Guide analytique – Portrait de la société canadienne : Les expériences pendant la pandémie

1.0 Description

La série d'enquêtes intitulée « Portrait de la société canadienne (PSC) » est une nouvelle initiative de Statistique Canada. Au moyen d'un panel en ligne fondé sur une approche probabiliste, cette initiative consiste à mener quatre courtes enquêtes en ligne auprès du même groupe de participants sur une période d'un an. Il s'agit, pour l'instant, d'un projet expérimental qui s'inscrit dans un effort plus vaste visant à moderniser nos méthodes et activités de collecte de données. L'objectif est de recueillir des renseignements importants sur la société canadienne de façon plus efficace, plus rapide et moins coûteuse comparativement aux méthodes d'enquête traditionnelles. Nous serons en mesure de mettre à l'essai cette méthode de collecte de données et de la peaufiner au fil du temps.

La nature expérimentale de ce projet et son degré élevé de non-réponse ont une incidence sur les estimations pouvant être produites à l'aide du panel en ligne. Les poids d'enquête ont été ajustés pour réduire le biais potentiel qui pourrait survenir en raison de la non-réponse. Des ajustements de non-réponse et de calage utilisant l'information auxiliaire disponible ont été appliqués et sont reflétés dans les poids d'enquête fournis avec le fichier de données. Malgré ces ajustements, le degré élevé de non-réponse au panel augmente le risque de biais résiduel, ce qui peut avoir une incidence sur les estimations produites au moyen des données par panel. À la section 5, vous trouverez de plus amples renseignements sur les méthodes de pondération utilisées pour ajuster la non-réponse. Les lignes directrices et les considérations concernant la qualité des données sont présentées à la section 6.

Chaque enquête de la série est menée auprès d'un sous-échantillon de répondants à l'Enquête sociale générale — Identité sociale (ESG-IS) ayant accepté de participer à des enquêtes supplémentaires.

Du 19 juillet au 1er août 2021, Statistique Canada a réalisé le Portrait de la société canadienne — Les expériences pendant la pandémie (PSC-EP). Il s'agit de la deuxième vague d'enquête sur le PSC.

L'objectif de cette enquête est de nous aider à mieux comprendre divers aspects de la vie des Canadiens pendant la pandémie, dont l'accès aux services de santé, les perceptions à l'égard de la sécurité et la communauté. Le PSC est conçu pour produire des données à l'échelle nationale.

Le présent manuel a été rédigé dans le but de faciliter la manipulation du fichier de microdonnées qui contient les résultats de l'enquête du PSC-EP.

Toutes les questions concernant l'ensemble de données ou son utilisation doivent être acheminées à :

Statistique Canada

Service à la clientèle
Centre de l'intégration et du développement des données sociales
Téléphone : 613-951-3321 ou composez sans frais le 1-800-461-9050
Télécopieur : 613-951-4527
Courriel : csdid-info-cidds@canada.ca

2.0 Méthodologie de l'enquête

2.1 Population cible et population observée

Le PSC-EP consiste en une enquête transversale par échantillon. Chaque enquête de la série est menée auprès d'un sous-échantillon de répondants à l'Enquête sociale générale — Identité sociale (ESG-IS) ayant accepté de participer à des enquêtes supplémentaires.

La population cible du Portrait de la société canadienne (PSC) est la même que celle de l'ESG-IS. La population cible comprend toutes les personnes de 15 ans et plus au Canada, sauf :

1. les résidents du Yukon, des Territoires du Nord-Ouest et du Nunavut;
2. les résidents à temps plein d'établissements institutionnels;
3. les résidents des réserves autochtones.

La base de sondage utilisée pour l'ESG-IS et la stratégie d'échantillonnage sont décrites à la section 5 du *Guide de l'utilisateur de l'ESG-IS* de 2020.

2.2 Plan d'échantillonnage et taille de l'échantillon

Afin d'établir l'échantillon pour le Portrait de la société canadienne (PSC), des questions de recrutement ont été ajoutées à la fin de l'Enquête sociale générale — Identité sociale (ESG-IS). Environ 22 % des répondants à l'ESG-IS ont accepté d'être sollicités pour participer à de futures enquêtes. Ils ont formé l'échantillon pour le PSC.

Le tableau ci-dessous rend compte du nombre de répondants pour chaque degré du plan d'échantillonnage du PSC-EP.

Degrés de l'échantillon	Nombre
Logements sélectionnés pour l'ESG-IS	86 804
Personnes ayant répondu à l'ESG-IS	34 044
Personnes ayant accepté d'être sollicitées pour participer à de futures enquêtes	7 502
Échantillon brut des enquêtes du PSC	7 502
Panélistes ayant participé au PSC-EP	3 330

Le tableau ci-dessous présente le nombre de répondants pour le PSC-EP selon la région, le groupe d'âge et le sexe.

Région	Domaine	Nombre
Région géographique	Canada	3 330
	Provinces de l'Atlantique	502
	Québec	635
	Ontario	1 125
	Prairies	645
	Colombie-Britannique	423

Groupe d'âge	Tous	3 330
	15 à 24 ans	131
	25 à 34 ans	459
	35 à 44 ans	699
	45 à 54 ans	671
	55 à 64 ans	621
	65 à 74 ans	558
	75 ans et plus	191
Sexe	Tous	3 330
	Hommes	1 720
	Femmes	1 610

3.0 Collecte des données

Le recrutement pour le PSC s'est fait en ajoutant deux questions de recrutement à la fin du questionnaire de l'ESG-IS. Celle-ci a été menée du 17 août 2020 au 8 février 2021. La première question servait à déterminer si les répondants souhaitaient participer à une série de courtes enquêtes, d'une durée de 15 à 20 minutes, sur des enjeux sociaux importants. On a demandé aux personnes ayant répondu « oui » à cette question de fournir leur adresse électronique et leur numéro de téléphone cellulaire. Ce sous-échantillon de l'ESG-IS a constitué l'échantillon du PSC.

PSC-EP — Les expériences pendant la pandémie

Tous les répondants de l'ESG-IS ayant répondu « oui » aux questions de recrutement ont reçu une invitation par courriel qui comprenait un lien vers le PSC-EP et un code d'accès sécurisé pour remplir l'enquête en ligne. La collecte de l'enquête a commencé le 19 juillet 2021.

Deux stratégies de collecte ont été suivies pour le PSC-EP. L'échantillon a été divisé en deux groupes, et chacun a fait l'objet de l'une des deux stratégies. Tous les répondants potentiels ont reçu le courriel d'invitation le 19 juillet. Pour ce qui est du premier groupe, les courriels de rappel ont été envoyés les 20, 22 et 24 juillet, et pour le deuxième groupe, les 22, 26 et 29 juillet. L'application est demeurée accessible jusqu'au 1er août 2021.

Couplage d'enregistrements

Afin d'améliorer la qualité des données provenant du PSC-EP et de réduire le fardeau de réponse, Statistique Canada a combiné les renseignements fournis par les répondants aux renseignements provenant de l'Enquête sociale générale — Identité sociale. L'ESG-IS est la source des variables sociodémographiques disponibles pour le PSC-EP.

3.1 Contrôle de la divulgation

La loi interdit à Statistique Canada de rendre publique toute donnée susceptible de révéler de l'information obtenue en vertu de la *Loi sur la statistique* et se rapportant à toute personne, entreprise ou organisation reconnaissable sans que cette personne, entreprise ou organisation le sache ou y consente par écrit. Diverses règles de confidentialité s'appliquent à toutes les données diffusées ou publiées afin d'empêcher la publication ou la divulgation de toute information jugée confidentielle. Au besoin, des données sont supprimées pour empêcher la divulgation directe ou par recoupement de données reconnaissables.

4.0 Qualité des données

Les erreurs d'enquête proviennent de diverses sources. Ces erreurs peuvent être classées en deux principales catégories : erreurs non dues à l'échantillonnage et erreurs d'échantillonnage.

4.1 Erreurs non dues à l'échantillonnage

Les erreurs non dues à l'échantillonnage peuvent se définir comme des erreurs qui se produisent pendant presque toutes les activités d'enquête, mis à part l'échantillonnage. Ces erreurs sont présentes à la fois dans les enquêtes-échantillons et les recensements (contrairement à l'erreur d'échantillonnage, qui est présente seulement dans les enquêtes-échantillons). Les principales sources d'erreurs non dues à l'échantillonnage sont les suivantes : la non-réponse, la couverture, la mesure et le traitement.

4.1.1 Non-réponse

La non-réponse est à la fois une source d'erreur d'échantillonnage et d'erreur non due à l'échantillonnage. Les résultats de non-réponse sont issus d'une incapacité à recueillir les renseignements exhaustifs auprès de toutes les unités dans l'échantillon sélectionné. La non-réponse est une source d'erreur non due à l'échantillonnage dans le sens où les non-répondants ont souvent des caractéristiques différentes de celles des répondants, ce qui peut entraîner des estimations d'enquête biaisées si le biais de non-réponse n'est pas entièrement éliminé par des corrections de la pondération. Plus le taux de réponse est faible, plus le risque de biais est élevé. La non-réponse est aussi une source d'erreur d'échantillonnage; il en est question à la section 6.2.

Le plan d'enquête du PSC-EP s'effectue en plusieurs étapes, chacune donnant lieu à de la non-réponse. Le tableau ci-dessous résume le taux de réponse à chacune de ces étapes et le taux de réponse cumulatif qui en résulte pour le PSC-EP.

Étape de l'enquête	Nombre de répondants	Taux de réponse
ESG-IS	34 044	40,3 %
Participation volontaire des répondants de l'ESG-IS à des enquêtes supplémentaires	7 502	22,0 %
Réponse des participants du panel au PSC-EP	3 330	44,4%
Taux de réponse cumulatif		3,8 %

4.1.2 Erreurs de couverture

Les erreurs de couverture comprennent les omissions, les inclusions erronées, les dédoublements et les erreurs de classification des unités figurant dans la base de sondage. Comme elles ont une incidence sur chaque estimation de l'enquête, elles constituent l'un des plus importants types d'erreurs. Les erreurs de couverture peuvent entraîner un biais dans les estimations, et leur incidence peut varier d'un sous-groupe de la population à l'autre. Il s'agit d'un type d'erreur très difficile à mesurer ou à quantifier avec exactitude.

Les données du PSC-EP sont recueillies auprès de personnes de 15 ans et plus vivant dans un logement privé dans l'une des 10 provinces. La couverture de l'enquête exclut les résidents du Yukon, des Territoires du Nord-Ouest et du Nunavut; les résidents à temps plein d'établissements institutionnels, et les résidents des réserves autochtones. Ces groupes exclus représentent ensemble moins de 2 % de la population canadienne âgée de 15 ans et plus.

Comme le PSC-EP est fondé sur l'échantillon de l'ESG-IS et que les données ont été recueillies du 19 juillet au 1er août 2021, il y a un sous-dénombrement dans les 10 provinces des résidents qui ont 15 ans depuis le 17 août 2020, ce qui correspond au début de la période de collecte de l'ESG-IS. Il y a également un sous-dénombrement des personnes qui n'ont pas accès à Internet, puisque les données du PSC-EP ont été recueillies entièrement en ligne. Le sous-dénombrement est plus important chez les personnes de 65 ans et plus.

4.1.3 Erreurs de mesure

Les erreurs de mesure (parfois appelées les erreurs de réponse) se produisent lorsque la réponse fournie diffère de la valeur réelle; de telles erreurs peuvent être attribuables au répondant, au questionnaire, à la méthode de collecte des données ou au système de tenue des dossiers des répondants. Ces erreurs peuvent être aléatoires ou entraîner un biais systématique si elles ne sont pas aléatoires.

4.1.4 Erreurs de traitement

Les erreurs de traitement désignent des erreurs associées aux activités menées une fois que les réponses ont été reçues. Elles comprennent toutes les activités de manipulation des données après la collecte et avant l'estimation. Comme pour toute autre erreur, elles peuvent être aléatoires et entraîner une hausse de la variance des estimations de l'enquête, ou elles peuvent être systématiques et introduire un biais. Il est difficile d'obtenir des mesures directes des erreurs de traitement ainsi que de leur incidence sur la qualité des données, surtout parce qu'on les confond avec d'autres types d'erreurs (non-réponse, erreur de mesure et erreur de couverture).

4.2 Erreurs d'échantillonnage

L'erreur d'échantillonnage est définie comme étant l'erreur résultant de l'estimation d'une caractéristique de la population en mesurant une partie au lieu de toute la population. Des méthodes de calcul de l'erreur d'échantillonnage s'appliquent à l'enquête-échantillon probabiliste. Ces méthodes découlent directement du plan d'échantillonnage et de la méthode d'estimation de l'enquête.

La mesure utilisée le plus souvent pour quantifier l'erreur d'échantillonnage est la variance d'échantillonnage. La variance d'échantillonnage mesure à quel point l'estimation d'une caractéristique de divers échantillons possibles de même taille et de même conception est différente l'une de l'autre. Dans le cas des plans d'échantillonnage qui utilisent l'échantillonnage probabiliste, l'ampleur de la variance d'échantillonnage d'une estimation peut être estimée.

Les éléments qui ont des répercussions sur l'ampleur de la variance d'échantillonnage comprennent ce qui suit :

1. La variabilité de la caractéristique d'intérêt dans la population : plus la caractéristique dans la population est variable, plus la variance d'échantillonnage est grande.
2. La taille de la population : en général, la taille de la population a des répercussions sur la variance d'échantillonnage seulement pour les populations de petite taille ou de taille moyenne.
3. Le taux de réponse : la variance d'échantillonnage augmente dans la mesure où la taille de l'échantillon diminue. Étant donné que les non-répondants diminuent en fait la taille de l'échantillon, les non-réponses augmentent la variance d'échantillonnage.
4. Le plan d'échantillonnage et les méthodes d'estimation : certains plans d'échantillonnage sont plus efficients que d'autres parce que, pour la même taille d'échantillon et la même méthode d'estimation, un plan peut mener à une variance d'échantillonnage moindre que l'autre.

L'erreur-type d'un estimateur est la racine carrée de sa variance d'échantillonnage. Cette mesure donne une indication de l'erreur d'échantillonnage à l'aide de la même échelle que l'estimation, tandis que la variance est basée sur les différences quadratiques.

Le coefficient de variation (CV) d'une estimation est une mesure relative de l'erreur d'échantillonnage. Il est défini comme l'estimation de l'erreur-type divisée par l'estimation elle-même. Il est très utile pour mesurer et comparer l'erreur d'échantillonnage de variables quantitatives avec de grandes valeurs positives. Cependant, il n'est pas recommandé pour des estimations telles que les proportions, les estimations de changements ou de différences, et les variables qui peuvent avoir des valeurs négatives.

Il est considéré comme une pratique exemplaire à Statistique Canada de faire état de l'erreur d'échantillonnage d'une estimation par l'entremise de son intervalle de confiance de 95 %. L'intervalle de confiance de 95 % d'une estimation signifie que si l'enquête était répétée à maintes reprises, 95 % du temps (ou 19 fois sur 20), l'intervalle de confiance couvrirait la véritable valeur de la population.

5.0 Pondération

Le principe qui sous-tend l'estimation dans un échantillon probabiliste repose sur le fait que chaque unité sélectionnée dans l'échantillon représente, en plus d'elle-même, d'autres unités non sélectionnées dans l'échantillon. Par exemple, si un échantillon aléatoire simple de 100 unités est sélectionné d'une population de 5 000 unités, chaque unité de l'échantillon représente 50 unités de la population. Le nombre d'unités que représente une unité de l'échantillon est le poids d'enquête de l'unité échantillonnée.

L'étape de la pondération consiste à calculer un poids d'échantillonnage associé à chaque personne. Ce poids, qui figure dans le fichier de microdonnées, doit être utilisé afin d'obtenir des estimations représentatives de la population cible à partir de l'enquête. Par exemple, si l'on doit estimer le nombre de personnes qui fument quotidiennement, on le fait en sélectionnant les enregistrements des personnes de l'échantillon qui présentent cette caractéristique et en faisant la somme des poids de ces enregistrements. L'étape de la pondération consiste à calculer ce nombre (soit le poids) pour chaque enregistrement. La présente section contient des renseignements sur la méthode utilisée pour calculer les poids d'échantillonnage du PSC-EP.

La pondération de l'échantillon pour le PSC-EP comporte plusieurs étapes pour refléter les étapes de l'échantillonnage, de la participation et de la réponse afin d'obtenir l'ensemble final de répondants. Les sections suivantes portent sur les étapes de pondération afin de déterminer les poids de l'enquête pour le PSC-EP.

5.1 Poids de sondage

Les poids initiaux du panel correspondent aux poids calés finaux de l'ESG-IS. Il s'agit des poids de sondage de l'ESG-IS ajustés selon les unités hors champ et la non-réponse à l'ESG-IS, puis calés sur les totaux de contrôle de la population. Pour en savoir plus sur ces poids, consultez la section 8.1 du *Guide de l'utilisateur de l'ESG-IS.*

5.2 Ajustement pour la non-réponse et la non-participation

Pendant la collecte des données du PSC-EP, les réponses sont obtenues uniquement à partir d'une proportion d'unités échantillonnées. Les personnes qui ont répondu à l'ESG-IS peuvent décider de ne pas participer volontairement à des enquêtes supplémentaires et, par conséquent, au panel. De plus, certaines personnes ayant choisi de participer au panel ne répondent pas pendant la collecte des données du PSC-EP. Les poids des unités de non-réponse ou de non-participation ont été redistribués aux unités participantes. Les poids des unités n'ayant pas participé au panel ont été redistribués aux unités participantes ayant des caractéristiques semblables dans les groupes de réponses homogènes (GRH).

Les variables possibles pour la création des GRH étaient disponibles pour les unités déclarantes et non déclarantes. Ces variables comprennent notamment des caractéristiques personnelles (âge, sexe, niveau de scolarité, groupe de population, orientation sexuelle, renseignements sur l'emploi, habitudes de vote et revenu personnel) et des caractéristiques du ménage (accession à la propriété et revenu du ménage), ainsi que des variables liées à la collecte dans le cadre de l'ESG-IS (comme le mois de la réponse à l'ESG et si la réponse a été fournie en ligne ou soumise avec l'aide d'un intervieweur). Un facteur d'ajustement a été calculé dans chacun des groupes de réponses de la façon suivante :

$$\frac{\text{Somme des poids des répondants et des non-répondants}}{\text{Somme des poids des répondants}}$$

Les poids des répondants ont été multipliés par ce facteur pour produire les poids ajustés pour la non-réponse. Les unités non participantes ont été retirées du processus de pondération à cette étape.

5.3 Calage

Les totaux de contrôle ont été calculés en utilisant les données de projection démographique de l'EPA. Durant le calage, un facteur d'ajustement est calculé et appliqué aux poids d'enquête. Cet ajustement est calculé afin que les sommes pondérées correspondent aux totaux de contrôle. Trois ensembles de totaux de contrôle de la population ont été utilisés pour le PSC-EP:

1) Région géographique, groupe d'âge et sexe : Les groupes de région géographique et d'âge sélectionnés pour le calage ont tenu compte du nombre parfois faible de répondants dans différentes catégories. Les cinq régions géographiques utilisées pour le calage étaient les provinces de l'Atlantique, le Québec, l'Ontario, les provinces des Prairies et la Colombie-Britannique. Les groupes d'âge utilisés étaient les 15 à 34 ans, les 35 à 64 ans, les 65 ans et plus.

2) Géographies infrarégionales : Les poids des répondants ont également été calés de sorte que la somme dans chaque province, ainsi que dans les RMR de Montréal, de

Toronto et de Vancouver, corresponde au contrôle de la population dans ces géographies infrarégionales.

3) Groupe d'âge à l'échelle nationale : Les poids des répondants ont été calés sur les totaux de population (à l'échelle nationale) dans des groupes d'âge plus détaillés. Ces groupes ont été définis comme les 15 à 24 ans, les 25 à 34 ans et ainsi de suite jusqu'aux répondants âgés de 75 ans et plus.

5.4 Poids bootstrap

Les poids bootstrap ont été produits pour les répondants à l'enquête du PSC-EP. Chaque échantillonnage bootstrap a été produit en fonction des poids de sondage initiaux du PSC-EP, puis ajusté pour la non-réponse et calé comme il a été décrit précédemment.

6.0 Lignes directrices pour la totalisation, l'analyse et la diffusion

Le présent chapitre de la documentation décrit les lignes directrices que doivent respecter les utilisateurs qui totalisent, analysent, publient ou diffusent de toute autre façon des données calculées à partir des fichiers de microdonnées de l'enquête. À l'aide de ces lignes directrices, les utilisateurs de microdonnées devraient être en mesure de produire les mêmes chiffres que ceux que produit Statistique Canada, tout en élaborant des chiffres actuellement non publiés conformément à ces directives établies.

6.1 Lignes directrices pour l'arrondissement

Nous conseillons vivement aux utilisateurs de respecter les lignes directrices qui suivent pour l'arrondissement des estimations calculées à partir des fichiers de microdonnées :

a) Les estimations dans le corps principal d'un tableau statistique doivent être arrondies à l'aide de la technique d'arrondissement normale. Selon cette technique, si le premier ou le seul chiffre à supprimer se situe entre 0 et 4, le dernier chiffre à conserver ne change pas. Si le premier ou le seul chiffre à supprimer se situe entre 5 et 9, le dernier chiffre à conserver est augmenté de 1.

b) Les totaux partiels marginaux et des totaux marginaux des tableaux statistiques doivent être calculés à partir de leurs composantes non arrondies correspondantes, puis ensuite être arrondis à leur tour à l'aide de la technique d'arrondissement normale. Les moyennes, les taux, les pourcentages, les proportions et les rapports doivent être calculés à partir de composantes non arrondies (c'est-à-dire numérateurs et/ou dénominateurs), puis être arrondis à leur tour à l'aide de la technique d'arrondissement normale. Les sommes et les différences doivent être calculées à partir de leurs composantes non arrondies correspondantes, puis être arrondies à leur tour à l'aide de la technique d'arrondissement normale.

c) Dans les cas où, en raison de limitations d'ordre techniques ou de toutes autres limites, une technique d'arrondissement autre que la technique normale est utilisée et produit des estimations à être publiées ou autrement diffusées différentes des estimations correspondantes publiées par Statistique Canada, nous conseillons vivement aux utilisateurs d'indiquer la raison de ces différences dans le ou les documents à publier ou à diffuser.

d) En aucun cas, les utilisateurs ne doivent publier ou autrement diffuser des estimations non arrondies. Des estimations non arrondies laissent entendre qu'elles sont plus précises qu'elles le sont en réalité.

6.2 Lignes directrices pour la pondération de l'échantillon en vue de la totalisation

Le PSC-EP utilise un plan d'échantillonnage et une méthode d'estimation complexes, ce qui implique que les poids d'enquête ne sont pas égaux pour toutes les unités de l'échantillon. Lorsque des estimations et tableaux statistiques sont produits, les utilisateurs doivent appliquer le poids d'enquête approprié. Si les poids appropriés ne sont pas utilisés, les estimations calculées à partir des fichiers de microdonnées ne peuvent être considérées comme représentatives de la population visée par l'enquête et ne correspondront pas à celles produites par Statistique Canada.

6.3 Lignes directrices pour la diffusion de la qualité

Avant de diffuser et/ou de publier toute estimation, les analystes doivent prendre en considération le niveau de qualité de l'estimation. Étant donné la nature expérimentale du PSC-EP et le haut degré de non-réponse, toutes les estimations produites à partir du panel devraient être accompagnées d'un avertissement de qualité indiquant d'utiliser les estimations avec prudence.

Malgré que les erreurs d'échantillonnage et non dues à l'échantillonnage influencent la qualité des données, la présente section porte sur la qualité en ce qui a trait aux erreurs d'échantillonnage. À Statistique Canada, il est considéré comme une pratique exemplaire de faire état de l'erreur d'échantillonnage d'une estimation par l'intermédiaire de son intervalle de confiance (IC) de 95 %. L'intervalle de confiance doit être publié avec l'estimation, dans le même tableau que celle-ci. En plus des intervalles de confiance, les estimations du PSC-EP sont classées dans l'une des deux catégories de diffusion suivantes :

Catégorie E
Les estimations et les intervalles de confiance doivent être signalés par la lettre E (ou un quelconque identificateur similaire) et accompagnés d'un avertissement sur la qualité d'utiliser les estimations avec prudence. Les utilisateurs de données devraient utiliser l'intervalle de confiance de 95% afin de déterminer si la qualité de l'estimation est suffisante.

Catégorie F
Les estimations et les intervalles de confiance ne sont pas recommandés pour la diffusion. Ils sont jugés de si piètre qualité, qu'ils ne se portent à aucune utilisation; ils sont très instables, ce qui les rend peu fiables et potentiellement trompeurs. Si les analystes insistent pour publier des estimations de piètre qualité, même après avoir été informés de leur exactitude, ces estimations doivent être accompagnées d'un avis de non-responsabilité. Les analystes doivent tenir compte des mises en garde reçues et s'engager à ne pas diffuser, présenter, ni rapporter les estimations, directement ou indirectement, sans cet avis de non-responsabilité. Ces estimations doivent être signalées par la lettre F (ou un quelconque identificateur semblable) et la mise en garde suivante doit accompagner les estimations et intervalles de confiance : « Nous informons l'utilisateur que ces estimations et intervalles de confiance (désignés par la lettre F) ne respectent pas les normes de qualité de Statistique Canada. Les conclusions fondées sur ces données ne seront pas fiables et peuvent être invalides.»

Les règles permettant d'assigner une estimation à une catégorie de diffusion dépendent du type d'estimation.

Règles de diffusion pour les estimations de proportions et de comptes
Les estimations de proportions et de comptes sont dérivées à partir de variables binaires. Les estimations de comptes sont des estimations du nombre total de personnes/ménages avec une caractéristique d'intérêt; en d'autres termes, elles représentent une somme pondérée d'une variable binaire (p. ex. nombre estimé d'immigrants). Les estimations de proportions sont des

estimations de proportions de personnes /ménages avec une caractéristique d'intérêt (p. ex. proportion estimée d'immigrants dans la population générale). Les estimations de comptes et de proportions peuvent aussi être dérivées à partir de variables catégorielles: c'est-à-dire l'estimation du nombre ou de la proportion des personnes/ménages qui appartiennent à une catégorie.

Les règles de diffusion pour les estimations de proportions et de chiffres sont fondées sur la taille de l'échantillon. Le tableau 1 présente les règles de diffusion du Portrait de la société canadienne – Perceptions de la vie (PSC-EP) pendant la pandémie, applicables à toutes les estimations de proportions et de chiffres, à l'exception des estimations de minorités visibles.

Tableau 1 : Règles générales pour les estimations de proportions et de chiffres, à l'exception des estimations de minorités visibles

Taille de l'échantillon (n)	Catégorie de diffusion	Mesure à prendre
n ≥ 200	E	Diffuser avec une mise en garde sur la qualité; les utilisateurs devraient utiliser l'IC comme indicateur de qualité
n < 200	F	Supprimer l'estimation et son IC pour des raisons de qualité

Pour les estimations de proportions, « n » est défini comme le nombre non pondéré de répondants dans le dénominateur (et non le numérateur) de la proportion. Pour les estimations de chiffres, « n » est défini comme le nombre non pondéré de répondants dont les valeurs sont différentes de zéro et qui contribuent à l'estimation.

Règles spéciales pour les estimations de minorités visibles

Le tableau 2 présente les règles de diffusion spéciales à suivre lorsque des estimations sont produites pour un groupe de minorités visibles (c.-à-d. en utilisant les catégories VISMIN ou VISMINFL). Des règles spéciales sont nécessaires en raison du plan d'échantillonnage de l'ESG-IS qui comprenait une surreprésentation de certains groupes de minorités visibles dans l'échantillon.

Tableau 2 : Règles spéciales pour les estimations de proportions et de chiffres des minorités visibles

Taille de l'échantillon (n)	Catégorie de diffusion	Mesure à prendre
n ≥ 350	E	Diffuser avec une mise en garde sur la qualité; les utilisateurs devraient utiliser l'IC comme indicateur de qualité
n < 350	F	Supprimer l'estimation et son IC pour des raisons de qualité

Compte tenu du nombre de répondants au PSC-EP, ces règles signifient que les groupes individuels de minorités visibles ne peuvent pas être utilisés comme domaines d'analyse selon le PSC-EP, mais que l'analyse par la catégorie VISMINFL est autorisée. D'autre part, étant donné que les expériences des différents groupes de minorités visibles peuvent être très différentes les unes des autres, il pourrait ne pas être convenable de produire une estimation pour tous les groupes de minorités visibles regroupés (VISMINFL = 1). Il est donc recommandé, même si ces estimations ne doivent pas être diffusées, de comparer les estimations des catégories VISMIN

plus désagrégées entre elles avant de décider de regrouper tous les groupes de minorités visibles.

Règles de diffusion pour les moyennes et les totaux de variables quantitatives

Les règles de diffusion pour les estimations de moyennes et de totaux de variables quantitatives ou de quantités sont fondées sur la taille de l'échantillon et sur le CV de l'estimation. Le tableau 3 présente les règles de diffusion du PSC-EP, à l'exception des estimations de minorités visibles.

Tableau 3 : Règles générales pour les moyennes et les totaux

Taille de l'échantillon (n)	Catégorie de diffusion	Mesure à prendre
n ≥ 200 et CV ≤ 50 %	E	Diffuser avec une mise en garde sur la qualité; les utilisateurs devraient utiliser l'IC comme indicateur de qualité
n < 200 ou CV > 50 %	F	Supprimer l'estimation et son IC pour des raisons de qualité

Pour les estimations de moyennes, « n » est défini comme le nombre non pondéré de répondants qui contribuent à l'estimation, y compris les répondants avec les valeurs de zéro. Pour les estimations de totaux, « n » est défini comme le nombre non pondéré de répondants dont les valeurs sont différentes de zéro et qui contribuent à l'estimation.

Règles spéciales pour les estimations de minorités visibles

Le tableau 4 présente les règles de diffusion spéciales à suivre lorsque des estimations sont produites pour un groupe de minorités visibles (c.-à-d. en utilisant les catégories VISMIN ou VISMINFL). Des règles spéciales sont nécessaires en raison du plan d'échantillonnage de l'ESG-IS qui comprenait une surreprésentation de certains groupes de minorités visibles dans l'échantillon.

Tableau 4 : Règles spéciales pour les estimations de moyennes et de totaux des minorités visibles

Taille de l'échantillon (n)	Catégorie de diffusion	Mesure à prendre
n ≥ 350 et CV ≤ 50 %	E	Diffuser avec une mise en garde sur la qualité; les utilisateurs devraient utiliser l'IC comme indicateur de qualité
n < 350 ou CV > 50 %	F	Supprimer l'estimation et son IC pour des raisons de qualité

Compte tenu du nombre de répondants au PSC-EP, ces règles signifient que les groupes individuels de minorités visibles ne peuvent pas être utilisés comme domaines d'analyse selon le PSC-EP, mais que l'analyse par la catégorie VISMINFL est autorisée. D'autre part, étant donné que les expériences des différents groupes de minorités visibles peuvent être très différentes les unes des autres, il pourrait ne pas être convenable de produire une estimation pour tous les groupes de minorités visibles regroupés (VISMINFL = 1). Il est donc recommandé, même si ces estimations ne doivent pas être diffusées, de comparer les estimations des catégories VISMIN plus désagrégées entre elles avant de décider de regrouper tous les groupes de minorités visibles.

Règles de diffusion pour les différences

Afin d'assigner une catégorie de diffusion pour une estimation de différence entre deux estimations, l'analyste doit d'abord déterminer la catégorie de diffusion de chacune des deux estimations en utilisant les règles décrites précédemment. Ensuite, la catégorie de diffusion de l'estimation de la différence ou de l'estimation du changement est assignée à la catégorie de diffusion la plus basse des deux estimations; ceci peut être spécifié ainsi :

- Si une estimation ou les deux sont de catégorie F, l'estimation de la différence est assignée à la catégorie F et doit être supprimée.
- Sinon, assigné l'estimation de la différence à la catégorie E.

Règles additionnelles au sujet des intervalles de confiance

Les règles de diffusion précédentes devraient supprimer la plupart des estimations et intervalles de confiance de piètre qualité. Il y a également deux conditions qui indiquent si un intervalle de confiance est de piètre qualité. Une estimation et son intervalle de confiance devraient être assignés à la catégorie de diffusion F si l'une ou l'autre des deux conditions suivantes est satisfaite :

- La borne inférieure de l'intervalle de confiance de 95 % est égale à la borne supérieure de l'intervalle; en d'autres termes, l'intervalle de confiance est de longueur nulle. (Une exception s'applique si l'estimation correspond à un total de contrôle de calage.)
- La borne inférieure ou supérieure de l'intervalle de confiance de 95 % n'est pas une valeur plausible pour l'estimation. Par exemple, la borne inférieure d'une estimation de proportion est négative.