieure au chiffre comparable du recensement pour toutes les provinces à l'exception de la Saskatchewan; cet écart est statistiquement significatif pour cinq provinces : l'Île-du-Prince-Édouard, le Québec, l'Ontario, le Manitoba et l'Alberta.

Lors des cycles précédents, des différences significatives ont également été observées. Ces écarts ont fait l'objet d'une vérification plus approfondie pour confirmer que la classification de l'ESoR ne comportait pas de biais (y compris la province de résidence le jour du recensement, par exemple). D'autres facteurs pourraient également jouer un rôle important dans les écarts observés. Outre l'erreur d'échantillonnage, des biais dans les ajustements (p. ex. Canadiens de retour) apportés aux chiffres du recensement publiés, afin d'obtenir des chiffres comparables du point de vue conceptuel, pourraient expliquer ces écarts. Le biais de non-réponse de l'ESoR pourrait aussi entrer en ligne de compte, car l'ajustement en fonction de la non-réponse a été choisi afin d'obtenir le meilleur résultat possible pour estimer les personnes omises et non les personnes dénombrées. Toutes les étapes de l'ESoR ont été régulièrement vérifiées et soumises à des contrôles de la qualité.

Du fait des écarts plus significatifs qu'habituellement, une enquête approfondie a été réalisée. Dans les recensements antérieurs, il y a toujours eu des personnes dénombrées dans le recensement qui n'étaient pas couvertes par les bases de sondage de l'ESoR, soit du fait des limites de ces bases de sondage soit qu'elles ne faisaient pas partie de la population cible du recensement. Outre les Canadiens de retour susmentionnés, la première catégorie comprend également les personnes à charge (enfants et conjoints) de résidents non permanents. La deuxième catégorie comprend les personnes décédées avant le jour du recensement, les naissances postcensitaires, les visiteurs étrangers, les immigrants arrivés au Canada après le jour du recensement, ainsi que les résidents non permanents et leurs personnes à charge arrivés après le jour du recensement ou ne possédant pas de permis valide. Pour les cycles antérieurs, le nombre de ces personnes était jugé suffisamment limité pour ne pas avoir d'incidence significative sur la comparaison entre l'estimation des personnes dénombrées de l'ESoR et le chiffre comparable du recensement. Du fait de l'importante hausse du nombre de résidents non permanents ces récentes années, il semble que cela ne soit plus le cas. Une grande partie de l'écart entre ces deux chiffres peut s'expliquer par le dénombrement des personnes à charge de résidents non permanents ou d'enregistrements hors du champ du recensement.

Une analyse des écarts par groupe d'âge a indiqué que les écarts négatifs étaient concentrés parmi les enfants (âgés de 0 à 14 ans; les écarts pour les groupes d'âge de 0 à 4 ans et de 5 à 9 ans étant statistiquement significatifs) et les jeunes âgés de 25 à 29 ans. Les écarts sont minimaux, ou au moins non significatifs, pour les autres groupes. Cela reflète le fait qu'une partie significative des personnes dénombrées non couvertes par les bases de sondage de l'ESoR serait les enfants ou conjoints de résidents non permanents ayant un permis.

Aucun travail d'enquête n'a soulevé d'inquiétude sur les étapes de classification, de pondération, ni d'estimation de l'ESoR.

10.1.3 Comparaison avec les estimations démographiques

Personnes décédées

Le [tableau 10.1.3.1a](#) permet de comparer l'estimation du nombre de personnes décédées pendant la période intercensitaire (entre le 10 mai 2016 et le 10 mai 2021) selon la province de classification de l'ESoR aux chiffres tirés des dossiers des statistiques de l'état civil. L'estimation de l'ESoR exclut les personnes décédées hors Canada lorsque le pays de décès est connu. À l'échelle nationale, l'estimation de l'ESoR dépasse de 6 492 personnes (0,5 %) les chiffres des statistiques de l'état civil; cette différence n'était pas statistiquement significative. À l'échelle provinciale, l'écart en pourcentage le plus important a été enregistré pour l'Île-du-Prince-Édouard (648 ou 9,7 %) et le Québec (12 731 ou 3,7 %); seul le premier était statistiquement significatif (valeur de t de 2,28). Dans les autres provinces, les différences relatives se situaient entre -1,7 % et 2,1 %. Elles n'étaient pas statistiquement significatives et la plupart de ces écarts étaient inférieurs à ceux observés en 2016.

Tableau 10.1.3.1a

Comparaison entre le nombre estimé de personnes décédées dans l'Étude sur le sous-dénombrement du recensement et le chiffre des statistiques de l'état civil pour les provinces

Province	Personnes décédées du 10 mai 2016 au 10 mai 2021		Chiffre des statistiques de l'état civil	Écart	Valeur t^1
	ESoR				
	Nombre estimé	Erreur- type			
Total	1 442 017	13 967	1 435 525	6 492	0,46
Terre-Neuve-et-Labrador	25 999	626	26 255	-256	-0,41
Île-du-Prince-Édouard	7 317	285	6 669	648	2,28
Nouvelle-Écosse	47 877	933	48 349	-472	-0,51
Nouveau-Brunswick	38 084	918	37 748	336	0,37
Québec	355 031	7 366	342 300	12 731	1,73

Tableau 10.1.3.1a

Comparaison entre le nombre estimé de personnes décédées dans l'Étude sur le sous-dénombrement du recensement et le chiffre des statistiques de l'état civil pour les provinces

Province	Personnes décédées du 10 mai 2016 au 10 mai 2021		Chiffre des statistiques de l'état civil	Écart	Valeur t ¹
	ESoR				
	Nombre estimé	Erreur- type			
Ontario	531 134	9 193	540 387	-9 253	-1,01
Manitoba	55 461	1 656	56 058	-597	-0,36
Saskatchewan	48 750	1 419	48 835	-85	-0,06
Alberta	132 746	3 842	133 471	-725	-0,19
Colombie-Britannique	199 619	5 001	195 453	4 166	0,83

ESoR = Étude sur le sous-dénombrement du recensement

1. Une valeur t supérieure à 1,96 ou inférieure à -1,96 signifie que l'écart est significatif au seuil de 95 %.

Note : En raison de l'arrondissement, la somme des estimations de la couverture pourrait ne pas correspondre au total.

Sources : Statistique Canada, Étude sur le sous-dénombrement du Recensement de 2021, Programme de la statistique de l'état civil et Programme des estimations démographiques.

Certaines raisons peuvent expliquer un écart significatif. Tout d'abord, l'estimation de l'ESoR peut inclure certains décès survenant à l'étranger; ce que n'incluent pas les statistiques de l'état civil. Dans l'ESoR, si le pays de décès est connu et qu'il se trouve à l'étranger, le décès ne fait pas partie de la comparaison des personnes décédées du [tableau 10.1.3.1a](#). Cependant, si la personne n'a pas été trouvée dans les dossiers des statistiques de l'état civil et que le pays de décès est inconnu, le décès sera alors classé par défaut dans la plus récente province de résidence de la personne au Canada. Seules 12 personnes choisies présentaient cette situation; aucune ne concernait l'Île-du-Prince-Édouard. Par conséquent, il ne s'agit pas de la principale raison de cet écart significatif pour cette province. Une autre raison pouvant expliquer l'écart est de ne pas pouvoir trouver la personne dans les statistiques de l'état civil du fait de différences dans les renseignements personnels. La dernière raison est une sous-déclaration dans les statistiques de l'état civil. Le décès de 34 personnes choisies a été confirmé comme ayant eu lieu au Canada sans figurer dans les statistiques de l'état civil; parmi ces cas, 10 concernaient l'Île-du-Prince-Édouard. Le [tableau 10.1.3.1b](#) permet de comparer l'estimation de l'ESoR du nombre de personnes décédées pendant la période intercensitaire (entre le 10 mai 2016 et le 10 mai 2021), selon la province de décès indiquée dans les dossiers des statistiques de l'état civil (donc seulement pour les personnes trouvées dans ces dossiers), aux chiffres tirés des statistiques de l'état civil. L'écart pour l'Île-du-Prince-Édouard qui était significatif ne l'était plus, avec une valeur t de 0,45. Pourtant, l'écart devient significatif en Ontario (-19 386, avec une valeur t de -2,47). Même si ces derniers résultats ne semblent pas démontrer d'enjeux liés aux estimations de l'ESoR du nombre de personnes décédées, des mesures de vérification plus approfondies ont été effectuées, afin de confirmer qu'aucune erreur de classification ou d'autre nature n'était en cause dans les opérations ou dans les estimations. Aucun problème à cet effet n'a été détecté.

Tableau 10.1.3.1b

Comparaison entre le nombre estimé de personnes décédées dans l'Étude sur le sous-dénombrement du recensement et liées aux statistiques de l'état civil et le chiffre des statistiques de l'état civil pour les provinces

Province	Personnes décédées du 10 mai 2016 au 10 mai 2021		Chiffre des statistiques de l'état civil	Écart	Valeur t ¹
	ESoR				
	Nombre estimé	Erreur- type			
Total	1 419 908	12 750	1 435 525	-15 617	-1,22
Terre-Neuve-et-Labrador	25 966	626	26 255	-289	-0,46
Île-du-Prince-Édouard	6 773	232	6 669	104	0,45

Tableau 10.1.3.1b

Comparaison entre le nombre estimé de personnes décédées dans l'Étude sur le sous-dénombrement du recensement et liées aux statistiques de l'état civil et le chiffre des statistiques de l'état civil pour les provinces

Province	Personnes décédées du 10 mai 2016 au 10 mai 2021		Chiffre des statistiques de l'état civil	Écart	Valeur t ¹
	ESoR				
	Nombre estimé	Erreur- type			
Nouvelle-Écosse	47 877	933	48 349	-472	-0,51
Nouveau-Brunswick	37 123	887	37 748	-625	-0,70
Québec	347 817	7 081	342 300	5 517	0,78
Ontario	521 001	7 859	540 387	-19 386	-2,47
Manitoba	54 605	1 628	56 058	-1 453	-0,89
Saskatchewan	48 750	1 419	48 835	-85	-0,06
Alberta	131 732	3 823	133 471	-1 739	-0,45
Colombie-Britannique	198 263	4 869	195 453	2 810	0,58

ESoR = Étude sur le sous-dénombrement du recensement

1. Une valeur t supérieure à 1,96 ou inférieure à -1,96 signifie que l'écart est significatif au seuil de 95 %.

Note : En raison de l'arrondissement, la somme des estimations de la couverture pourrait ne pas correspondre au total.

Sources : Statistique Canada, Étude sur le sous-dénombrement du Recensement de 2021, Programme de la statistique de l'état civil et Programme des estimations démographiques.

Migration interprovinciale

Le [tableau 10.1.3.2](#) compare les estimations de l'ESoR du solde de migration interprovinciale intercensitaire avec les chiffres correspondants calculés par le PED selon des fichiers de l'Agence du revenu du Canada. En général, les données sur les migrants interprovinciaux ne sont pas comparables, car l'ESoR ne prend en compte que les migrations survenues entre la date de référence de la base de sondage (p. ex. le 10 mai 2016 pour la base du recensement) et le jour du Recensement de 2021, alors que les estimations du PED prennent en compte les migrations annuelles. Pour cette raison, seules les estimations du solde de migration interprovinciale sont présentées.

Même si ces estimations diffèrent, les estimations démographiques et de l'ESoR du solde de migration interprovinciale sont de même direction (solde migratoire positif ou négatif) pour chaque province. L'écart entre les estimations démographiques et de l'ESoR est inférieur à celui de 2016 pour 6 des 10 provinces et sensiblement identique pour une autre; toutefois, les erreurs-types des estimations de l'ESoR sont également bien inférieures qu'elles ne l'étaient en 2016. Par conséquent, deux provinces enregistrent un écart statistiquement significatif entre les estimations démographiques et de l'ESoR du solde migratoire : l'Ontario (valeur t de 3,27) et le Manitoba (-2,27). Un écart significatif avait été observé pour une province en 2016.

Tableau 10.1.3.2

Comparaison entre les estimations du solde de migration interprovinciale de l'Étude sur le sous-dénombrement du recensement et les estimations démographiques (définition de l'Agence du revenu du Canada) pour les provinces

Province	Solde de migration interprovinciale			Estimation démographique	Écart	Valeur t ²
	ESoR ¹					
	Taille de l'échantillon	Nombre estimé	Erreur- type			
Terre-Neuve-et-Labrador	27 245	-6 477	2 405	-8 555	2 078	0,86
Île-du-Prince-Édouard	13 650	1 616	1 093	2 993	-1 377	-1,26
Nouvelle-Écosse	69 283	14 315	4 119	21 720	-7 406	-1,80

Tableau 10.1.3.2

Comparaison entre les estimations du solde de migration interprovinciale de l'Étude sur le sous-dénombrement du recensement et les estimations démographiques (définition de l'Agence du revenu du Canada) pour les provinces

Province	Solde de migration interprovinciale					
	ESoR ¹			Estimation démographique	Écart	Valeur t ²
	Taille de l'échantillon	Nombre estimé	Erreur-type			
Nouveau-Brunswick	50 933	6 938	2 918	7 491	-553	-0,19
Québec	111 790	-39 080	9 831	-26 852	-12 227	-1,24
Ontario	291 826	73 340	15 334	23 157	50 183	3,27
Manitoba	57 416	-41 156	4 598	-30 715	-10 441	-2,27
Saskatchewan	65 324	-42 617	5 066	-42 019	-599	-0,12
Alberta	232 028	-58 339	13 084	-35 046	-23 293	-1,78
Colombie-Britannique	211 119	91 460	13 209	87 826	3 634	0,28

ESoR = Étude sur le sous-dénombrement du recensement

1. L'Étude sur le sous-dénombrement du recensement (ESoR) exclut les personnes habitant dans une province le 11 mai 2021 qui vivaient dans l'un des trois territoires cinq ans auparavant, soit le 10 mai 2016.

2. Une valeur t supérieure à 1,96 ou inférieure à -1,96 signifie que l'écart est significatif au seuil de 95 %.

Sources : Statistique Canada, Étude sur le sous-dénombrement du Recensement de 2021 et Programme des estimations démographiques.

10.2 Étude sur le surdénombrement du recensement

La validation des résultats des deux études sur la couverture reposait sur la Directive sur la validation des produits statistiques de Statistique Canada. Parmi diverses étapes de validation, il a été possible d'évaluer les résultats de l'Étude sur le surdénombrement du recensement (ESuR) de 2021 en évaluant la façon dont chacune des composantes menant à la construction de sa base de sondage contribuait à l'estimation globale du surdénombrement du recensement. Il a également été possible d'examiner les raisons possibles pour lesquelles des personnes ont été dénombrées plusieurs fois dans le recensement. Reportez-vous à la [section 8.7](#) pour obtenir de plus amples détails.

Au cours du processus de validation, une autre mesure a été de comparer la base de sondage et les estimations de l'ESuR avec d'autres sources disponibles, pour analyser ce qui pouvait manquer dans la base de sondage de l'ESuR et relever tout problème systémique, le cas échéant. Une telle source est l'Environnement de couplage de données sociales (ECDS). L'équipe de l'ECDS de Statistique Canada est chargée de tous les couplages du recensement utilisés pour dériver des renseignements de sources de données administratives (p. ex. afin d'obtenir des renseignements sur le revenu). Comme première étape de ces couplages, l'équipe de l'ECDS effectue un couplage probabiliste interne de la Base de données des réponses du recensement (BDR) entière avec elle-même pour relever des doublons potentiels de personnes. Des différences importantes existent entre l'objectif du couplage interne des enregistrements de la BDR effectué par l'ECDS et celui de l'ESuR. L'objectif de l'ECDS est de relever les enregistrements en double avec certitude et une approche conservatrice est donc employée dans la détermination des doublons. L'objectif de l'ESuR est de construire une base de sondage de toutes les paires de doublons possibles du couplage interne de la BDR.

La liste de l'ECDS des doublons de personnes a été comparée à la base de sondage de l'ESuR; tous les doublons de l'ECDS se trouvaient dans la base de sondage de l'ESuR de 2021. De plus, une comparaison des estimations de l'ESuR et des doublons potentiels de personnes tirés de l'ECDS a été effectuée, afin d'étudier les tendances entre 2016 et 2021. Même si les résultats de l'ESuR sont une estimation du surdénombrement et qu'ils comprennent l'erreur d'échantillonnage, alors que les résultats de l'ECDS sont des dénombrements et devraient représenter un sous-ensemble de l'estimation de l'ESuR de par leur nature, comparer la variation en pourcentage entre 2016 et 2021 pour les deux sources était tout de même un exercice utile. Les deux sources ont présenté des résultats cohérents : une hausse du surdénombrement entre 2016 et 2021 et une tendance similaire observée pour l'ensemble des provinces et des territoires.

Une autre évaluation a été effectuée en examinant la corrélation entre l'état de surdénombrement d'une paire d'enregistrements de la BDR potentiellement en double et le poids de couplage dérivé de cette paire, pour les paires relevées par le couplage probabiliste de la BDR avec elle-même (reportez-vous à la [section 8.2.3](#) pour obtenir de plus amples détails sur cette étape du couplage probabiliste). Dans le cadre d'un couplage probabiliste, un poids de couplage est calculé pour chaque paire d'enregistrements couplés, en fonction de la robustesse de ce lien. Comme le décrit la [section 8](#), dans l'ESuR, un échantillon des paires relevées comme doublons potentiels par le couplage probabiliste de la BDR à elle-même a fait l'objet d'un processus de vérification manuelle, selon lequel les codeurs devaient déterminer si chaque paire était un véritable doublon (surdénombrement vérifié) ou non. On a comparé la répartition des poids de couplage pour les cas de surdénombrement et de non-surdénombrement. Si le couplage s'est déroulé comme prévu, il devrait exister une différence entre les poids des deux groupes. Comme prévu, les poids de couplage étaient bien supérieurs en moyenne pour les cas de surdénombrement; cela s'observait pour chaque province et territoire.

De plus, les cas supplémentaires de surdénombrement relevés au cours des opérations de vérification manuelle, mais ne faisant pas partie de la base de sondage finale de l'ESuR, ont été évalués pour comprendre pourquoi ils n'étaient pas capturés. Il ne semble pas exister de raison systémique expliquant pourquoi les paires supplémentaires du surdénombrement ne figuraient pas dans la base de l'ESuR. En général, ces paires étaient trop différentes; ce qui signifie qu'elles présentaient de multiples coquilles, erreurs ou trop de différences dans les champs utilisés pendant les processus de couplage, entraînant leur absence de la base de sondage de l'ESuR. Comme pour chaque cycle, les critères de sélection initiaux et règles de couplage seront examinés, révisés et testés avant le prochain cycle.

Les activités de validation de l'ESuR de 2021 n'ont pas soulevé de préoccupations quant à la méthodologie de l'étude et l'évaluation des estimations de l'ESuR a montré que les résultats étaient cohérents avec les résultats passés et avec les attentes.

10.3 Estimations démographiques

10.3.1 Erreur en fin de période

Le PED de Statistique Canada établit la population des provinces et des territoires au jour du recensement en additionnant les chiffres de la population du recensement, les estimations du sous-dénombrement net du recensement (SDNR) et l'estimation des populations des réserves et établissements partiellement dénombrés. Le PED reconduit ensuite au 1^{er} juillet 2021 ces chiffres rajustés qui deviennent le point de départ des estimations postcensitaires de la population.

Au moment de l'établissement de ces chiffres ajustés, le PED procède à une évaluation de la qualité des estimations postcensitaires qu'il a produites dans la période quinquennale précédant le recensement. L'évaluation porte sur l'écart entre le chiffre des estimations postcensitaires au jour du recensement et le chiffre ajusté de la population de ce recensement. Cet écart est appelé erreur en fin de période. L'examen détaillé de cette erreur constitue la principale mesure de la qualité des estimations postcensitaires.

Le [tableau 10.3.1](#) présente les erreurs en fin de période pour 2006, 2011, 2016 et 2021, selon la province et le territoire ainsi que pour l'ensemble du Canada. Rappelons qu'une erreur en fin de période est dite positive lorsque l'estimation démographique postcensitaire est supérieure au chiffre du recensement ajusté. À l'échelle nationale, l'erreur en fin de période pour 2021 s'élève à -41 269 personnes, pour un taux d'erreur de -0,11 %. Les estimations démographiques à l'échelle nationale ont donc sous-estimé la population du Canada. Le taux d'erreur de 2021 était inférieur à celui de 2006 et 2016¹³. Quatre provinces et un territoire enregistraient une erreur en fin de période supérieure à 1 % ou inférieure à -1 % en 2021 : Terre-Neuve-et-Labrador (-1,24 %), Île-du-Prince-Édouard (1,59 %), Nouvelle-Écosse (-1,00 %), Saskatchewan (1,06 %) et Territoires du Nord-Ouest (2,57 %). Aux fins de comparaison, en 2016, cinq provinces et un territoire enregistraient des erreurs en fin de période similaires. En 2021, huit provinces et un territoire enregistraient une erreur en fin de période inférieure (en valeur absolue) à celle de 2016.

Tableau 10.3.1
Erreur en fin de période pour le Canada, les provinces et les territoires, 2006, 2011, 2016 et 2021

Province ou territoire	2006		2011		2016		2021	
	nombre	taux (%)	nombre	taux (%)	nombre	taux (%)	nombre	taux (%)
Canada	39 409	0,12	158 558	0,46	120 044	0,33	-41 269	-0,11
Terre-Neuve-et-Labrador	-1 821	-0,36	-11 121	-2,12	1 097	0,21	-6 540	-1,24
Île-du-Prince-Édouard	-31	-0,02	2 096	1,46	2 906	1,99	2 564	1,59
Nouvelle-Écosse	-3 997	-0,43	5 075	0,54	7 395	0,79	-9 944	-1,00
Nouveau-Brunswick	2 673	0,36	1 432	0,19	-5 992	-0,79	-317	-0,04
Québec	19 776	0,26	-23 207	-0,29	89 035	1,08	33 890	0,40
Ontario	24 532	0,19	121 217	0,92	68 329	0,49	-43 978	-0,30
Manitoba	-5 977	-0,51	21 464	1,74	5 358	0,41	-3 084	-0,22
Saskatchewan	-3 691	-0,37	-7 779	-0,73	12 492	1,10	12 402	1,06
Alberta	-50 869	-1,49	-3 345	-0,09	43 891	1,05	2 013	0,05
Colombie-Britannique	61 120	1,44	52 325	1,16	-104 201	-2,15	-29 372	-0,56
Yukon	-1 027	-3,19	103	0,29	-391	-1,02	150	0,35
Territoires du Nord-Ouest	-857	-1,99	758	1,74	-47	-0,11	1 146	2,57
Nunavut	-422	-1,37	-460	-1,35	172	0,47	-199	-0,50

Source : Statistique Canada, Centre de démographie.

10.3.2 Exactitude des estimations postcensitaires

Aux fins de la production des estimations du PED, les études sur la couverture du recensement sont utilisées pour ajuster les chiffres du recensement pour tenir compte du SDNR. Or, comme ces études se basent en partie sur des enquêtes par sondage, les résultats du SDNR comportent une certaine variabilité statistique attribuable à l'échantillonnage. Afin de déterminer si les erreurs en fin de période présentées précédemment sont statistiquement significatives, il faut donc tenir compte de l'erreur-type des chiffres du recensement ajustés. Par ailleurs, étant donné que les chiffres ajustés du Recensement de 2016 ont servi de base aux estimations postcensitaires de 2016 à 2021, une erreur-type combinant la variabilité statistique des chiffres ajustés des recensements de 2016 et de 2021 a été calculée pour le Canada et pour chaque province et territoire.

13. Les erreurs en fin de période de 2006 à 2016 sont calculées à l'aide des estimations postcensitaires des recensements de 2006, 2011 et 2016, modifiées en 2023 à la suite d'une révision des composantes.

Rapport technique sur la couverture

Le tableau 10.3.2 présente l'erreur en fin de période de 2021 pour le Canada, les provinces et les territoires, l'erreur-type combinée des chiffres ajustés des recensements de 2016 et de 2021 et la valeur t^{14} . L'erreur en fin de période est statistiquement significative au niveau de confiance de 95 % pour Terre-Neuve-et-Labrador, l'Île-du-Prince-Édouard, la Nouvelle-Écosse, la Saskatchewan et les Territoires du Nord-Ouest. Pour ces secteurs de compétence, la variabilité attribuable à l'échantillonnage des chiffres ajustés des recensements de 2016 et de 2021 ne permet donc pas d'expliquer la plus grande part de l'erreur en fin de période.

Tableau 10.3.2
Incidence de la variabilité statistique des recensements rajustés sur l'exactitude des estimations de la population, Canada, provinces et territoires, 2016 à 2021

Province ou territoire	Erreur en fin de période	Erreur-type combinée des recensements rajustés de 2016 et de 2021	Valeur t^1
		nombre	
Canada	-41 269	49 096	-0,84
Terre-Neuve-et-Labrador	-6 540	2 456	-2,66
Île-du-Prince-Édouard	2 564	1 238	2,07
Nouvelle-Écosse	-9 944	4 372	-2,27
Nouveau-Brunswick	-317	3 542	-0,09
Québec	33 890	23 487	1,44
Ontario	-43 978	38 965	-1,13
Manitoba	-3 084	6 420	-0,48
Saskatchewan	12 402	5 646	2,20
Alberta	2 013	16 334	0,12
Colombie-Britannique	-29 372	20 052	-1,46
Yukon	150	262	0,57
Territoires du Nord-Ouest	1 146	320	3,58
Nunavut	-199	331	-0,60

1. Une valeur t supérieure à 1,96 ou inférieure à -1,96 signifie que l'écart est significatif au seuil de 95 %.

Source : Statistique Canada, Division de la démographie.

Les composantes de l'accroissement démographique estimées par le PED ont été comparées à celles tirées d'autres sources, principalement de l'ESoR, afin de déterminer les composantes pouvant être plus étroitement liées à l'erreur en fin de période. Cette analyse s'est concentrée sur les cinq secteurs de compétence pour lesquels cette erreur était statistiquement significative. La migration interprovinciale, notamment celle des immigrants récents, pourrait expliquer une partie de l'erreur en fin de période calculée pour l'Île-du-Prince-Édouard, la Saskatchewan et les Territoires du Nord-Ouest. Les effets de l'échantillonnage de l'ESoR et de plusieurs composantes de l'accroissement démographique pourraient contribuer à expliquer l'erreur calculée pour Terre-Neuve-et-Labrador, ainsi que pour la Nouvelle-Écosse. Toutefois, il est difficile de déterminer un facteur principal pour ces deux provinces. Enfin, l'émigration et le nombre de résidents non permanents demeurent généralement les phénomènes démographiques particulièrement difficiles à mesurer.

14. Si la valeur t est supérieure à 1,96 ou inférieure à -1,96, alors l'estimation du PED est statistiquement différente du chiffre du recensement ajusté à un niveau de confiance de 95 %.

11. Estimations historiques des erreurs de couverture de la population

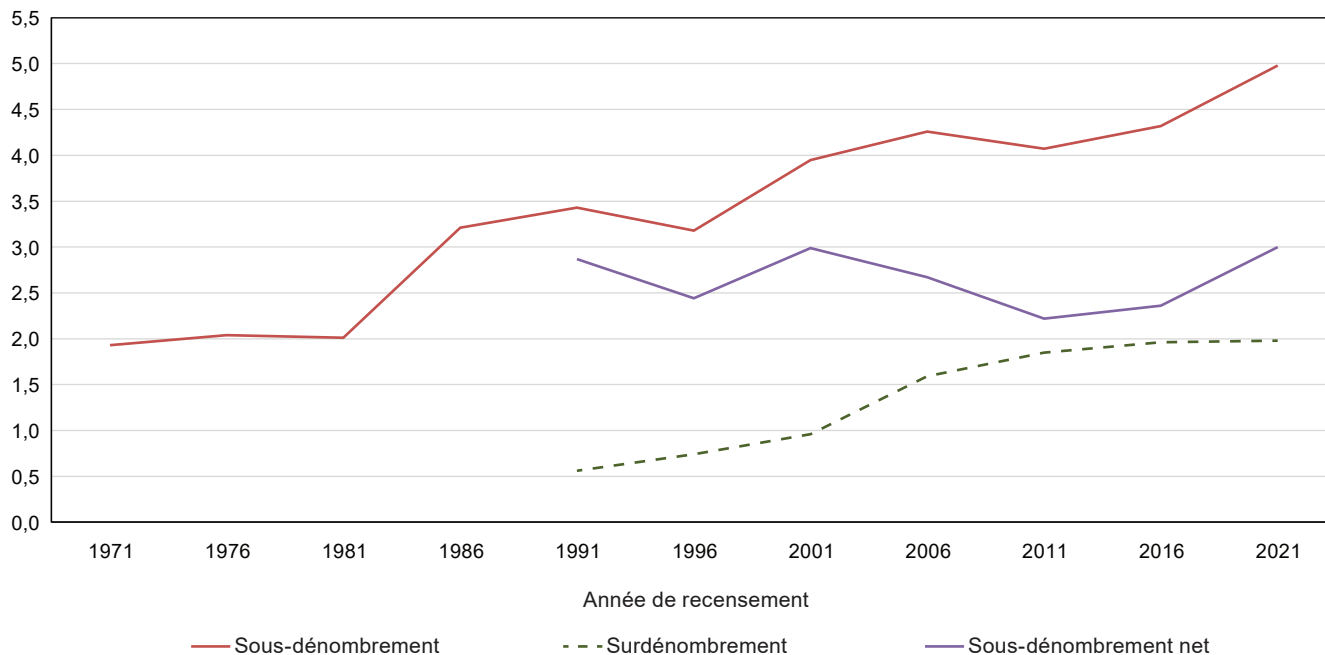
11.1 Estimations

La présente section propose des estimations historiques des erreurs de couverture de la population. Le [graphique 11.1a](#) présente le taux de sous-dénombrement \hat{T}_S de la population estimée du Recensement de 1971 à celui de 2021, ainsi que le taux de surdénombrement \hat{T}_R de la population estimée et le taux de sous-dénombrement net \hat{T}_N de la population estimée du Recensement de 1991 à celui de 2021. Le surdénombrement et le sous-dénombrement net remontent à 1991, étant donné que le taux estimé de surdénombrement a été établi pour la première fois lors du Recensement de 1991, après une étude expérimentale portant sur le Recensement de 1986.

Graphique 11.1a

Estimations du taux d'erreur de couverture de la population pour le Canada, recensements de 1971 à 2021

taux (pourcentage)



Sources : Statistique Canada, études de couverture des recensements de 1971 à 2021.

L'erreur de couverture de la population est une préoccupation majeure en matière de qualité des données; le sous-dénombrement a augmenté d'un facteur de 2,5 depuis 1981, et le surdénombrement est 2,5 fois supérieur à ce qu'il était en 1996, même s'il est demeuré relativement stable au cours des trois derniers recensements. Les variations du sous-dénombrement net d'un recensement à l'autre reflètent les changements du sous-dénombrement et du surdénombrement, qui reflètent à leur tour des changements touchant la situation démographique, la situation des particuliers dans les ménages, la méthodologie du recensement et les modifications apportées à la méthodologie relative aux études sur la couverture. Cette dernière question est abordée à la [section 11.2](#).

Comme le montre le [graphique 11.1a](#), le taux de sous-dénombrement a accusé une hausse marquée en 2021, l'une des augmentations les plus importantes d'un recensement à un autre, alors que le taux de surdénombrement a très peu varié. Le taux de sous-dénombrement a peu varié lors des recensements de 1971 à 1981, avant d'augmenter de 1,2 point de pourcentage lors du Recensement de 1986. Il s'est ensuite relativement maintenu de 1986 à 1996, avant d'augmenter de nouveau de 0,78 point de pourcentage en 2001. Il est redevenu relativement stable lors des recensements de 2001 à 2016, alternant entre de faibles augmentations et diminutions, puis a

enregistré à nouveau une hausse, de 0,66 point de pourcentage en 2021. Comme cela a été mentionné, une partie de ces fluctuations peut être expliquée par les améliorations apportées à la méthodologie des études sur la couverture décrite à la [section 11.2](#).

Le taux de surdénombrement a augmenté à chaque recensement depuis sa première mesure. L'augmentation la plus importante enregistrée du taux de surdénombrement a été lorsqu'il est passé de 0,96 % à 1,59 % entre 2001 et 2006. En 2021, ce taux a atteint un sommet à 1,98 %, même si la faible hausse de 0,02 point de pourcentage après 2016 n'était pas statistiquement significative.

En 2021, le sous-dénombrement net a augmenté de manière significative (hausse pour deux recensements de suite), suivant des reculs lors des deux recensements précédents. Il s'agit du plus important taux de sous-dénombrement net observé depuis sa première mesure, même s'il était approximativement équivalent au taux observé en 2001. La hausse par rapport à 2016 était statistiquement significative.

Si l'on examine le sous-dénombrement depuis le Recensement de 1981, l'augmentation constatée lors du Recensement de 1986 a donné lieu à la création du Registre des adresses (RA) dans le cadre du Recensement de 1991. Le RA constituait une liste distincte des logements urbains qui auraient dû être dénombrés. Lors du Recensement de 1996, l'intervention d'un agent recenseur pour remplacer l'autodénombrement dans certains centres de grandes villes a permis de réduire le taux de sous-dénombrement. Par ailleurs, le fait de devancer le jour du recensement, du début du mois de juin au milieu du mois de mai, a amélioré le contrôle du sous-dénombrement, étant donné que les personnes étaient plus susceptibles d'être chez elles et moins enclines à être sur le point de déménager. En 2006, l'envoi postal des questionnaires dans les zones urbaines a permis de diminuer le nombre d'employés nécessaires à la collecte. L'introduction des réponses par Internet a également permis de diminuer les problèmes de saisie des données. En 2011, l'adoption de la méthodologie des vagues a permis de mieux concentrer les efforts de suivi du recensement, et une forte augmentation des réponses par Internet a une fois de plus permis de réduire les problèmes de saisie des données. La disparition du questionnaire détaillé a probablement contribué à diminuer légèrement le taux de non-réponse au recensement. En 2016, le retour du questionnaire détaillé obligatoire, la tendance à l'augmentation des réponses par Internet et le mouvement d'appui de la population envers le recensement ont permis de diminuer le taux de non-réponse au recensement et d'en améliorer la qualité globale. En 2021, une forte augmentation du nombre de résidents non permanents, population représentant généralement un taux de sous-dénombrement bien plus élevé, a contribué à la hausse du sous-dénombrement. La pandémie a aussi pu jouer un rôle dans ce résultat.

Les tableaux [11.1a](#) et [11.1b](#) présentent les estimations du sous-dénombrement. Il convient de signaler que l'année 1971 n'est pas comprise dans le tableau 11.1b, étant donné que des estimations pour des groupes d'âge différents ont été établies pour les personnes de plus de 24 ans.

Rapport technique sur la couverture

Tableau 11.1a

Estimations du taux de sous-dénombrement de la population et de l'erreur-type pour le Canada, les provinces et les territoires, Recensement de 1971 au Recensement de 2021

Province ou territoire	1971		1976		1981			
	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type		
	pourcentage							
Canada	1,93	0,09	2,04	0,10	2,01	0,09		
Terre-Neuve-et-Labrador	2,25	0,72	1,10	0,39	1,74	0,45		
Île-du-Prince-Édouard	1,23	1,13	0,38	0,25	1,17	0,54		
Nouvelle-Écosse	1,33	0,45	0,86	0,34	1,05	0,34		
Nouveau-Brunswick	1,65	0,56	2,16	0,37	1,81	0,30		
Québec	2,10	0,19	2,95	0,25	1,91	0,21		
Ontario	1,68	0,12	1,52	0,17	1,94	0,14		
Manitoba	1,13	0,38	1,07	0,33	0,98	0,35		
Saskatchewan	1,00	0,37	1,33	0,34	0,99	0,37		
Alberta	2,55	0,44	1,49	0,26	2,54	0,36		
Colombie-Britannique	2,89	0,39	3,13	0,31	3,16	0,33		
Yukon		
Territoires du Nord-Ouest		
Nunavut		
Province ou territoire	1986		1991		1996		2001	
	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type
	pourcentage							
Canada	3,21	0,13	3,43	0,12	3,18	0,09	3,95	0,13
Terre-Neuve-et-Labrador	1,92	0,33	2,47	0,30	2,45	0,29	2,43	0,32
Île-du-Prince-Édouard	2,14	0,80	1,67	0,23	1,76	0,28	1,89	0,53
Nouvelle-Écosse	2,15	0,34	2,25	0,36	2,70	0,27	3,44	0,41
Nouveau-Brunswick	2,71	0,33	3,71	0,42	2,49	0,28	3,57	0,42
Québec	2,91	0,31	3,18	0,20	2,46	0,18	2,93	0,26
Ontario	3,43	0,19	4,23	0,28	3,40	0,18	4,56	0,25
Manitoba	2,94	0,40	2,31	0,36	2,55	0,29	3,49	0,43
Saskatchewan	2,38	0,37	2,15	0,32	3,30	0,32	3,18	0,37
Alberta	3,00	0,32	2,51	0,27	2,99	0,24	3,18	0,33
Colombie-Britannique	4,48	0,36	3,42	0,24	4,58	0,24	5,30	0,34
Yukon	4,12	0,58	3,92	0,51	5,59	1,16
Territoires du Nord-Ouest	5,73	0,57	4,28	0,67	9,10	0,80
Nunavut	6,54	0,63	5,07	1,39

Tableau 11.1a

Estimations du taux de sous-dénombrement de la population et de l'erreur-type pour le Canada, les provinces et les territoire, Recensement de 1971 au Recensement de 2021

Province ou territoire	2006		2011		2016		2021	
	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type
	pourcentage							
Canada	4,26	0,17	4,07	0,16	4,32	0,11	4,98	0,09
Terre-Neuve-et-Labrador	2,62	0,54	3,70	0,53	3,94	0,36	5,02	0,32
Île-du-Prince-Édouard	3,04	0,52	3,90	0,62	4,01	0,57	6,32	0,57
Nouvelle-Écosse	4,02	0,54	4,04	0,54	3,70	0,30	4,73	0,32
Nouveau-Brunswick	3,56	0,43	2,64	0,43	4,24	0,34	3,82	0,32
Québec	2,46	0,32	2,99	0,29	2,57	0,21	2,64	0,18
Ontario	5,18	0,34	4,47	0,32	4,63	0,22	5,78	0,18
Manitoba	4,32	0,57	3,11	0,48	3,95	0,35	4,52	0,35
Saskatchewan	3,81	0,50	4,43	0,57	4,99	0,37	5,06	0,36
Alberta	4,74	0,49	5,11	0,45	4,51	0,30	5,24	0,26
Colombie-Britannique	4,83	0,41	4,31	0,41	6,31	0,31	6,50	0,28
Yukon	7,23	0,64	6,30	0,81	8,42	0,45	7,71	0,37
Territoires du Nord-Ouest	5,74	0,57	5,99	0,69	7,83	0,53	9,91	0,46
Nunavut	5,55	0,60	7,39	1,65	4,25	0,59	9,70	0,56

.. indisponible pour une période de référence précise

... n'ayant pas lieu de figurer

Notes : À l'exclusion des réserves partiellement dénombrées. Les chiffres du Recensement de 1991 et ceux des recensements subséquents comprennent les résidents non permanents et les territoires. Les chiffres tiennent compte des modifications apportées à la publication initiale de 1986. Les chiffres ne comprennent pas les estimations des personnes omises dans les logements classés par erreur comme étant inoccupés en 1971 et en 1976.

Sources : Statistique Canada, études sur la couverture des recensements de 1971 à 2021.

Rapport technique sur la couverture

Tableau 11.1b

Estimations du taux de sous-dénombrement de la population et de l'erreur-type, selon le sexe et le groupe d'âge pour le Canada, Recensement de 1976 au Recensement de 2021

Sexe ¹ et groupe d'âge	1976		1981		1986	
	Taux estimé	Erreur- type	Taux estimé	Erreur- type	Taux estimé	Erreur- type
	pourcentage					
Les deux sexes	2,04	0,10	2,01	0,09	3,21	0,13
0 à 4 ans	2,31	0,28	1,21	0,22	2,14	0,49
5 à 14 ans	1,20	0,16	1,23	0,21	2,08	0,26
15 à 17 ans ²	1,99	0,38	2,96	0,52	3,58	0,60
18 à 19 ans
20 à 24 ans	5,31	0,38	5,51	0,29	8,66	0,46
25 à 34 ans	2,85	0,28	2,31	0,28	4,51	0,35
35 à 44 ans	1,54	0,26	2,20	0,26	2,32	0,31
45 à 54 ans	1,22	0,33	0,81	0,23	1,58	0,29
55 à 64 ans	0,92	0,20	0,91	0,29	2,06	0,31
65 ans et plus	1,20	0,25	0,71	0,30	1,76	0,31
Hommes³	2,46	0,17	2,37	0,13	3,75	0,16
0 à 4 ans	2,53	0,46	1,32	0,33	2,22	0,67
5 à 14 ans	1,14	0,21	1,27	0,29	1,98	0,32
15 à 17 ans ²	1,93	0,48	3,12	0,68	4,09	0,74
18 à 19 ans
20 à 24 ans	5,99	0,52	6,03	0,48	10,36	0,57
25 à 34 ans	3,64	0,46	2,70	0,44	5,43	0,45
35 à 44 ans	2,33	0,48	3,42	0,40	3,29	0,51
45 à 54 ans	1,63	0,41	1,21	0,38	1,95	0,52
55 à 64 ans	1,28	0,34	0,91	0,40	1,88	0,47
65 ans et plus	1,90	0,44	0,69	0,47	1,57	0,50

Rapport technique sur la couverture

Tableau 11.1b

Estimations du taux de sous-dénombrement de la population et de l'erreur-type, selon le sexe et le groupe d'âge pour le Canada, Recensement de 1976 au Recensement de 2021

Sexe ¹ et groupe d'âge	1976		1981		1986			
	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type		
	pourcentage							
Femmes⁴	1,61	0,10	1,65	0,12	2,68	0,17		
0 à 4 ans	2,07	0,36	1,10	0,33	2,06	0,62		
5 à 14 ans	1,26	0,27	1,19	0,31	2,20	0,33		
15 à 17 ans ²	2,05	0,51	2,80	0,73	3,05	0,76		
18 à 19 ans		
20 à 24 ans	4,62	0,48	4,98	0,43	6,89	0,72		
25 à 34 ans	2,03	0,38	1,92	0,32	3,59	0,45		
35 à 44 ans	0,72	0,24	0,93	0,31	1,33	0,32		
45 à 54 ans	0,81	0,38	0,41	0,26	1,20	0,35		
55 à 64 ans	0,58	0,25	0,92	0,34	2,23	0,50		
65 ans et plus	0,64	0,38	0,71	0,42	1,89	0,44		
Sexe ¹ et groupe d'âge	1991		1996		2001		2006	
	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type
	pourcentage							
Les deux sexes	3,43	0,12	3,18	0,09	3,95	0,13	4,26	0,17
0 à 4 ans	3,55	0,49	2,89	0,36	4,42	0,71	4,07	0,65
5 à 14 ans	2,49	0,27	1,45	0,14	2,90	0,38	3,10	0,46
15 à 17 ans ²	3,75	0,42	3,48	0,42	4,36	0,53	1,56	0,60
18 à 19 ans	8,86	1,58
20 à 24 ans	8,18	0,52	8,00	0,34	9,85	0,62	10,50	0,74
25 à 34 ans	5,65	0,35	5,81	0,29	8,07	0,36	9,43	0,56
35 à 44 ans	2,84	0,29	2,78	0,24	4,04	0,33	5,36	0,50
45 à 54 ans	1,61	0,27	1,90	0,21	1,79	0,29	2,64	0,43
55 à 64 ans	1,69	0,28	2,23	0,34	1,22	0,37	0,95	0,53
65 ans et plus	1,51	0,28	1,52	0,26	1,29	0,34	0,21	0,40

Rapport technique sur la couverture

Tableau 11.1b

Estimations du taux de sous-dénombrement de la population de d'erreur-type, selon le sexe et le groupe d'âge pour le Canada, Recensement de 1976 au Recensement de 2021

Sexe ¹ et groupe d'âge	1991		1996		2001		2006	
	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type
	pourcentage							
Hommes³	3,95	0,16	3,89	0,14	4,90	0,19	5,51	0,26
0 à 4 ans	2,79	0,58	2,56	0,47	3,36	0,89	4,24	0,95
5 à 14 ans	2,32	0,34	1,46	0,24	2,38	0,49	3,04	0,64
15 à 17 ans ²	3,55	0,60	3,68	0,43	5,49	0,80	1,88	0,88
18 à 19 ans	10,06	2,45
20 à 24 ans	8,98	0,81	9,48	0,50	11,68	0,92	12,21	1,12
25 à 34 ans	7,28	0,56	7,74	0,42	10,67	0,55	11,42	0,86
35 à 44 ans	3,65	0,41	3,94	0,39	5,71	0,51	7,77	0,79
45 à 54 ans	2,05	0,45	2,12	0,27	2,50	0,44	4,14	0,69
55 à 64 ans	2,04	0,44	2,50	0,54	1,35	0,54	2,13	0,77
65 ans et plus	1,41	0,50	1,64	0,45	1,50	0,53	-0,05	0,56
Femmes⁴	2,93	0,17	2,49	0,12	3,02	0,18	3,04	0,23
0 à 4 ans	4,35	0,71	3,24	0,55	5,50	1,14	3,88	0,92
5 à 14 ans	2,65	0,39	1,45	0,22	3,44	0,58	3,17	0,66
15 à 17 ans ²	3,96	0,54	3,28	0,55	3,13	0,69	1,23	0,83
18 à 19 ans	7,58	1,96
20 à 24 ans	7,36	0,71	6,45	0,48	7,91	0,84	8,70	0,98
25 à 34 ans	3,98	0,37	3,84	0,40	5,41	0,46	7,43	0,73
35 à 44 ans	2,01	0,35	1,62	0,28	2,35	0,43	2,90	0,61
45 à 54 ans	1,16	0,34	1,68	0,33	1,09	0,37	1,13	0,51
55 à 64 ans	1,35	0,33	1,97	0,40	1,09	0,52	-0,22	0,73
65 ans et plus	1,58	0,36	1,43	0,32	1,13	0,45	0,40	0,56

Rapport technique sur la couverture

Tableau 11.1b

Estimations du taux de sous-dénombrement de la population de d'erreur-type, selon le sexe et le groupe d'âge pour le Canada, Recensement de 1976 au Recensement de 2021

Sexe ¹ et groupe d'âge	2011		2016		2021	
	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type
	pourcentage					
Les deux sexes	4,07	0,16	4,32	0,11	4,98	0,09
0 à 4 ans	3,36	0,62	3,79	0,51	5,19	0,29
5 à 14 ans	2,61	0,42	2,94	0,37	3,35	0,29
15 à 17 ans ²	3,83	0,85	4,35	0,78	3,40	0,68
18 à 19 ans	6,28	0,93	7,71	1,25	5,91	1,03
20 à 24 ans	9,60	0,69	9,64	0,58	12,91	0,66
25 à 34 ans	8,96	0,48	8,60	0,42	9,89	0,39
35 à 44 ans	4,66	0,45	5,07	0,41	5,49	0,36
45 à 54 ans	2,95	0,42	3,86	0,39	4,01	0,38
55 à 64 ans	1,02	0,41	2,21	0,42	3,10	0,34
65 ans et plus	1,19	0,45	0,77	0,29	1,25	0,24
Hommes³	5,07	0,24	5,27	0,18	6,01	0,17
0 à 4 ans	3,14	0,82	4,22	0,68	4,80	0,56
5 à 14 ans	3,00	0,62	2,35	0,54	3,57	0,39
15 à 17 ans ²	4,31	1,13	4,55	1,11	3,53	1,02
18 à 19 ans	5,42	1,19	7,38	1,82	5,70	1,34
20 à 24 ans	9,37	0,88	11,46	0,88	12,97	0,92
25 à 34 ans	10,54	0,73	10,44	0,63	12,00	0,60
35 à 44 ans	6,34	0,68	6,24	0,63	7,48	0,58
45 à 54 ans	4,69	0,66	5,16	0,59	4,89	0,60
55 à 64 ans	2,58	0,69	3,11	0,67	4,42	0,54
65 ans et plus	1,32	0,60	1,23	0,41	1,64	0,39
Femmes⁴	3,08	0,22	3,39	0,18	3,95	0,14
0 à 4 ans	3,59	0,95	3,34	0,87	5,60	0,57
5 à 14 ans	2,20	0,57	3,56	0,56	3,12	0,44
15 à 17 ans ²	3,31	1,28	4,13	1,06	3,26	0,94
18 à 19 ans	7,17	1,45	8,07	1,77	6,14	1,59
20 à 24 ans	9,83	1,07	7,66	0,77	12,86	1,06
25 à 34 ans	7,37	0,63	6,71	0,56	7,66	0,53
35 à 44 ans	2,99	0,58	3,93	0,53	3,49	0,39
45 à 54 ans	1,21	0,52	2,56	0,55	3,14	0,50
55 à 64 ans	-0,52	0,44	1,32	0,50	1,80	0,42
65 ans et plus	1,08	0,66	0,39	0,42	0,91	0,33

.. indisponible pour une période de référence précise

1. Pour 2021, les estimations ont été produites en fonction du genre plutôt que du sexe.

2. Les données pour toutes les années, sauf celles de 2006 à 2021, concernent les personnes de 15 à 19 ans.

3. Pour 2021, les données concernent les hommes+. Le terme « hommes+ » comprend les hommes (et les garçons) et certaines personnes non binaires.

4. Pour 2021, les données concernent les femmes+. Le terme « femmes+ » comprend les femmes (et les filles) et certaines personnes non binaires.

Notes : À l'exclusion des réserves partiellement dénombrées. Les chiffres du Recensement de 1991 et ceux des recensements subséquents comprennent les résidents non permanents et les territoires. Les taux tiennent compte des modifications apportées à la publication initiale de 1986. Les taux ne comprennent pas les estimations des personnes omises dans les logements classés par erreur comme étant inoccupés en 1976.

Sources : Statistique Canada, études sur la couverture des recensements de 1976 à 2021.

Ces tableaux présentent les observations suivantes.

Le sous-dénombrement est généralement supérieur dans les trois territoires. Parmi les provinces, le sous-dénombrement est généralement supérieur en Colombie-Britannique et en Ontario. Toutefois, dans les récents recensements, le taux de sous-dénombrement en Alberta et en Saskatchewan était proche de celui de l'Ontario et l'a même dépassé à quelques occasions. Entre 1971 et 2001, la Colombie-Britannique a été la province ayant enregistré le taux de sous-dénombrement le plus élevé lors de chaque recensement, sauf en 1991, lorsque l'Ontario a occupé ce rang. L'Ontario a enregistré le taux le plus élevé en 2006 également, alors qu'en 2011, cela a été le tour de l'Alberta. En 2016 et 2021, la Colombie-Britannique a à nouveau enregistré le taux le plus élevé, suivie en 2021 par l'Île-du-Prince-Édouard, province n'ayant jamais enregistré l'un des taux les plus élevés par le passé. Le Québec et les provinces de l'Atlantique ont tendance à afficher un taux de sous-dénombrement inférieur au taux national.

Le sous-dénombrement était supérieur pour les jeunes adultes et pour les hommes+ par rapport aux femmes+ en 2021, deux tendances démographiques persistantes. Avant les études sur la couverture de 2021, les erreurs de couverture étaient estimées selon le sexe plutôt que selon le genre. Le [tableau 11.1b](#) montre que le sous-dénombrement était plus élevé pour les hommes que pour les femmes pour tous les recensements depuis 1971. Le tableau 11.1b révèle également que le taux de sous-dénombrement des jeunes hommes âgés de 20 à 24 ans était supérieur au taux de sous-dénombrement pour l'ensemble des hommes. Il en va de même pour les femmes âgées de 20 à 24 ans, mais le taux pour ces jeunes femmes était inférieur au taux pour les hommes du même groupe d'âge pour tous les recensements, à l'exception de celui de 2011, pour lequel on constate des taux respectifs de 9,83 % et de 9,37 %. En 2021, le taux de sous-dénombrement pour les femmes+ âgées de 20 à 24 ans était très proche de celui des hommes+; tous deux dépassant tous les taux observés pour les hommes et les femmes dans les précédents recensements. Les taux de sous-dénombrement des adultes âgés de 25 à 34 ans étaient également élevés. Le taux de sous-dénombrement était plus élevé dans le cas des jeunes adultes; ce qui est en partie attribuable à leur mode de vie moins stable. En effet, les jeunes adultes sont plus susceptibles de changer leurs modalités de vie que les adultes plus âgés ou que les enfants, étant donné qu'ils quittent la maison pour fréquenter un établissement d'enseignement postsecondaire, travailler ou emménager avec des amis ou des conjoints.

Le tableau 11.1c et le [tableau 11.1d](#) présentent les estimations des taux de surdénombrement.

Rapport technique sur la couverture

Tableau 11.1c

Estimations du taux de surdénombrement de la population et de l'erreur-type pour le Canada, les provinces et les territoires, Recensement de 1991 au Recensement de 2021

Province ou territoire	1991		1996		2001	
	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type
	pourcentage					
Canada	0,56	0,04	0,74	0,04	0,96	0,05
Terre-Neuve-et-Labrador	0,48	0,09	0,77	0,12	0,63	0,10
Île-du-Prince-Édouard	0,74	0,15	0,91	0,14	0,92	0,18
Nouvelle-Écosse	0,36	0,09	0,47	0,07	0,81	0,14
Nouveau-Brunswick	0,46	0,09	0,60	0,10	0,89	0,19
Québec	0,51	0,07	0,85	0,08	1,03	0,10
Ontario	0,59	0,07	0,67	0,07	0,88	0,09
Manitoba	0,45	0,11	0,88	0,15	0,80	0,15
Saskatchewan	0,35	0,08	0,55	0,11	1,06	0,20
Alberta	0,51	0,09	0,59	0,10	0,89	0,13
Colombie-Britannique	0,68	0,10	0,89	0,09	1,26	0,12
Yukon	0,29	0,07	0,70	0,17	0,86	0,16
Territoires du Nord-Ouest	0,29	0,07	1,32	0,22	1,00	0,11
Nunavut	0,99	0,22	0,59	0,10

Province ou territoire	2006		2011		2016		2021	
	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type
	pourcentage							
Canada	1,59	0,01	1,85	0,02	1,96	0,04	1,98	0,02
Terre-Neuve-et-Labrador	1,63	0,05	1,76	0,04	2,09	0,07	1,94	0,08
Île-du-Prince-Édouard	1,66	0,06	1,54	0,04	1,65	0,07	2,04	0,12
Nouvelle-Écosse	1,40	0,03	1,72	0,04	1,81	0,07	1,94	0,07
Nouveau-Brunswick	1,41	0,03	2,12	0,05	2,18	0,08	2,10	0,08
Québec	1,66	0,02	2,07	0,03	2,15	0,12	2,14	0,07
Ontario	1,49	0,02	1,67	0,04	1,87	0,08	1,82	0,05
Manitoba	1,42	0,04	1,35	0,04	1,51	0,06	1,68	0,05
Saskatchewan	1,53	0,04	1,65	0,04	1,91	0,14	2,07	0,06
Alberta	1,47	0,02	1,70	0,05	1,74	0,06	1,77	0,06
Colombie-Britannique	1,96	0,03	2,28	0,05	2,24	0,07	2,41	0,05
Yukon	1,62	0,08	2,45	0,05	2,22	0,08	1,94	0,09
Territoires du Nord-Ouest	1,98	0,08	1,44	0,11	1,25	0,06	1,87	0,03
Nunavut	1,44	0,07	1,17	0,07	1,73	0,07	1,88	0,04

... n'ayant pas lieu de figurer

Notes : À l'exclusion des réserves partiellement dénombrées. Les chiffres comprennent les résidents non permanents.

Sources : Statistique Canada, études sur la couverture des recensements de 1991 à 2021.

Rapport technique sur la couverture

Tableau 11.1d

Estimations du taux de surdénombrement de la population et de l'erreur-type, selon le sexe et le groupe d'âge pour le Canada, Recensement de 1996 au Recensement de 2021

Sexe ¹ et groupe d'âge	1996		2001		2006	
	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type
	pourcentage					
Les deux sexes	0,74	0,04	0,96	0,05	1,59	0,01
0 à 4 ans	0,61	0,10	0,96	0,18	1,35	0,07
5 à 14 ans	0,96	0,09	1,52	0,15	2,24	0,07
15 à 17 ans ²	1,24	0,15	1,85	0,26	2,33	0,14
18 à 19 ans	2,65	0,17
20 à 24 ans	2,44	0,28	2,66	0,32	2,88	0,11
25 à 34 ans	0,66	0,08	0,92	0,09	1,43	0,06
35 à 44 ans	0,38	0,06	0,49	0,06	1,05	0,05
45 à 54 ans	0,48	0,11	0,39	0,04	1,13	0,05
55 à 64 ans	0,52	0,11	0,38	0,05	1,24	0,06
65 ans et plus	0,36	0,07	0,77	0,21	1,60	0,06
Hommes³	0,70	0,04	0,92	0,06	1,62	0,02
0 à 4 ans	0,52	0,09	0,69	0,07	1,35	0,09
5 à 14 ans	0,99	0,15	1,59	0,21	2,25	0,10
15 à 17 ans ²	1,12	0,24	1,45	0,31	2,37	0,20
18 à 19 ans	2,28	0,21
20 à 24 ans	2,34	0,34	2,44	0,45	2,75	0,15
25 à 34 ans	0,65	0,11	1,03	0,14	1,51	0,08
35 à 44 ans	0,38	0,06	0,46	0,06	1,10	0,06
45 à 54 ans	0,35	0,07	0,34	0,03	1,16	0,07
55 à 64 ans	0,37	0,12	0,33	0,04	1,30	0,09
65 ans et plus	0,33	0,02	0,74	0,21	1,69	0,10
Femmes⁴	0,77	0,06	1,00	0,08	1,56	0,01
0 à 4 ans	0,69	0,18	1,25	0,36	1,35	0,10
5 à 14 ans	0,92	0,14	1,44	0,21	2,23	0,10
15 à 17 ans ²	1,36	0,29	2,27	0,43	2,28	0,19
18 à 19 ans	3,04	0,28
20 à 24 ans	2,55	0,46	2,89	0,46	3,01	0,17
25 à 34 ans	0,66	0,11	0,81	0,12	1,35	0,08
35 à 44 ans	0,37	0,10	0,53	0,11	0,99	0,06
45 à 54 ans	0,61	0,20	0,43	0,07	1,11	0,06
55 à 64 ans	0,66	0,19	0,42	0,09	1,18	0,07
65 ans et plus	0,38	0,11	0,80	0,33	1,53	0,08

Rapport technique sur la couverture

Tableau 11.1d

Estimations du taux de surdénombrement de la population et de l'erreur-type, selon le sexe et le groupe d'âge pour le Canada, Recensement de 1996 au Recensement de 2021

Sexe ¹ et groupe d'âge	2011		2016		2021	
	Taux estimé	Erreur-type	Taux estimé	Erreur-type	Taux estimé	Erreur-type
	pourcentage					
Les deux sexes	1,85	0,02	1,96	0,04	1,98	0,02
0 à 4 ans	1,61	0,10	1,76	0,19	1,74	0,08
5 à 14 ans	2,79	0,10	3,49	0,20	3,21	0,09
15 à 17 ans ²	2,98	0,23	3,25	0,35	3,58	0,19
18 à 19 ans	3,37	0,27	3,27	0,40	3,71	0,26
20 à 24 ans	3,11	0,13	3,51	0,26	3,52	0,15
25 à 34 ans	1,69	0,08	2,19	0,17	2,06	0,08
35 à 44 ans	1,23	0,06	1,12	0,11	1,25	0,06
45 à 54 ans	1,36	0,06	1,18	0,10	1,32	0,07
55 à 64 ans	1,50	0,07	1,44	0,14	1,51	0,07
65 ans et plus	1,64	0,08	1,55	0,10	1,59	0,05
Hommes³	1,86	0,04	1,95	0,07	2,01	0,04
0 à 4 ans	1,65	0,13	1,56	0,24	1,69	0,11
5 à 14 ans	2,77	0,14	3,50	0,28	3,23	0,12
15 à 17 ans ²	3,04	0,27	4,14	0,58	3,72	0,29
18 à 19 ans	3,06	0,33	2,56	0,44	3,55	0,35
20 à 24 ans	3,03	0,19	3,27	0,37	3,53	0,21
25 à 34 ans	1,75	0,11	2,29	0,24	2,12	0,12
35 à 44 ans	1,26	0,09	1,21	0,16	1,31	0,11
45 à 54 ans	1,32	0,09	1,17	0,13	1,27	0,07
55 à 64 ans	1,54	0,11	1,22	0,20	1,52	0,09
65 ans et plus	1,68	0,14	1,47	0,15	1,59	0,08
Femmes⁴	1,83	0,04	1,98	0,07	1,95	0,04
0 à 4 ans	1,57	0,15	1,96	0,31	1,78	0,12
5 à 14 ans	2,81	0,15	3,49	0,29	3,20	0,12
15 à 17 ans ²	2,93	0,38	2,32	0,38	3,44	0,25
18 à 19 ans	3,69	0,43	4,02	0,68	3,89	0,38
20 à 24 ans	3,19	0,18	3,77	0,38	3,50	0,21
25 à 34 ans	1,63	0,11	2,08	0,23	1,99	0,10
35 à 44 ans	1,20	0,10	1,04	0,16	1,18	0,06
45 à 54 ans	1,39	0,09	1,19	0,14	1,36	0,12
55 à 64 ans	1,46	0,10	1,66	0,20	1,50	0,10
65 ans et plus	1,60	0,10	1,63	0,14	1,60	0,08

.. indisponible pour une période de référence précise

1. Pour 2021, les estimations ont été produites en fonction du genre plutôt que du sexe.

2. Les données de 1996 et de 2001 concernent les personnes de 15 à 19 ans.

3. Pour 2021, les données concernent les hommes+. Le terme « hommes+ » comprend les hommes (et les garçons) et certaines personnes non binaires.

4. Pour 2021, les données concernent les femmes+. Le terme « femmes+ » comprend les femmes (et les filles) et certaines personnes non binaires.

Note : À l'exclusion des réserves partiellement dénombrées.

Sources : Statistique Canada, études sur la couverture des recensements de 1996 à 2021.

Ces tableaux présentent les observations suivantes.

Le surdénombrement est constamment plus élevé en Colombie-Britannique que dans les autres provinces. La Colombie-Britannique est la province où le taux de surdénombrement de la population a été le plus élevé pour les cinq derniers recensements. Des trois territoires, le Yukon a enregistré le taux de surdénombrement le plus élevé pour un troisième recensement consécutif, alors que les Territoires du Nord-Ouest ont affiché le taux le plus élevé de 1996 à 2006. Le surdénombrement a augmenté dans huit des provinces et territoires en 2021 et a baissé dans les cinq autres, mais l'écart était statistiquement significatif dans moins de la moitié de ces cas. Le taux de surdénombrement national est demeuré pratiquement identique entre 2016 et 2021. Le Manitoba a enregistré le taux le plus faible de toutes les provinces à chacun des trois derniers recensements; se classant au deuxième rang le plus bas uniquement devant les Territoires du Nord-Ouest en 2016 parmi l'ensemble des provinces et territoires.

Comme l'indique le [tableau 11.1d](#), le surdénombrement était plus courant pour les enfants d'âge scolaire et les jeunes adultes. Dans le cas des enfants d'âge scolaire, cette situation était attribuable en grande partie au fait que les enfants dont les parents ne vivent pas ensemble sont souvent dénombrés par chacun d'eux. Le surdénombrement dans le cas des jeunes adultes est vraisemblablement attribuable aux modes de vie moins stables, qui peuvent également donner lieu à un sous-dénombrement. À l'échelle nationale, les taux de surdénombrement dépassaient les 3 % chez les jeunes adultes (âgés de 18 à 24 ans), ainsi que chez les jeunes de 5 à 17 ans. Le groupe d'âge des 15 à 19 ans a enregistré la plus forte hausse du taux de surdénombrement par rapport à 2016.

11.2 Modifications apportées à la conception des études sur la couverture de la population

En raison des différences apportées à la conception des études sur la couverture au fil des ans, les taux indiqués dans les tableaux [11.1a](#), [11.1b](#), [11.1c](#) et [11.1d](#) ne sont pas strictement comparables. Vous trouverez ci-dessous une liste des modifications apportées à la méthodologie depuis 1976. Il convient de noter que les notions fondamentales de l'approche de mesure du sous-dénombrement dans le cadre de l'Étude sur le sous-dénombrement du recensement (ESoR) (appelée la contre-vérification des dossiers jusqu'au Recensement de 2021) n'ont pas beaucoup changé depuis le Recensement de 1966. On prélève un échantillon à partir de bases de sondage couvrant la population visée indépendamment du recensement. On vérifie ensuite les dossiers du recensement pour déterminer si les personnes échantillonnées ont effectivement été dénombrées. D'autres modifications ont été apportées à la mesure du surdénombrement. Plusieurs études ont été réalisées en 1991, 1996 et 2001. En 1996, la portée de l'ESoR a été élargie de façon à prendre en compte le surdénombrement. En 2006, l'ESoR a cessé d'être utilisée pour estimer le surdénombrement et l'on a mis en place une nouvelle étude visant à mesurer tous les cas de surdénombrement en fonction de couplages probabilistes et exacts à partir du nom, de la date de naissance et du sexe.

Études sur la couverture du Recensement de 2021

- a) La stratification des bases de sondage pour le recensement de l'ESoR et les territoires a été améliorée. Ces bases de sondage complètes ont été couplées à la Base de données des réponses du recensement (BDR) de 2021 et aux fichiers de données fiscales, ainsi qu'aux enregistrements de décès des statistiques de l'état civil dans le cas de la base du recensement. Selon les renseignements obtenus de ces couplages, les unités de ces bases de sondage ont alors été stratifiées en fonction de leur probabilité à être dénombrées, omises ou hors du champ du recensement. Cela a conduit à un plan de sondage plus efficace.
- b) Des modifications ont été apportées aux opérations de collecte de l'ESoR. L'ESoR est passée du système Blaise au Système intégré de collecte et des opérations, application standard sur le Web élaborée par Statistique Canada et utilisée depuis le Recensement de 2016. Le questionnaire papier envoyé sur demande aux répondants ou lors des opérations de suivi des cas de non-réponse a été remplacé par un questionnaire électronique d'autodéclaration. De plus, des mises à jour de dépistage ont été fournies pendant la collecte pour aider à trouver et interviewer les répondants sélectionnés.

- c) Pour améliorer la classification des personnes dont la résidence principale ne se trouvait pas au Canada le jour du Recensement de 2021, on a eu recours à une imputation par donneur pour la non-réponse partielle au module hors du Canada du questionnaire de l'ESoR.
- d) Dans l'ESoR, on a ajouté à la méthodologie de pondération un ajustement de certains poids influents et un étalonnage des poids pour les bases des naissances et des immigrants.
- e) Un couplage déterministe a été ajouté aux étapes utilisées pour créer la base de sondage des potentiels doublons du recensement de l'Étude sur le surdénombrement du recensement (ESuR). Ce couplage déterministe se fonde sur la modification de la méthodologie de l'Étude par appariement automatisé (EAA).
- f) La Méthode généralisée de partage des poids (MGPP) a été utilisée pour calculer un poids pour les paires de surdénombrement ne figurant pas dans la base de sondage de l'ESuR détectées lors de la vérification manuelle des paires de ménages associées avec des paires de personnes échantillonnées, afin que ces paires puissent être incluses dans les estimations de l'ESuR.
- g) Les estimations des erreurs de couverture par groupe d'âge et sexe ont été remplacées par des estimations par groupe d'âge et genre, reposant sur deux catégories de genre (hommes+ et femmes+).

Études sur la couverture du Recensement de 2016

- a) Dans le cas de la base de sondage de l'ESoR, la stratification se fondait sur la province de résidence mise à jour à l'aide des données fiscales. De plus, une strate de décès a été déterminée avant la sélection de l'échantillon. Ces améliorations ont permis de rendre l'échantillon plus efficace.
- b) Un ménage « administratif » a été créé à l'aide de données fiscales pour chaque personne choisie de l'ESoR, afin d'accroître l'efficacité des processus de traitement et de dépistage.
- c) L'ajustement relatif à la non-réponse de l'ESoR pour les personnes non dépistées ne fait plus appel au sous-échantillon. L'ajustement se fait au sein de groupes de réponses homogènes formés en regroupant des personnes ayant des probabilités semblables d'être dans la population cible et de répondre à l'enquête, telles qu'estimées à l'aide de données disponibles à la fois pour les répondants et les non-répondants. Le modèle d'ajustement de la non-réponse ayant été finalisé après la collecte, le sous-échantillon qui servait à l'ajustement pour les personnes non dépistées lors des ESoR précédentes a tout de même été envoyé à la collecte, mais n'a pas été utilisé. Cela ne sera pas nécessaire pour les ESoR suivantes.
- d) Un calage a été effectué, afin d'ajuster les poids de la base du recensement dans le cas où un échantillon provincial contiendrait trop ou trop peu de personnes dénombrées ou décédées. Pour les trois territoires, un calage au nombre comparable de personnes dénombrées a été effectué comme lors des ESoR précédentes, sauf qu'il a été réalisé séparément pour six groupes d'âge-sexe dans chaque territoire, plutôt que d'utiliser un seul groupe de calage global pour chacun d'eux, comme cela était le cas par le passé.
- e) Des vagues déterministes hiérarchiques ont été utilisées pour former les blocs servant aux couplages probabilistes des étapes 1 et 2 de l'ESuR, afin d'améliorer la couverture de la base.
- f) Les règles de couplage des étapes 1 et 2 ont été élargies, afin de détecter le plus de cas possible de surdénombrement potentiels.
- g) Alors qu'en 2011, l'étape 3 (extension) de l'ESuR se limitait aux paires de ménages privés, elle a été étendue aux paires de ménages produites pour chaque paire de personnes créée aux étapes 1 et 2 et pour lesquelles l'un des deux ménages se trouvait dans un logement collectif.
- h) Les groupes d'enregistrements pouvant représenter la même personne ont été créés non pas séparément pour chaque étape, mais en combinant les étapes 1, 2 et 3 de l'ESuR, afin de mieux tenir compte des cas de surdénombrement multiples.

Tout comme chaque ESoR depuis 1996, celle de 2016 n'a pas permis d'estimer le nombre de personnes omises dans le cas de réserves et d'établissements partiellement dénombrés. Pour obtenir de plus amples renseignements à ce sujet, voir la [section 12.2](#).

Études sur la couverture du Recensement de 2011

- a) L'ESoR de 2011 a été très semblable à celle de 2006. Pour la rendre plus efficiente, certaines modifications ont été apportées, y compris des améliorations au programme de méga-appariement, des stratégies de recherche dans la BDR du recensement plus efficaces et l'utilisation de nouvelles bases de sondage pour les naissances.
- b) Pour la première fois, la pondération de l'échantillon de la base du recensement a tenu compte du surdénombrement présent dans cette base.
- c) Le recours à des méthodes automatisées a permis d'utiliser des paramètres provinciaux et territoriaux au lieu de paramètres nationaux lors de l'élaboration de la base de sondage de l'ESuR.

Études sur la couverture du Recensement de 2006

Dans le cadre de l'ESoR et de l'ESuR, le champ de nom ajouté à la BDR du Recensement de 2006 a été utilisé de façon optimale lors des opérations d'appariement et de recherche. Les modifications suivantes ont également été apportées.

- a) La mesure du surdénombrement a été limitée à l'ESuR. La méthodologie de l'ESoR a été modifiée par la suite de façon à ce que les cas ne fassent pas tous l'objet d'une collecte sur le terrain. Depuis 2006, l'ESoR comporte une étape de traitement antérieure à la collecte, qui permet de déterminer si des activités de collecte sont nécessaires. On effectue une recherche dans la version de la BDR du recensement utilisée pour réaliser les études de couverture du recensement (BDR-ECR) afin de trouver les personnes échantillonnées à partir des renseignements des bases de sondage et de diverses sources de mise à jour, comme les données fiscales. Si la recherche permet de retrouver la personne échantillonnée dans la BDR-ECR, aucune mesure de collecte n'est nécessaire. La seule exception vise un échantillon de personnes trouvées, afin de recueillir les données requises aux fins de l'ajustement pour la non-réponse.
- b) Les trois études de couverture menées en 2001 pour mesurer le surdénombrement ont été remplacées par l'ESuR en 2006. Cette étude a fait appel à une méthodologie différente de celle employée dans les études antérieures sur le surdénombrement. Elle reposait essentiellement sur une technique de couplage utilisant les noms et prénoms, le sexe et la date de naissance ainsi que sur des vérifications manuelles permettant de mesurer le surdénombrement.

Études sur la couverture du Recensement de 2001

- a) La composante institutionnelle de l'Étude sur les logements collectifs a été abandonnée, et l'estimation du surdénombrement visant cette population a été établie à l'aide de l'ESoR.
- b) L'Étude sur la classification des logements a remplacé la vérification des logements inoccupés. Cette dernière était effectuée lors des recensements précédents pour vérifier la classification des logements que l'agent recenseur avait déterminés comme étant inoccupés. L'Étude sur la classification des logements était un prolongement de la vérification des logements inoccupés qui permettait d'estimer le nombre de personnes vivant dans des logements non répondants.

Études sur la couverture du recensement de 1996

- a) L'ESoR de 1996 n'a pas estimé le nombre de personnes omises dans le cas de réserves partiellement dénombrées.
- b) L'Étude sur les résidents temporaires a été annulée en raison de préoccupations à propos de la qualité des données et parce qu'on a reconnu que l'ESoR permettrait de mesurer ce type de sous-dénombrement de façon appropriée.

- c) On a procédé à une évaluation du surdénombrement plus exhaustive qu'en 1991 en intégrant l'Étude sur les logements privés à l'ESoR, de façon à pouvoir déterminer que chaque personne échantillonnée était une personne dénombrée plus d'une fois. Cette approche a donné lieu à une augmentation des adresses à traiter pour lesquelles il pourrait y avoir eu surdénombrement. D'autre part, la portée de l'EAA a été considérablement élargie par rapport à 1991, de façon à ce que le surdénombrement puisse être déterminé non seulement dans le cas d'un secteur de dénombrement, mais aussi dans une grande région (provinces de l'Atlantique, Québec, Ontario, Ouest canadien et territoires).

Études sur la couverture du recensement de 1991

- a) Les résidents non permanents ont été inclus dans la population cible pour la première fois.
- b) Après les études expérimentales de 1986, l'évaluation du taux de surdénombrement de la population a commencé en 1991. Les résultats de trois études (l'Étude sur les logements privés, l'Étude sur les logements collectifs et l'Étude par appariement automatisé) ont été combinés pour former une estimation complète.

Études sur la couverture du recensement de 1986

Les taux présentés dans les [tableaux 11.1a](#) et [11.1b](#) pour le Recensement de 1986 diffèrent des résultats publiés dans le *Guide à l'intention des utilisateurs sur la qualité des données du recensement de 1986 : Couverture*, qui comprenait des révisions effectuées après la publication de 1986, dans les cas où les réserves partiellement dénombrées avaient été considérées comme des omissions. Lors de la publication initiale de 1986, elles étaient comptées à titre de cas « dénombrés » étant donné que les données provinciales prenaient en compte d'une estimation des personnes omises dans le cas des réserves.

Études sur la couverture du recensement de 1976

Les données du recensement ne comprenaient pas d'estimation de la vérification des logements inoccupés dans le cas des personnes omises en cas de logements classés par erreur comme étant des logements inoccupés. Le taux de sous-dénombrement de la population aurait été de 1,78 % en 1976 si l'on avait pris en compte les résultats de la vérification des logements inoccupés de 1976. Aucune vérification des logements inoccupés n'a été effectuée lors du Recensement de 1971.

Pour obtenir de plus amples détails sur l'historique des études de couverture, voir [Dolson \(2010\)](#).

12. Sujets spéciaux

12.1 Sous-dénombrement de collecte

Jusqu'à présent, il a été question du sous-dénombrement des chiffres de population du recensement. La présente section traite du concept de sous-dénombrement de la collecte du Recensement de la population. Il est utile d'élargir le concept de sous-dénombrement de façon à tenir compte des personnes non dénombrées pour une raison quelconque. Le sous-dénombrement désigne le nombre de personnes non incluses dans les chiffres du recensement. Comme le présente la [section 3.3](#), les chiffres de population C comprennent deux éléments : $C = D + I$, où D = le nombre de dénombremments et I = le nombre de personnes imputées.

Le sous-dénombrement représente donc seulement une partie de toutes les personnes non inscrites dans un questionnaire du recensement, mais qui auraient dû l'être. Il n'inclut pas les personnes qui n'ont pas été dénombrées, soit parce qu'aucun questionnaire du recensement n'a été retourné pour le logement (logement non répondant), soit parce que le logement n'a pas fait l'objet d'un suivi des cas de non-réponse en ayant été classé par erreur comme étant un logement inoccupé (logement occupé classé par erreur comme étant inoccupé).

Toujours selon la [section 3.3](#), une estimation du nombre réel de personnes visées par le recensement est :

$$\hat{P} = C + \hat{N} = C + \hat{S} - \hat{R}$$

En combinant ces deux équations, on obtient :

$$\hat{P} = C + \hat{N} = C + \hat{S} - \hat{R} = D + (I + \hat{S}) - \hat{R}$$

Cette formulation de \hat{P} comporte trois éléments :

D = le nombre de personnes figurant dans un questionnaire du recensement¹⁵ (dénombremments)

\hat{R} = une estimation du nombre de dénombremments excédentaires¹⁶

$(I + \hat{S})$ = une estimation du nombre de personnes non inscrites dans un questionnaire du recensement, mais qui auraient dû l'être.

Le dernier élément, $(I + \hat{S})$, désigne une estimation du nombre de personnes omises dans le cadre du recensement pour une raison quelconque. Le sous-dénombrement de la collecte du Recensement de la population, (L), désigne les personnes non dénombrées pour une raison quelconque. L'estimation du sous-dénombrement de la collecte du recensement est la suivante :

$$\hat{L} = (I + \hat{S})$$

L'estimation correspondante du taux de sous-dénombrement de la collecte du Recensement de la population est la suivante :

$$\hat{T}_L = 100 * \frac{\hat{L}}{\hat{P}} = 100 * \left(\frac{I + \hat{S}}{C + \hat{N}} \right)$$

15. Il se pourrait que certaines personnes inscrites dans un questionnaire ne figurent pas dans la base de données définitive du recensement. Ainsi, les « personnes inscrites dans un questionnaire du recensement » est une expression utilisée dans la présente section pour indiquer les personnes figurant dans la base de données définitive du recensement.

16. La plupart des cas de surdénombrement désignent des dénombremments en double; c.-à-d. lorsque la même personne figure deux fois dans la base de données. Dans un petit nombre de cas, cependant, la même personne peut y figurer plus de deux fois. La variable \hat{R} désigne l'estimation du nombre de dénombremments excédentaires plutôt que le nombre de personnes visées par un dénombrement multiple.

Rapport technique sur la couverture

Le sous-dénombrement net de la collecte du Recensement de la population peut être calculé en soustrayant le surdénombrement \hat{R} de \hat{L} :

$$\hat{P} = C + \hat{N} = D + (I + \hat{S}) - \hat{R} = D + \hat{L} - \hat{R}$$

Bien que l'on ne puisse pas appliquer le sous-dénombrement net de la collecte aux données du recensement pour effectuer un rajustement à titre d'erreur de couverture, \hat{L} et \hat{T}_L donnent une idée plus générale de la façon dont le recensement a permis de dénombrer la population cible. En fait, ils incluent les personnes non dénombrées, qu'elles aient été prises en compte ou non dans le recensement au moyen des imputations. Le [tableau 12.1](#) présente les estimations du sous-dénombrement de la collecte de la population du Recensement de 2021 \hat{L} et \hat{T}_L ainsi que du sous-dénombrement de la population \hat{S} et \hat{T}_S (voir également le [tableau 1.3](#)). Il présente en outre l'estimation des erreurs-types. Les résultats sont présentés selon la province ou le territoire ainsi que pour l'ensemble du Canada, de même que selon le groupe d'âge et le genre. À l'échelle nationale, la collecte du Recensement de 2021 a donné lieu au dénombrement de 92,58 % de la population cible ($100 - \hat{T}_L$), alors que le chiffre officiel et publié du Recensement de 2021 représentait 95,02 % des personnes qui auraient dû être comprises dans ce chiffre ($100 - \hat{T}_S$). La différence entre ces deux taux est tout simplement attribuable à l'inclusion de l'imputation dans \hat{T}_L . Le Recensement de 2016 avait donné lieu au dénombrement de 93,63 % de la population cible, alors que 95,68 % des personnes qui auraient dû être incluses dans les chiffres officiels du Recensement de 2016 l'ont été. Ainsi, la collecte du Recensement de 2021 a dénombré un pourcentage légèrement inférieur à la population cible que le recensement précédent, alors que davantage d'imputations ont eu lieu qu'en 2016.

Tableau 12.1

Estimations du sous-dénombrement de la collecte de la population, du sous-dénombrement de la population et de l'erreur-type pour diverses caractéristiques, Recensement de 2021

Caractéristique	Sous-dénombrement de la collecte de la population				Sous-dénombrement de la population			
	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)
Canada	2 829 222	37 004	7,42	0,09	1 897 876	37 004	4,98	0,09
Province ou territoire								
Terre-Neuve-et-Labrador	39 055	1 775	7,41	0,31	26 428	1 775	5,02	0,32
Île-du-Prince-Édouard	13 177	977	8,17	0,56	10 190	977	6,32	0,57
Nouvelle-Écosse	70 323	3 373	7,05	0,31	47 196	3 373	4,73	0,32
Nouveau-Brunswick	50 361	2 655	6,38	0,31	30 165	2 655	3,82	0,32
Québec	410 770	15 756	4,81	0,18	225 381	15 756	2,64	0,18
Ontario	1 162 640	27 795	7,85	0,17	855 470	27 795	5,78	0,18
Manitoba	106 952	5 089	7,74	0,34	62 443	5 089	4,52	0,35
Saskatchewan	106 650	4 411	9,14	0,34	59 075	4 411	5,06	0,36
Alberta	357 305	12 225	8,09	0,25	231 249	12 225	5,24	0,26
Colombie-Britannique	488 394	15 827	9,37	0,28	338 678	15 827	6,50	0,28
Yukon	4 768	170	11,17	0,35	3 294	170	7,71	0,37
Territoires du Nord-Ouest	7 571	227	16,95	0,42	4 426	227	9,91	0,46
Nunavut	11 254	250	28,14	0,45	3 879	250	9,70	0,56

Rapport technique sur la couverture

Tableau 12.1

Estimations du sous-dénombrement de la collecte de la population, du sous-dénombrement de la population et de l'erreur-type pour diverses caractéristiques, Recensement de 2021

Caractéristique	Sous-dénombrement de la collecte de la population				Sous-dénombrement de la population			
	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)
Genre et groupe d'âge								
Total, genre, tous les âges	2 829 222	37 004	7,42	0,09	1 897 876	37 004	4,98	0,09
0 à 4 ans	143 827	5 809	7,58	0,28	98 438	5 809	5,19	0,29
5 à 14 ans	231 406	12 630	5,53	0,29	140 303	12 630	3,35	0,29
15 à 17 ans	65 929	8 516	5,45	0,67	41 091	8 516	3,40	0,68
18 à 19 ans	66 568	8 933	8,13	1,00	48 426	8 933	5,91	1,03
20 à 24 ans	375 100	18 381	15,43	0,64	313 912	18 381	12,91	0,66
25 à 34 ans	673 702	22 964	12,57	0,37	530 315	22 964	9,89	0,39
35 à 44 ans	402 817	19 332	7,85	0,35	281 659	19 332	5,49	0,36
45 à 54 ans	301 743	19 130	6,28	0,37	192 317	19 130	4,01	0,38
55 à 64 ans	299 714	18 459	5,65	0,33	164 211	18 459	3,10	0,34
65 à 74 ans	158 130	13 112	3,91	0,31	49 955	13 112	1,23	0,32
75 ans et plus	110 285	11 803	3,74	0,39	37 248	11 803	1,26	0,40
Hommes¹, tous les âges	1 606 471	34 194	8,46	0,16	1 140 736	34 194	6,01	0,17
0 à 4 ans	69 726	5 668	7,20	0,54	46 527	5 668	4,80	0,56
5 à 14 ans	123 836	8 684	5,75	0,38	76 859	8 684	3,57	0,39
15 à 17 ans	35 012	6 595	5,61	1,00	22 004	6 595	3,53	1,02
18 à 19 ans	33 034	5 962	7,85	1,31	23 968	5 962	5,70	1,34
20 à 24 ans	194 168	13 315	15,44	0,90	163 089	13 315	12,97	0,92
25 à 34 ans	403 628	18 903	14,66	0,59	330 497	18 903	12,00	0,60
35 à 44 ans	254 696	16 057	9,90	0,56	192 459	16 057	7,48	0,58
45 à 54 ans	171 964	15 044	7,24	0,59	116 139	15 044	4,89	0,60
55 à 64 ans	183 775	14 720	6,99	0,52	116 176	14 720	4,42	0,54
65 à 74 ans	89 756	10 291	4,59	0,50	37 494	10 291	1,92	0,52
75 ans et plus	46 876	7 677	3,68	0,58	15 524	7 677	1,22	0,60
Femmes², tous les âges	1 222 751	28 656	6,39	0,14	757 140	28 656	3,95	0,14
0 à 4 ans	74 101	5 609	7,99	0,56	51 911	5 609	5,60	0,57
5 à 14 ans	107 570	9 289	5,29	0,43	63 444	9 289	3,12	0,44
15 à 17 ans	30 918	5 677	5,28	0,92	19 088	5 677	3,26	0,94
18 à 19 ans	33 534	6 747	8,42	1,55	24 458	6 747	6,14	1,59
20 à 24 ans	180 931	14 260	15,42	1,03	150 822	14 260	12,86	1,06
25 à 34 ans	270 074	15 010	10,36	0,52	199 818	15 010	7,66	0,53
35 à 44 ans	148 120	10 306	5,80	0,38	89 199	10 306	3,49	0,39
45 à 54 ans	129 779	12 572	5,35	0,49	76 178	12 572	3,14	0,50

Tableau 12.1

Estimations du sous-dénombrement de la collecte de la population, du sous-dénombrement de la population et de l'erreur-type pour diverses caractéristiques, Recensement de 2021

Caractéristique	Sous-dénombrement de la collecte de la population				Sous-dénombrement de la population			
	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)
55 à 64 ans	115 940	11 524	4,34	0,41	48 036	11 524	1,80	0,42
65 à 74 ans	68 375	8 934	3,26	0,41	12 462	8 934	0,60	0,42
75 ans et plus	63 409	8 843	3,78	0,51	21 724	8 843	1,30	0,52

1. Le terme « hommes+ » comprend les hommes (et les garçons) et certaines personnes non binaires.

2. Le terme « femmes+ » comprend les femmes (et les filles) et certaines personnes non binaires.

Note : En raison de l'arrondissement, la somme des estimations de la couverture pourrait ne pas correspondre au total.

Sources : Statistique Canada, Étude sur le sous-dénombrement du Recensement de 2021.

12.2 Participation des réserves et des établissements

12.2.1 Introduction

En 2021, on comptait 1 026 réserves et établissements, dont 63 ont été jugés être partiellement dénombrés. Pour ces 63 réserves et établissements partiellement dénombrés, le dénombrement des logements n'a soit pas été permis, soit a été interrompu avant d'être terminé. Aucune donnée de 2021 n'existait pour les réserves et établissements partiellement dénombrés; par conséquent, ils n'ont pas été inclus dans les calculs. Dans le Recensement de 2016, 14 réserves et établissements avaient été déclarés partiellement dénombrés; 1 ayant pris part au Recensement de 2021. En 2011, 31 réserves et établissements ont été partiellement dénombrés. Cette importante augmentation du nombre de réserves et établissements partiellement dénombrés pour le Recensement de 2021 par rapport à celui de 2016 a principalement été due aux mesures sanitaires liées à la pandémie de COVID-19.

De plus amples renseignements sont disponibles à la page : [Réserves et établissements partiellement dénombrés](#).

Les estimations des réserves et des établissements partiellement dénombrés sont basées sur un modèle. On doit utiliser ces estimations avec circonspection étant donné qu'il n'existe aucune source fiable permettant de confirmer les hypothèses utilisées dans ce modèle.

12.2.2 Réserves et établissements partiellement dénombrés

Pour 63 réserves et établissements partiellement dénombrés, le Recensement de 2021 n'a pas été en mesure d'inclure des chiffres de population et les études de couverture n'ont pas pu estimer directement le sous-dénombrement net de la population, en raison de la disponibilité limitée des données du Recensement de 2021. Les chiffres et le sous-dénombrement net ont été estimés à l'aide d'autres méthodes.

Afin d'estimer les chiffres de population du recensement, on a utilisé une méthodologie fondée sur un modèle pour ces réserves et établissements partiellement dénombrés. Une régression linéaire sans ordonnée à l'origine a été choisie pour ce modèle d'estimation. Le paramètre du modèle a été estimé en fonction des réserves entièrement dénombrées dans le cadre des recensements de 2016 et de 2021. Ce modèle supposait que le chiffre du Recensement de 2021 était une fonction linéaire du chiffre du Recensement de 2016 pour toutes les provinces, avec estimations distinctes de la pente pour chaque province. En particulier, on a supposé que les réserves et établissements partiellement dénombrés suivaient la même tendance de croissance que les réserves et établissements entièrement dénombrés. Le modèle a été évalué pour les hypothèses de régression de base que sont l'indépendance des erreurs, l'homogénéité des variances et la normalité des erreurs. Pour chacun des 63 réserves et établissements partiellement dénombrés, la variable d'entrée pour le modèle de régression était soit le chiffre réel du Recensement de 2016 soit le chiffre estimé du recensement selon le modèle de 2016. Le résultat du modèle était le chiffre estimé du Recensement de 2021 pour ces 63 communautés. Les estimations

Rapport technique sur la couverture

obtenues doivent être utilisées avec prudence, puisqu'elles reposent entièrement sur un modèle dont les hypothèses ne peuvent être vérifiées. La validité de ces estimations fondées sur un modèle dépend de la mesure dans laquelle les hypothèses du modèle peuvent saisir la véritable situation sous-jacente.

Au Recensement de 2016, 14 réserves, comptant environ 28 000 personnes, ont été classées comme partiellement dénombrées. Parmi les 63 réserves et établissements utilisés dans les estimations fondées sur un modèle au Recensement de 2021, 50 ont été jugés entièrement dénombrés au Recensement de 2016, alors que les 13 autres ont été classés comme partiellement dénombrés. La population totale des 63 réserves et établissements partiellement dénombrés a été estimée à 58 480 personnes, c'est-à-dire plus du double de celle de 2016.

Les chiffres de population estimés par ce modèle pour les 63 réserves et établissements partiellement dénombrés sont sujets à des erreurs de couverture, de la même façon que les chiffres de population du recensement pour le reste du pays. Le sous-dénombrement net de ces 63 lieux a été estimé en calculant le taux de sous-dénombrement net pour toutes les réserves entièrement dénombrées dans le pays. Ce taux a ensuite été appliqué aux chiffres de population estimés de l'ensemble des réserves et établissements partiellement dénombrés. La population totale estimée de 58 480 pour les réserves et établissements partiellement dénombrés en 2021 comprend cet ajustement visant à tenir compte des erreurs de couverture.

Les estimations de la population des réserves et établissements partiellement dénombrés ne sont pas incluses dans les estimations de sous-dénombrement, de sous-dénombrement de collecte, de surdénombrement et de sous-dénombrement net présentées dans ce rapport, car elles sont établies au moyen d'un modèle et non d'études de couverture du recensement. De plus, elles ne fournissent pas le niveau de détails (p. ex. estimations selon la langue maternelle ou l'état matrimonial) présenté pour les autres estimations.

Les estimations de la population des réserves et des établissements partiellement dénombrés sont incluses dans les estimations démographiques. Ces estimations, réparties par province et territoire, figurent au tableau 1 de la section « Qualité des données, concepts et méthodologie » de la publication des estimations démographiques suivante : [Estimations démographiques annuelles : Canada, provinces et territoires \(Population totale seulement\), 2023](#), produit n° 91-215-X au catalogue de Statistique Canada.

Annexe A — Qui inscrire dans le questionnaire du recensement

Les instructions suivantes étaient fournies à la page 3 des questionnaires abrégé et détaillé du Recensement de 2021 pour permettre de déterminer les personnes à inscrire dans le questionnaire.

1. Qui inscrire à l'étape B

- Toutes les personnes qui ont leur **résidence principale** à cette adresse le 11 mai 2021, y compris les nouveau-nés, les colocataires et les personnes temporairement absentes;
- Les **citoyens canadiens**, les **immigrants reçus** (résidents permanents), les personnes qui demandent **le statut de réfugié** (demandeurs d'asile), les personnes d'un **autre pays** ayant un **permis de travail** ou **d'études** et les membres de leur famille qui habitent ici avec elles;
- Les personnes qui demeurent temporairement à cette adresse le 11 mai 2021 et **qui n'ont pas de résidence principale ailleurs**.

2. Où inscrire les personnes ayant plus d'une résidence

- Les **enfants en garde partagée** doivent être inscrits au domicile du parent où ils vivent la plupart du temps. Les enfants qui passent autant de temps avec chaque parent doivent être inscrits au domicile du parent où ils se trouvent le 11 mai 2021.
- Les **étudiants** qui retournent vivre chez leurs parents durant l'année doivent être inscrits à l'adresse de leurs parents, même s'ils résident ailleurs pendant leurs études ou leur emploi d'été.
- Les **conjointes temporairement absents** qui demeurent ailleurs en raison de leur travail ou de leurs études doivent être inscrits à la résidence principale de leur famille, s'ils y retournent périodiquement.
- Les **personnes en établissement institutionnel** depuis **moins de six mois** (par exemple, dans un foyer pour personnes âgées, un hôpital ou une prison) doivent être inscrites à leur résidence habituelle.

Bibliographie

Abeysondera, M., Zhao, Z. et Farr, H. 2023. *Rapport méthodologique de l'Étude sur le surdénombrement du Recensement (ESuR) de 2021*. Statistique Canada.

Dasylyva, A., Abeysondera, M., Haddou, M. et Lachance, M. 2015. *Sampling and Estimation Based on a Frame of Connected Record-Groups Methodology* (Disponible en anglais seulement). Statistique Canada.

Dolson, D. 2010. « Études de la couverture du recensement au Canada : un historique mettant l'accent sur le Recensement de 2011. » *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

Fellegi, I.P. et Sunter, A.B. 1969. « A Theory for Record Linkage » *Journal of the American Statistical Association* (Disponible en anglais seulement). vol. 64, p. 1183 à 1210.

Hong, R. (2023). *L'Enquête sur la classification des logements de 2021*. Statistique Canada.

Lavallée, P. 2007. *Indirect Sampling* (Disponible en anglais seulement). Springer, New York.

Parenteau, M. 2023. « Processus de traitement de l'ESoR 2021. » *Rencontre du Comité fédéral-provincial-territorial de la démographie*, 31 janvier au 2 février 2023, Statistique Canada, Ottawa, Ontario, 38 diapositives. (Réunion en anglais du 24 au 26 janvier 2023).