

Enquête québécoise sur la découverte de produits culturels et le numérique

Méthodologie de l'enquête



Pour tout renseignement concernant l'Institut de la statistique du Québec (ISQ) et les données statistiques dont il dispose, s'adresser à :

Institut de la statistique du Québec
200, chemin Sainte-Foy, 3^e étage
Québec (Québec) G1R 5T4

Téléphone :
418 691-2401
1 800 463-4090 (Canada et États-Unis)

Site Web : statistique.quebec.ca

Ce document est disponible seulement en version électronique.

Dépôt légal
Bibliothèque et Archives nationales du Québec
2^e trimestre 2024
ISBN 978-2-550-98591-4 (en ligne)

© Gouvernement du Québec, Institut de la statistique du Québec, 2024

Toute reproduction autre qu'à des fins de consultation personnelle est interdite sans l'autorisation du gouvernement du Québec.
statistique.quebec.ca/fr/institut/nous-joindre/droits-auteur-permission-reproduction

Septembre 2024

Publication réalisée à
l'Institut de la statistique du Québec par : Oussama Dabdoubi

Sous la coordination de : Marie-Eve Tremblay

Sous la direction de : Éric Gagnon

Révision linguistique et édition : Direction de la diffusion et des communications

Comité de lecture interne : Marie-Eve Tremblay, Éric Gagnon et Lysandre Champagne

Comité de lecture externe :

Enquête sous la responsabilité de : Institut de la statistique du Québec

Étude financée par : Ministère de la Culture et des Communications

Photo en couverture : Eva-Katalin / iStock

Pour tout renseignement concernant
le contenu de cette publication :

Direction de la méthodologie
Institut de la statistique du Québec
200, chemin Sainte-Foy, 3^e étage
Québec (Québec) G1R 5T4
Téléphone :
418 691-2401
1 800 463-4090 (Canada et États-Unis)
Site Web : statistique.quebec.ca

Avertissement

Les proportions de ce rapport sont arrondies à une décimale dans les tableaux et à l'unité dans le texte. Exceptionnellement, dans le texte, les proportions inférieures à 5 % et celles qui ne sont pas reportées dans les tableaux sont présentées avec une décimale. En raison de l'arrondissement des données, le total ne correspond pas nécessairement à la somme des parties.

Notice bibliographique suggérée

DABDOUBI, Oussama (2024). *Enquête québécoise sur la découverte de produits culturels et le numérique – Méthodologie de l'enquête*, [En ligne], Québec, Institut de la statistique du Québec, 20 p. [statistique.quebec.ca/fr/fichier/methodologie-enquete-quebecoise-decouverte-produits-culturels-numerique.pdf].

Table des matières

Introduction	5
1 Plan d'échantillonnage	6
1.1 Population visée	6
1.2 Base de sondage	6
1.3 Méthode de sélection, taille et répartition de l'échantillon	7
2 Collecte de données	8
2.1 Prétest	8
2.2 Méthodes de collecte	8
2.3 Résultats de la collecte	9
3 Traitement des données	11
3.1 Validation et saisie	11
3.2 Pondération	12
3.3 Non-réponse totale	13
3.4 Non-réponse partielle	13
4 Analyse des données, précision et tests statistiques	14
4.1 Précision des estimations et tests statistiques	14
4.2 Indicateur du niveau de revenu du ménage	15
5 Présentation des résultats	16
6 Portée et limites de l'enquête	17
Références bibliographiques	18
Annexe – Indicateurs présentant un taux de non-réponse partielle pondéré supérieur à 5%	19

Introduction

Afin de faire une utilisation adéquate des données et des résultats issus de *l'Enquête québécoise sur la découverte de produits culturels et le numérique (EQDPCN)*, il importe de connaître la méthodologie d'enquête utilisée. En effet, le plan de sondage, les procédures de collecte et le traitement des données sont tous des éléments qui ont une incidence sur les résultats d'une enquête. La connaissance des aspects méthodologiques aide à interpréter adéquatement les résultats et à en apprécier la qualité, la portée et les limites.

Dans le cadre du Plan culturel numérique du Québec, le ministère de la Culture et des Communications a confié à l'Institut de la Statistique du Québec la mission de mener une enquête visant à établir un portrait statistique des pratiques de découverte de contenus culturels à l'aide des réseaux numériques. Plus précisément, l'enquête permet de :

- déterminer les proportions de la population ayant diverses pratiques de découverte de contenus culturels par l'intermédiaire des réseaux numériques ;
- vérifier l'importance de l'influence des algorithmes sur les pratiques de découverte et de consommation ;
- produire des statistiques sur les pratiques de découverte de contenus culturels québécois en français dans l'environnement numérique.

Structure du rapport méthodologique

Les six sections du présent document abordent les principaux éléments de la méthodologie de l'EQDPCN. La section 1 présente le plan d'échantillonnage de l'enquête, alors que la section 2 décrit les stratégies utilisées et les résultats obtenus lors de la collecte des données. La section 3 porte sur le traitement des données : le processus de validation, la méthode de pondération utilisée afin que les résultats puissent être inférés à la population visée, de même que l'examen de la non-réponse partielle et le recours à l'imputation des données pour en atténuer l'incidence sur les analyses. La section 4 traite des méthodes utilisées pour l'analyse des données, de l'estimation de la précision et des tests statistiques. Les normes de présentation des résultats sont précisées à la section 5, alors qu'un aperçu de la portée et des limites de l'enquête est offert à la section 6.

Le présent document accompagne le rapport de l'enquête intitulé *La découverte de produits culturels et le numérique, Québec, 2023*.

1

Plan d'échantillonnage

Cette section comprend une description de la population visée et de la base de sondage, ainsi que toutes les informations utiles sur la sélection de l'échantillon à partir de cette base.

1.1 Population visée

La population visée par l'EQDPCN 2023 correspond à l'ensemble des personnes de 15 ans et plus vivant dans un logement non institutionnel au Québec. Sont comprises les personnes qui vivent dans un logement privé, celles qui vivent dans un logement collectif non institutionnel (résidence pour personnes âgées, couvent, etc.) ainsi que celles qui vivent dans des communautés autochtones autres que les communautés qui font partie des régions sociosanitaires du Nunavik et des Terres-Cries-de-la-Baie-James.

Les personnes vivant dans un logement collectif institutionnel (hôpital, CHSLD, centre jeunesse, centre de réadaptation, etc.) de même que celles résidant dans les régions sociosanitaires du Nunavik et des Terres-Cries-de-la-Baie-James ne sont pas visées par l'enquête.

1.2 Base de sondage

La base de sondage utilisée pour sélectionner l'échantillon de l'enquête a été élaborée à partir des données du Fichier d'inscription des personnes assurées (FIPA) de la Régie de l'assurance maladie du Québec (RAMQ). Ce fichier présente plusieurs avantages : il permet notamment l'exclusion de la majeure partie des personnes non visées par l'enquête, tout en assurant une excellente couverture de la population ciblée. Il permet également la sélection de personnes en fonction de l'âge, du sexe¹ et du lieu de résidence, et contient les renseignements nécessaires pour communiquer avec celles-ci.

Un des principaux inconvénients du FIPA est que certains numéros de téléphone sont inexacts ou même absents, ce qui complique le contact et la collecte auprès de la population échantillonnée. Cette situation est plus fréquente chez les jeunes de 18 à 24 ans. L'utilisation d'autres sources administratives, notamment celle du ministère de l'Éducation du Québec, a permis de pallier cette lacune en conformité avec la *Loi sur l'Institut de la statistique du Québec*. La base de sondage présente une légère sous-couverture, puisque certaines personnes admissibles à l'enquête ne sont pas inscrites au régime québécois d'assurance maladie (les personnes n'ayant

1. Pour les personnes transgenres ayant obtenu un changement de la mention du sexe figurant à leur acte de naissance auprès du Directeur de l'état civil, le sexe indiqué dans le FIPA est celui indiqué dans l'acte de naissance après le changement.

pas renouvelé leur carte de la RAMQ, par exemple) et ne figurent donc pas au FIPA. De même, la base de sondage pourrait compter quelques personnes inadmissibles à l'enquête, par exemple celles n'ayant pas encore informé la RAMQ d'un déménagement à l'extérieur du territoire visé par l'enquête. Il est toutefois difficile de quantifier de manière précise la couverture nette de la base de sondage, étant donné que la taille de la population visée ne peut pas être établie de façon exacte. En comparant les effectifs de la base de sondage aux plus récentes estimations de population produites par l'ISQ, ajustées pour tenir compte de l'exclusion des personnes en institution de la base de sondage, on peut établir que la couverture de la population visée par le FIPA est d'au moins 95 %. En fait, le pourcentage de couverture est probablement un peu plus élevé, puisque les estimations de population prises en compte dans le calcul incluent des personnes n'étant pas visées par l'enquête, par exemple des résidents et résidentes non permanents qui ne sont pas couverts par le régime québécois d'assurance maladie.

1.3 Méthode de sélection, taille et répartition de l'échantillon

L'enquête vise à fournir des estimations à l'échelle de la province, de la région métropolitaine de recensement (RMR) de Montréal, du groupe des autres RMR (Saguenay, Québec, Sherbrooke, Trois-Rivières, Ottawa Gatineau, Drummondville) et du reste du territoire (hors RMR). L'enquête vise également à fournir des estimations pour plusieurs groupes d'âge et pour des sous-domaines, par exemple l'auditoire internaute des contenus culturels. Le plan de sondage a donc été construit de manière à estimer une proportion de :

- 17 % ou plus avec une bonne précision relative, c'est-à-dire un coefficient de variation² (CV) d'au plus 15 % selon le groupe d'âge (15 à 29 ans, 30 à 44 ans, 45 à 64 ans et 65 ans et plus) et la RMR (RMR de Montréal, autres RMR et hors RMR) ;
- 8 % ou plus avec une bonne précision relative selon le groupe d'âge (15 à 29 ans, 30 à 44 ans, 45 à 64 ans et 65 ans et plus) ;
- 6 % ou plus associée à un CV d'au plus 15 % selon la RMR (RMR de Montréal, autres RMR et hors RMR) ;
- 12 % et plus associée à un CV d'au plus 15 % pour une population représentant 15 % de la population visée.

Le nombre total de répondants et de répondantes à l'EQDPCN nécessaire à l'atteinte de ces objectifs de précision a été fixé à 2 940. Des hypothèses quant aux taux d'admissibilité et de réponse ont par la suite été posées. Ces hypothèses ont ensuite permis de déterminer la taille d'échantillon pour chacune des strates de l'enquête, définies à la fin de la présente section.

Plus précisément, nous avons présumé que nous aurions un taux de réponse de 55 % et un taux d'admissibilité à l'enquête de 99 % ;

Le tirage de l'échantillon s'est fait selon une stratification de la base de sondage visant à assurer un certain contrôle de la composition de l'échantillon. Les strates sont formées par le croisement entre une variable géographique à trois catégories (RMR de Montréal, autre RMR et hors RMR) le groupe d'âge (4 catégories : 15 à 29 ans, 30 à 44 ans, 45 à 64 ans, 65 ans et plus) et le sexe de la personne. L'échantillon a ensuite été réparti de manière égale sur les 24 strates, totalisant ainsi 5400 personnes.

2. Des renseignements sur le calcul du coefficient de variation sont fournis à la section 4.1.

2 Collecte de données

Cette section porte sur les méthodes et les résultats de la collecte des données. On y présente notamment les instruments et le mode de collecte, ainsi que les taux de réponse obtenus.

2.1 Prétest

Du 21 au 29 novembre 2022, un prétest a été effectué auprès de 61 répondants et répondantes à partir d'un échantillon de 300 personnes. Les entrevues ont été réalisées uniquement au moyen d'interviews téléphoniques assistées par ordinateur (ITAO) réalisées avec le logiciel Interviewer de Voxco. L'objectif était de valider les questions du questionnaire et d'estimer le temps requis pour répondre aux différentes sections. À la suite de l'analyse des résultats du prétest, certaines modifications mineures ont été apportées au questionnaire.

2.2 Méthodes de collecte

Instruments de collecte

Le questionnaire de l'*Enquête québécoise sur la découverte de produits culturels et le numérique 2023* a été élaboré par l'Observatoire de la culture et des communications du Québec en collaboration avec le ministère de la Culture et des Communications. Si on exclut l'introduction et les remerciements, le questionnaire compte 4 sections qui abordent les produits culturels suivant :

- Section 1 : Films et séries télévisées
- Section 2 : Musique
- Section 3 : Balados
- Section 4 : Livres en format papier, numérique ou audio.

Il compte aussi une section sur les caractéristiques individuelles.

Modes de collecte

La collecte des données a été effectuée par interview téléphonique assistée par ordinateur (ITAO) et/ou par interview Web assistée par ordinateur (IWAO) à partir du logiciel Interviewer de Voxco. La collecte était multimode, ce qui présente des avantages pour ce qui est de la qualité des données. Premièrement, le libellé des questions (y compris les pronoms), est personnalisé automatiquement en fonction de facteurs comme l'âge, le genre de la personne et les réponses aux questions précédentes. Deuxièmement, des mesures de contrôle isolent les réponses incohérentes ou hors norme, et des instructions apparaissent à l'écran lorsqu'une telle situation se présente. La personne remplissant le questionnaire reçoit une rétroaction immédiate, et une correction de l'incohérence peut être apportée. Troisièmement, le processus fait que les questions qui ne concernent pas la personne ne leur sont pas posées.

Période de collecte

La collecte des données s'est déroulée du 27 mars 2023 au 26 juin 2023.

Stratégies de collecte

Une lettre de présentation de l'enquête a d'abord été envoyée aux personnes sélectionnées pour leur expliquer les objectifs de l'enquête et les inviter à se rendre sur le site Web pour remplir le questionnaire électronique. Pour les personnes de 15 à 17 ans, une lettre a d'abord été envoyée aux parents afin de les informer que leur jeune avait été sélectionné pour participer à l'enquête. Environ deux semaines après l'envoi de la lettre initiale, chaque personne échantillonnée n'ayant pas déjà rempli son questionnaire Web a reçu une lettre de rappel lui demandant de le faire. Deux à trois semaines plus tard, une deuxième lettre de relance a été envoyée à toutes celles qui n'avaient pas encore rempli leur questionnaire. En cours de collecte, des envois postaux ont également été faits aux participants et participantes que nous n'avions pas été capables de joindre.

Une formation a été donnée aux intervieweurs et intervieweuses assignées à la collecte téléphonique de l'enquête afin de bien leur expliquer les tenants et aboutissants de la collecte de renseignements auprès des individus. Cette équipe a été formée non seulement pour mener des entrevues, mais aussi pour effectuer des relances et faire le suivi des dossiers. Après la formation, nous avons commencé les appels téléphoniques afin d'inciter les personnes n'ayant toujours pas rempli leur questionnaire Web à le faire. L'entrevue téléphonique était offerte à celles n'ayant pas accès à Internet ou ne voulant pas remplir le questionnaire en ligne.

2.3 Résultats de la collecte

À l'instar des taux de réponse aux enquêtes généralement présentés par l'ISQ, ceux de l'EQDPCN sont pondérés³. Ils tiennent compte, entre autres, de la non-proportionnalité de l'échantillon de l'enquête. Au total, 2 754 personnes ont participé à l'EQDPCN 2023, ce qui correspond à un taux de réponse pondéré de 53 % (tableau 2.1). Ce taux de réponse est légèrement inférieur à celui qui était visé au départ (55 %). Par sexe, 1 446 femmes ont rempli un questionnaire pour un taux de réponse de 55 %, ce qui correspond bel et bien à la cible fixée. Toutefois, du côté des hommes, le taux de réponse et le nombre de répondants ont été plus faibles que ceux anticipés. Le taux de réponse a atteint 51 %, soit 1 308 questionnaires remplis. On note également des différences selon le groupe d'âge, notamment chez les personnes de 15 à 29 ans, chez qui le taux de réponse est de 40 % (42 % chez les femmes et 37 % chez les hommes). À l'échelle des RMR, le taux de réponse est entre 52 et 57 % chez les femmes, alors qu'il varie de 50 à 52 % chez les hommes.

La première raison qui a contribué à la non-réponse est le fait que le numéro de téléphone identifié après une recherche soit non valide⁴. Cette raison concerne 874 personnes, soit 33 % des non-répondants et non-répondantes.

3. Plus de renseignements sur la pondération sont fournis à la section 3.2.

4. Il s'agit d'un numéro de téléphone non valide, en ce sens qu'il ne permet pas de joindre la personne échantillonnée. Après avoir effectué une recherche, on n'a pas pu trouver de nouveau numéro.

Une deuxième raison, soit le refus de participer à l'enquête, contribue aussi à la non-réponse. Elle représente 25 % des non-répondants et non-répondantes. Ce sont les personnes de 65 ans et plus qui ont proportionnellement refusé le plus de répondre au questionnaire (28 %), alors que chez les autres groupes d'âge, la proportion de refus varie de 22 % à 25 %.

Comme prévu, le taux d'admissibilité pondéré à l'enquête est très élevé (98,9 %). Seules quelques personnes ont été jugées inadmissibles au moment de la collecte :

celles qui avaient déménagé dans un ménage collectif institutionnel ou à l'extérieur du Québec et celles qui étaient décédées. Finalement, 79 % des répondants et répondantes ont opté pour le questionnaire en ligne.

La durée moyenne d'un questionnaire rempli par entrevue téléphonique est estimée à environ 17 minutes, et celle pour un questionnaire rempli sur le Web, à environ 13 minutes.

Tableau 2.1

Nombre de personnes répondantes et taux de réponse pondéré selon l'âge et la région métropolitaine de recensement au global et par sexe, Québec, 2023

Caractéristiques	Nombre de personnes répondantes			Taux de réponse pondéré		
	Femme ³	Homme ²	Total	Femme	Homme	Total
	n			%		
Âge¹						
15-29 ans	278	236	514	42,0	37,2	39,6
30-44 ans	342	303	645	49,8	46,9	48,4
45-64 ans	409	344	753	59,2	51,3	55,2
65 ans +	417	425	842	62,0	64,2	63,0
RMR²						
RMR de Montréal	462	464	926	52,2	52,4	52,3
Autres RMR	496	454	950	57,1	51,7	54,5
Hors RMR	488	390	878	57,1	46,9	52,0
Total	1 446	1 308	2 754	54,6	50,7	52,7

1. Il s'agit de la catégorie d'âge selon le FIPA.
2. Il s'agit d'une variable dérivée de la région administrative et du code postal du FIPA.
3. Il s'agit du sexe selon le FIPA.

Source : Institut de la statistique du Québec, *Enquête québécoise sur la découverte de produits culturels et le numérique*, 2023.

3 Traitement des données

Cette section aborde les étapes de traitement des données recueillies. Elle traite de la validation de la qualité des données, de la pondération nécessaire à l'inférence à la population visée ainsi que de l'examen de la non-réponse totale et partielle et des traitements d'imputation effectués.

3.1 Validation et saisie

Une validation du statut de réponse des personnes sélectionnées (répondants et répondantes, non répondants et non répondantes, ou inadmissibles) a d'abord été effectuée. Cela a permis de considérer comme répondantes certaines personnes qui n'avaient pas rempli le questionnaire en entier et de rejeter certains questionnaires comportant une importante proportion de questions sans réponse. Des questions posées en début d'entrevue visaient à assurer que le questionnaire soit bel et bien rempli par la personne sélectionnée. Des validations ont été effectuées en cours de collecte et a posteriori pour comparer les réponses fournies aux renseignements contenus dans la base de sondage, plus particulièrement pour la date de naissance et le code postal du lieu de résidence. La concordance étant exacte à quelques exceptions près, rien n'indiquait que certains questionnaires devaient être rejetés. Pour la faible proportion de personnes répondantes pour lesquelles la date de naissance fournie à l'entrevue correspondait aux renseignements de la base de sondage, mais qui ont déclaré un code postal différent de celui inscrit dans le FIPA, on a considéré que le questionnaire était bien rempli par la personne sélectionnée ; le nouveau code postal fourni a été conservé s'il s'agissait d'un code postal valide.

Étant entièrement informatisées, les collectes téléphoniques et sur le Web permettent d'effectuer plusieurs validations de base pendant l'entrevue, notamment une vérification du respect des choix de réponse ou de l'adéquation des sauts de section, pour les blocs de questions ne s'adressant qu'à un sous-groupe de répondants et répondantes. Elles permettent aussi une saisie automatique des réponses fournies par les personnes répondantes. La très grande majorité des questions sont donc ainsi autocodées, c'est-à-dire que l'intervieweur ou l'intervieweuse, ou la personne répondante n'a qu'à cocher le choix de réponse : la saisie s'en trouve donc réduite au minimum. Une validation a posteriori a été effectuée afin d'assurer la cohérence des réponses fournies par une même personne, ce qui a permis de relever quelques erreurs qui ont été corrigées. Enfin, les réponses ont été colligées de manière à distinguer les personnes ayant omis de répondre à une question de celles qui n'étaient pas concernées par une question.

3.2 Pondération

La pondération est essentielle pour la production des résultats de l'enquête. Elle permet de faire des inférences adéquates à la population visée, bien que celle-ci n'ait pas été sondée dans sa totalité. Elle consiste à attribuer un poids statistique à chaque personne répondante. Ce poids correspond au nombre de personnes qu'un répondant ou une répondante représente au sein de la population visée. Il doit tenir compte, entre autres, de la probabilité de sélection de la personne, prédéterminée par le plan d'échantillonnage, et de la non-réponse à l'enquête. En effet, en raison des objectifs de diffusion de résultats, certains groupes d'âge ont été surreprésentés, le plan de sondage a inévitablement entraîné des probabilités de sélection très variables. De plus, il est connu que dans les enquêtes, la probabilité de répondre varie selon plusieurs caractéristiques sociodémographiques. Il est donc nécessaire de tenir compte de ces éléments en les intégrant à la pondération.

De façon plus détaillée, le poids initial de chaque personne faisant partie de l'échantillon de l'EQDPCN 2023 a été déterminé en fonction de l'inverse de sa probabilité de sélection. Ensuite, un ajustement a été fait pour l'admissibilité à l'enquête, qui varie significativement selon l'âge. Le statut d'admissibilité étant inconnu pour la plupart des non-répondants et non-répondantes, il fallait effectivement réduire leur poids afin de refléter le fait qu'une partie d'entre eux était probablement inadmissible. Le poids de ces personnes a été multiplié par le taux d'admissibilité pondéré observé à l'enquête parmi les personnes ayant des caractéristiques semblables pour lesquelles l'admissibilité a pu être déterminée. L'ajustement a été fait selon deux catégories d'âge (moins de 65 ans, 65 ans et plus), car la proportion de personnes inadmissibles est plus importante pour les personnes plus âgées (il y a notamment une plus grande proportion de décès chez ces personnes).

L'étape suivante est l'ajustement pour la non-réponse totale. La méthode du score de propension à répondre a été employée (Haziza et Beaumont 2007 ; Eltinge et Yansaneh 1997). Elle consiste à modéliser le fait d'avoir répondu ou non à l'enquête selon les renseignements disponibles dans la base de sondage, entre autres l'âge, le sexe, la langue de correspondance et l'indice de défavorisation matérielle et sociale (Gamache et autres 2017). Des classes composées de personnes ayant des caractéristiques et une propension à répondre semblables ont ainsi été formées. À l'intérieur de chaque classe, le poids des répondants et répondantes a été ajusté par l'inverse du taux de réponse observé à l'enquête.

Par la suite, on a vérifié qu'aucune personne n'avait de poids très élevé comparativement aux poids des personnes de la même RMR, du même sexe et groupe d'âge afin de s'assurer qu'un répondant ou une répondante n'ait pas une influence indue sur les statistiques produites. Pour ce faire, une méthode appelée « écart-sigma » a été utilisée (Bernier et Nobrega 1998). Lorsque jugé trop élevé, le poids d'une personne est abaissé au poids inférieur le plus près dans le même croisement RMR-sexe-groupe d'âge. Le poids de 28 répondants et répondantes a ainsi été modifié (le poids de 26 d'entre eux n'a toutefois été que légèrement modifié, le facteur de correction étant de 0,9). La diminution des poids au cours de cette étape sera compensée à l'étape suivante.

La dernière étape, soit le calage aux marges⁵, consiste à ajuster la pondération afin que la somme des poids des répondants et répondantes corresponde bien aux effectifs connus de la population visée par l'enquête⁶, et cela pour :

- le croisement du sexe et de la catégorie d'âge (13 catégories : 15-19 ans, 20-24 ans, 25-29 ans, 30-34 ans, 35-39 ans, 40-44 ans, 45-49 ans, 50-54 ans, 55-59 ans, 60-64 ans, 65-69 ans, 70-74 ans et 75 ans et plus)⁷ ;
- la RMR (3 catégories : Montréal, autres RMR et hors RMR)

5. Le calage aux marges est une procédure que l'on peut appliquer pour incorporer des données auxiliaires. Cette procédure rajuste les poids d'échantillonnage au moyen de multiplicateurs appelés facteurs de calage, lesquels font correspondre les estimations aux totaux connus.

6. Ces effectifs sont dérivés des plus récentes estimations de population de l'ISQ (1^{er} juillet 2022).

7. On peut poser l'hypothèse que pour une personne donnée, le sexe qui se trouve dans la base de sondage et le sexe qui a été déclaré au recensement de 2021 ont de fortes chances d'être les mêmes, et, dans les faits, de correspondre au genre dans le cas des personnes transgenres. Sur la base de cette hypothèse, le calage aux marges a été réalisé à partir de la variable « sexe » de la base de sondage.

3.3 Non-réponse totale

Définition

La non-réponse totale survient lorsqu'une personne sélectionnée et admissible ne remplit pas le questionnaire de l'enquête pour diverses raisons. Cette non-réponse peut entraîner des biais dans les estimations si les personnes ayant répondu présentent des caractéristiques différentes de celles ne l'ayant pas fait, et que ces caractéristiques sont liées au sujet de l'enquête. La pondération est ajustée pour la non-réponse à l'enquête (section 3.2), ce qui permet de réduire le risque de biais dû à celle-ci. Toutefois, seules les informations contenues dans la base de sondage, donc connues pour l'ensemble de l'échantillon, peuvent être prises en considération pour cet ajustement. Ainsi, malgré l'utilisation de la pondération, des résultats de l'enquête peuvent quand même être biaisés si la non-réponse totale est liée à une ou plusieurs caractéristiques non disponibles dans la base de sondage, et que ces caractéristiques sont fortement corrélées à certains indicateurs.

3.4 Non-réponse partielle

Définition

La non-réponse partielle désigne l'absence de réponse à une question pour certaines personnes ayant rempli le questionnaire. Il est connu qu'une non-réponse partielle importante peut entraîner certains biais dans les estimations, au même titre que la non-réponse totale, s'il s'avère que les non-répondants et non-répondantes présentent des caractéristiques différentes de celles des personnes répondantes et que ces caractéristiques sont de surcroît liées au thème étudié. La pondération ne tient pas compte de la non-réponse partielle comme elle le fait pour la non-réponse totale.

Taux de non-réponse partielle

Le taux de non-réponse partielle pondéré à une question est défini comme le rapport entre le nombre pondéré de personnes n'ayant pas répondu à celle-ci et le nombre pondéré de personnes admissibles à y répondre. Plus ce taux est élevé, plus le risque de biais induits par la non-réponse partielle est grand. On fait généralement l'hypothèse qu'une non-réponse partielle inférieure à 5 % a une incidence négligeable sur les estimations.

La non-réponse partielle étant peu élevée pour la grande majorité des questions de l'EQDPCN 2023, son incidence sur le risque de biais dans les résultats est faible. Les indicateurs qui affichent une non-réponse partielle supérieure à 5 % concernent certains moyens et certaines sources d'information utilisés pour découvrir les séries ou les films (annexe A). Ces indicateurs ont fait l'objet d'une analyse de la non-réponse partielle plus approfondie afin de déterminer si les personnes ayant omis de répondre présentaient des caractéristiques différentes des autres répondants et répondantes et, le cas échéant, l'ampleur du biais potentiel a été estimée. Cette analyse laisse supposer que la non-réponse partielle n'a qu'un effet négligeable sur les estimations pour ces indicateurs.

4 Analyse des données, précision et tests statistiques

Cette section porte sur certains aspects de l'analyse, dont l'estimation de la précision et les tests statistiques. On y présente également l'indicateur de niveau de revenu du ménage pouvant être utilisé comme variable de croisement dans l'EQDPCN.

4.1 Précision des estimations et tests statistiques

La plupart des enquêtes statistiques comportent des erreurs dites d'échantillonnage, dues au fait que seule une partie des unités de la population visée est sélectionnée pour y participer. Ces erreurs se répercutent sur les estimations produites, dont la précision est par ailleurs influencée par la complexité du plan de sondage. Il est donc nécessaire de mesurer la précision de chaque estimation, et d'en tenir compte dans l'interprétation des résultats inférés à la population visée.

Dans l'EQDPCN, l'estimation de la variance et les tests statistiques sont effectués à l'aide de poids d'autoamorçage. Une série de 500 poids d'autoamorçage a été créée afin de tenir compte adéquatement, dans l'estimation de la variance et la production des tests statistiques, du plan de sondage complexe et de tous les ajustements de non-réponse et de calage apportés à la pondération. Pour ce faire, on a d'abord sélectionné 500 échantillons d'autoamorçage selon un plan de sondage avec remise à partir de l'échantillon initial. Ensuite, toutes les étapes de la pondération ont été appliquées à chacun des

échantillons, générant ainsi 500 poids d'autoamorçage (Rust et Rao 1996). Ces poids doivent être utilisés dans l'estimation de la variance et dans les tests statistiques à l'aide de logiciels tels que SAS et SUDAAN. Le CV⁸ a été retenu comme indicateur de précision relative pour les diffusions de résultat de l'ISQ ; les estimations dont le CV est supérieur à 15 % sont annotées dans les tableaux et les figures, comme on l'indique à la section 5, ainsi que dans le texte s'il y a lieu.

Un test statistique d'indépendance du khi-deux⁹ peut être utilisé pour faire une comparaison globale des proportions entre différents sous-groupes (p. ex. les catégories d'âge). En présence d'un écart significatif au seuil de signification fixé, soit 0,05, et lorsque la variable d'analyse ou la variable de croisement compte plus de deux catégories, des tests de comparaison des proportions peuvent être menés afin de repérer les écarts les plus importants. Ces tests reposent sur une statistique de Wald construite à partir de la différence de la transformation « logit » des proportions (Korn et Graubard 1999).

8. Le coefficient de variation est obtenu en divisant l'erreur-type de l'estimation par l'estimation elle-même.

9. On utilise une version modifiée du test du khi-deux habituel qui tient compte du plan de sondage de l'enquête : il s'agit de l'ajustement de Satterthwaite du test du khi-deux. Plus précisément, c'est la statistique F correspondant à cette correction du test du khi-deux qui est utilisée dans les analyses.

Des résultats non significatifs peuvent être signalés s'ils offrent un intérêt particulier ou s'ils font exception dans une série de résultats significatifs. Ils peuvent être présentés sous forme de tendance.

4.2 **Indicateur du niveau de revenu du ménage**

Un indicateur du niveau de revenu a été construit en utilisant les seuils de l'indicateur de faible revenu basé sur la mesure de faible revenu (MFR) avant impôt. Cette mesure permet d'estimer la situation économique d'un ménage d'après son revenu avant impôts et sa taille. Ainsi, la MFR tient compte du fait que les ménages de grande taille ont des besoins de base plus importants. Elle tient également compte du fait que l'ajout d'une personne est moins coûteux dans un grand ménage que dans un ménage plus petit¹⁰. Toutefois, il est important de noter que, puisqu'une seule question sur le revenu est posée dans l'EQDPCN, contrairement à d'autres enquêtes portant sur le revenu, l'indicateur peut surestimer la défavorisation des ménages, causée par une sous-déclaration des revenus. Pour une estimation plus fiable de la prévalence de faible revenu des personnes de 15 ans et plus, le lectorat est invité à consulter les publications de l'ISQ sur le sujet¹¹. L'indicateur construit permet tout de même de créer des groupes homogènes de personnes selon le niveau de revenu de leur famille. Le premier groupe contient les personnes de la population visée qui sont les moins bien nanties en termes relatifs.

10. Pour plus d'information sur la MFR, consultez le site Web de l'Institut de la statistique du Québec à l'adresse suivante : statistique.quebec.ca/fr/document/faible-revenu-a-lechelle-du-quebec-selon-la-mesure-du-faible-revenu-mfr.

11. Pour plus d'information sur les revenus au Québec, veuillez consulter les publications de l'Institut de la statistique du Québec sur le sujet à l'adresse suivante : statistique.quebec.ca/fr/document/revenu-et-faible-revenu-quebec-2019-donnees-et-tendances-depuis-25-ans.

5 Présentation des résultats

Dans le cadre de l'EQDPCN, les estimations de proportions publiées par l'ISQ sont arrondies à la décimale près dans les tableaux et figures et à l'unité près dans le texte, à l'exception des proportions inférieures à 5 % pour lesquelles une décimale a été conservée. En raison de l'arrondissement, la somme des proportions présentées dans certains tableaux ou certaines figures peut différer légèrement de 100 %.

Dans les tableaux et figures, les estimations dont le CV est inférieur ou égal à 15 % sont suffisamment précises pour être présentées sans indication, alors que les estimations dont le CV est supérieur à 15 % sont accompagnées d'un astérisque (*) (voir tableau 5.1). Fait à signaler, les estimations provinciales sans croisement dont le CV est supérieur à 25 %, marquées d'un double astérisque (**), ne sont généralement pas interprétées dans le texte.

Dans les tableaux et figures présentant des analyses bivariées, en présence d'une association significative selon le test du khi-deux, des lettres en exposant ajoutées aux statistiques présentées indiquent quelles sont les paires de catégories d'une variable de croisement pour lesquelles la variable d'analyse diffère significativement. Une même lettre révèle un écart significatif entre deux catégories.

Tableau 5.1

Système de classification des estimations, en termes de précision, en vigueur à l'Institut de la statistique du Québec

Précision de l'estimation (CV)	Signe distinctif	Réserve à mentionner lors de la diffusion
CV ≤ 15 %	(aucun signe)	Estimation diffusée sans mention.
15 % < CV ≤ 25 %	*	Estimation à interpréter avec prudence.
CV > 25 %	**	Estimation à utiliser avec circonspection, fournie à titre indicatif seulement.

Source : Institut de la statistique du Québec.

12. Pour plus de détail sur la prise en compte du genre dans les enquêtes de l'ISQ, consulter le : statistique.quebec.ca/fr/institut/methodologie/classifications-et-definitions/methodologie-prise-en-compte-du-genre.

6

Portée et limites de l'enquête

La participation des individus à l'enquête a été relativement bonne, comme le démontre le taux de réponse obtenu lors de la collecte des données (53 %), qui est près de celui prévu (55 %).

Les données des 2 754 répondants et répondantes de l'EQDPCN 2023 offrent un bon potentiel analytique à l'échelle provinciale. Tout a été mis en place pour maximiser la qualité et la représentativité des résultats produits. La précision de la grande majorité des proportions d'intérêt est bonne pour plusieurs domaines d'analyse. Cependant, le questionnaire contient un grand nombre de questions destinées à des sous-populations en fonction de leurs pratiques culturelles. Par conséquent, pour de nombreuses estimations, les tailles d'échantillons sont plus petites, ce qui pourrait limiter la capacité à réaliser des croisements et à obtenir des résultats significatifs.

Toutes les estimations produites dans le cadre de l'EQDPCN 2023 sont pondérées et tiennent compte non seulement du plan de sondage, mais aussi de la non-réponse totale, de manière à assurer la fiabilité de l'inférence à la population visée. De plus, toutes les mesures de précision et les tests statistiques ont été produits en considérant la complexité du plan de sondage de l'enquête.

Il est important de tenir compte des mises en garde et des recommandations mentionnées dans les sections précédentes afin d'utiliser adéquatement les données de cette enquête. La section 4.1 explique, entre autres, comment tenir compte du plan de sondage complexe lors de la mesure de la précision des résultats produits.

Il est important de mentionner que des données d'observation telles que celles recueillies dans le cadre de l'EQDPCN 2023 ne permettent généralement pas d'établir de lien de causalité. Les associations ou les différences significatives observées peuvent suggérer un tel lien, mais elles ne permettent pas de le confirmer. Compte tenu des objectifs spécifiques de l'enquête, les analyses présentées par l'ISQ s'appuient sur des méthodes bivariées. L'interprétation de certains résultats doit donc être faite avec prudence. Une analyse multivariée aurait été appropriée dans certains cas pour contrôler des facteurs confondants. L'approche retenue permet néanmoins d'explorer les données recueillies de façon utile et de fournir un portrait des phénomènes étudiés.

Références bibliographiques

- BERNIER, J., et K. NOBREGA (1998). "Outlier detection in asymmetric samples: A comparison of an inter-quartile range method and a variation of a sigma-gap method". *Congrès annuel de la société Statistique du Canada*, [En ligne], p. 137-141. [ssc.ca/sites/default/files/survey/documents/SSC1998_J_Bernier.pdf] (Consulté le 25 mai 2023).
- ELTINGE, J. L., et I. S. YANSANEH (1997). « Méthodes diagnostiques pour la construction de cellules de correction pour la non-réponse, avec application à la non-réponse aux questions sur le revenu de la U.S. Consumer Expenditure Survey », *Techniques d'enquête*, [En ligne], produit n° 12-001-X19970013103 au catalogue de Statistique Canada, vol. 23, n° 1, juin, p. 37-45. [www150.statcan.gc.ca/n1/fr/pub/12-001-x/1997001/article/3103-fra.pdf?st=YRiuivNK] (Consulté le 27 février 2023).
- GAMACHE, P., D. HAMEL et R. PAMPALON (2017). *L'indice de défavorisation matérielle et sociale : en bref*, [En ligne], [Québec], Institut national de santé publique du Québec, 9 p. [www.inspq.qc.ca/sites/default/files/santescope/indice-defavorisation/guidemethodologiquefr.pdf] (Consulté le 27 février 2023).
- Haziza, D., et J.-F. BEAUMONT (2007). "On the Construction of Imputation Classes in Surveys", *International Statistical Review*, [En ligne], vol. 75, n° 1, avril, p. 25-43. doi : [10.1111/j.1751-5823.2006.00002.x](https://doi.org/10.1111/j.1751-5823.2006.00002.x). (consulté le 25 mai 2023).
- INSTITUT DE LA STATISTIQUE DU QUÉBEC (2022). *Revenu et faible revenu au Québec en 2019 : les plus récentes données et les tendances depuis 25 ans*, [En ligne], n° 2, 20 p. [statistique.quebec.ca/fr/fichier/revenu-faible-revenu-quebec-2019-donnees-tendances-depuis-25-ans.pdf] (consulté le 15 mars 2023).
- KORN, E. L., et B. I. GRAUBARD (1999). *Analysis of Health Surveys*, New York, John Wiley & Sons, 382 p.
- RUST, K. F., et J. N. K. RAO. (1996). "Variance estimation for complex surveys using replication techniques", *Statistical Methods in Medical Research*, [En ligne], vol. 5, n° 3, septembre, p. 283-310. doi : [10.1177/096228029600500305](https://doi.org/10.1177/096228029600500305). (Consulté le 25 mai 2023).

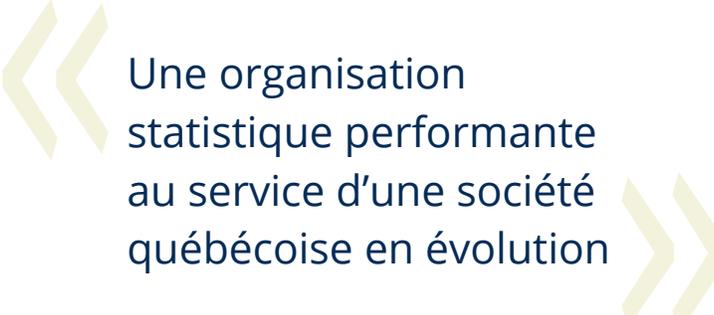
Annexe

Indicateurs présentant un taux de non-réponse partielle pondéré supérieur à 5 %

Tableau A.1

Indicateurs présentant un taux de non-réponse partielle pondéré supérieur à 5 %

Nom de la variable	Mesure/Indicateur	Taux de non-réponse partielle (%)	Effet possible sur l'analyse
Q3B	Proportion de la population découvrant les films ou les séries en se rendant dans un magasin ou une bibliothèque pour consulter ce qui est offert.	5,7	Semble négligeable
Q3F	Proportion de la population découvrant les films ou les séries en suivant les différents hyperliens sur le Web	5,2	Semble négligeable
Q3H	Proportion de la population découvrant les films ou les séries en s'informant sur les acteurs ou les réalisateurs	5,1	Semble négligeable
Q14K	Proportion de personnes découvrant les films ou les séries d'origine québécoise par l'entremise des associations d'artistes ou des professionnels de la culture	5,3	Semble négligeable



Une organisation
statistique performante
au service d'une société
québécoise en évolution

statistique.quebec.ca