
ENQUÊTE SUR LES DÉPARTS AÉRIENS
RÉSIDENTS DES ÉTATS-UNIS ET DES PAYS D'OUTRE-MER EN VOYAGE AU
CANADA
ARRIVÉES ET DÉPARTS PAR VOLS COMMERCIAUX
FICHIERS DE MICRODONNÉES À GRANDE DIFFUSION (FMGD)
2023



Pour obtenir plus de renseignements

Les demandes de renseignements concernant ce produit et des statistiques ou des services connexes doivent être adressées à :

Statistique Canada

150, promenade Tunney's Pasture, Ottawa, Ontario, K1A 0T6

Téléphone : 1-800-263-1136

Courriel : infostats@statcan.gc.ca

Accéder aux microdonnées et passer une commande

L'Enquête sur les départs aériens (EDA) produit deux types de fichiers de microdonnées : les fichiers maîtres et les fichiers de microdonnées à grande diffusion (FMGD).

Fichiers maîtres

Les fichiers maîtres contiennent toutes les variables et tous les enregistrements de l'enquête recueillis au cours d'une période de référence. Ces fichiers sont accessibles à Statistique Canada pour usage interne et peuvent aussi faire l'objet de demandes de totalisations personnalisées.

Totalisations personnalisées

Le personnel de la Section des services à la clientèle du **Programme de la statistique du tourisme** peut préparer des totalisations personnalisées à partir du fichier maître. Ce service est offert moyennant le recouvrement des coûts. Il permet aux utilisateurs qui n'ont pas les connaissances nécessaires à l'utilisation de logiciels de totalisation d'obtenir des résultats personnalisés. La confidentialité et la fiabilité des résultats sont vérifiées avant leur diffusion. Pour obtenir plus de renseignements, veuillez communiquer avec Statistique Canada.

Fichiers de microdonnées à grande diffusion

Les fichiers de microdonnées à grande diffusion (FMGD) sont élaborés à partir des fichiers maîtres au moyen d'une technique qui vise à concilier l'impératif de confidentialité des répondants et la nécessité de produire les données les plus utiles possibles. Des normes rigoureuses en matière de sécurité et de confidentialité exigées en vertu de la *Loi sur la statistique* doivent être respectées avant que le FMGD soit diffusé au public. Pour s'assurer que ces normes sont respectées, chaque FMGD fait l'objet d'un processus officiel d'examen et d'approbation à Statistique Canada. Les variables les plus susceptibles de mener à l'identification d'une personne sont supprimées du fichier de données ou agrégées en catégories moins détaillées.

À noter qu'il peut y avoir des différences entre les estimations publiées par Statistique Canada dans ses produits statistiques et les estimations reproduites par les utilisateurs au moyen des fichiers de microdonnées à grande diffusion, puisque Statistique Canada utilise des fichiers maîtres pour produire ses estimations. Pour obtenir une copie du FMGD, veuillez communiquer avec Statistique Canada.

Initiative de démocratisation des données

Le programme de l'Initiative de démocratisation des données permet aux étudiants et aux chercheurs d'utiliser les fichiers de microdonnées à grande diffusion dans plusieurs universités du Canada. Pour obtenir plus de renseignements, veuillez consulter la page Web suivante :

<https://www.statcan.gc.ca/fr/microdonnees/idd>.

Table des matières

1.0 CONTEXTE	5
2.0 COLLECTE DES DONNÉES.....	5
3.0 CONCEPTS ET DÉFINITIONS SE RAPPORTANT AU QUESTIONNAIRE	6
3.1 Voyages.....	6
3.2 Visites pendant un voyage	7
3.3 Autres caractéristiques des voyages et des visites	8
3.4 Voyageurs.....	10
3.5 Groupes de voyage.....	11
4.0 TRAITEMENT DES DONNÉES.....	12
4.1 Contrôle	12
4.2 Imputation	13
4.3 Contrôle de la divulgation.....	13
5.0 PONDÉRATION	14
6.0 ESTIMATION	14
7.0 VARIABILITÉ D'ÉCHANTILLONNAGE	15
7.1 Qualité des données	15
7.2 Méthode bootstrap pour estimer la variance	17
8.0 CONTENU DU FICHIER DE MICRODONNÉES À GRANDE DIFFUSION (FMGD)	17
Annexe A : Liste des régions géographiques types pour le lieu de résidence et leur constitution pour les résidents des pays d'outre-mer	19
Annexe B : Groupements d'États de résidence américains utilisés dans le FMGD de l'Enquête sur les départs aériens de 2023	20
Annexe C : Estimation de la variance bootstrap pour les FMGD	21

1.0 CONTEXTE

L'Enquête sur les voyages des visiteurs (EVV) est menée par Statistique Canada pour répondre aux exigences de la balance des paiements du Système de comptabilité nationale du Canada. L'enquête est également utilisée par le Compte satellite du tourisme, l'Agence des services frontaliers du Canada (ASFC), Destination Canada, les organismes touristiques provinciaux et territoriaux et le secteur privé. Elle est en outre utilisée dans le cadre des rapports destinés à des organisations internationales comme ONU Tourisme et l'Organisation de coopération et de développement économiques.

L'EVV fournit un large éventail de statistiques sur le volume de voyageurs internationaux et les caractéristiques détaillées de leurs voyages, comme les dépenses, les activités, les endroits visités et la durée du séjour.

Avant 2018, les renseignements sur les visiteurs au Canada étaient recueillis dans le cadre de l'Enquête sur les voyages internationaux (EVI). Au fil du temps, diverses questions portant sur des caractéristiques précises des visiteurs se sont ajoutées à l'EVI à des fins d'études de marché et de planification sectorielle. En 2018, l'EVI a été divisée en deux parties : l'Enquête nationale sur les voyages (ENV), qui porte sur les voyages internationaux et intérieurs des résidents canadiens, et l'EVV, qui porte sur les déplacements au Canada des résidents étrangers (États-Unis et pays d'outre-mer).

L'Enquête sur les départs aériens (EDA) est une composante de l'EVI ou de l'EVV depuis 2000 qui porte sur les visiteurs quittant le Canada au moyen d'un vol commercial. Avant la pandémie de COVID-19, les données étaient également recueillies auprès des voyageurs qui ne se déplaçaient pas par voie aérienne au moyen d'une distribution de cartes d'invitation aux points frontaliers. Les restrictions en matière de santé publique ont forcé l'arrêt de toutes les activités de collecte de données de l'EVV de mars 2020 à mars 2023, et bien que la collecte de données de l'EDA ait repris en avril 2023, la collecte concernant les voyageurs utilisant des modes de transport autres que les vols commerciaux n'a pas été rétablie, en raison du nombre insuffisant de réponses recueillies avant la pandémie.

Étant donné qu'un fichier de microdonnées à grande diffusion (FMGD) exige des microdonnées des répondants, la version du FMGD de 2023 se limite à la composante des vols commerciaux recueillie dans le cadre de l'EDA et elle est donc désignée sous le nom de FMGD de l'EDA de 2023. Ce guide a été rédigé pour faciliter l'utilisation du présent FMGD.

2.0 COLLECTE DES DONNÉES

L'EDA permet de recueillir des renseignements mensuels sur les caractéristiques des voyageurs et des voyages internationaux, qui sont diffusés tous les trimestres. Ces renseignements comprennent le lieu de résidence, le motif du voyage, la taille du groupe de voyage, les endroits visités, les types d'hébergement utilisés, les activités réalisées au cours du voyage, la durée du voyage et les dépenses de voyage.

Les données recueillies dans le cadre de l'EDA portent sur les visiteurs résidents des États-Unis et des pays d'outre-mer quittant le Canada au moyen de vols commerciaux, à partir des aéroports internationaux de Vancouver, de Calgary, de Toronto-Pearson, de Montréal-Trudeau et de Halifax. Des intervieweurs de Statistique Canada mènent des interviews en personne à l'aide de tablettes dans les aires de départ de ces cinq aéroports pendant que les voyageurs attendent leur vol de retour.

Les interviews sont réalisées chaque mois dans le cadre de l'EDA durant une période maximale de sept jours. Depuis la reprise de la collecte en avril 2023, environ 10 500 interviews (7 400 interviews

menées auprès de visiteurs des pays d'outre-mer et 3 100 interviews auprès de visiteurs des États-Unis) ont été menées jusqu'en décembre 2023. Environ 2 500 questionnaires supplémentaires imputés (1 400 pour les résidents des pays d'outre-mer et 1 100 pour les résidents des États-Unis) couvrant la période de janvier à mars 2023 ont été ajoutés au fichier afin d'assurer la disponibilité de données pour l'ensemble de l'année.

Les pays étrangers ciblés correspondent à ceux d'où le Canada reçoit le plus de visiteurs, tels que le Royaume-Uni, la France, l'Allemagne, le Mexique, l'Australie, la Chine, Hong Kong, le Japon, l'Inde, la Corée du Sud, les Pays-Bas et la Suisse. L'équipe de l'EDA se compose d'intervieweurs maîtrisant plusieurs langues, ce qui permet de réaliser les interviews dans la langue maternelle des voyageurs lorsque cela est possible. Si nécessaire, une version papier du questionnaire est également mise à disposition dans huit autres langues afin de faciliter la collecte à l'aide d'une tablette.

3.0 CONCEPTS ET DÉFINITIONS SE RAPPORTANT AU QUESTIONNAIRE

3.1 Voyages

Voyage

Un voyage commence lorsqu'un voyageur non résident passe par un point d'entrée de l'ASFC, entre au Canada et se termine lorsque le voyageur quitte le Canada pour retourner à son lieu de résidence habituel. Le programme de dénombrement à la frontière de Statistique Canada utilise les données de l'ASFC pour dénombrer le nombre de voyages effectués par les voyageurs internationaux chaque mois. Les données de ce programme servent de référence pour les renseignements sur les voyages recueillis dans le cadre de l'EDA. Un voyageur peut être inclus dans le dénombrement à la frontière plus d'une fois au cours de la même période de référence s'il entre au Canada plus d'une fois par un point d'entrée de l'ASFC pendant cette période.

Durée du voyage

La durée du voyage se mesure par le nombre de nuits passées par un voyageur au Canada. La durée des voyages est également divisée en deux groupes mutuellement exclusifs : les voyages de même jour et les voyages d'une nuit ou plus.

Voyage de même jour

Un voyage de même jour comprend les voyageurs qui entrent au Canada et en sortent le même jour civil. Cette catégorie comprend les personnes qui résident dans un autre pays et qui se rendent quotidiennement au travail ou en reviennent.

Voyage d'une nuit ou plus

Les voyages d'une nuit ou plus comprennent les voyageurs qui séjournent au Canada pendant au moins une nuit avant de quitter le pays. Cette classification se rapproche beaucoup du terme

« touristique », bien qu'elle comprenne les travailleurs temporaires et saisonniers, les étudiants et les voyageurs qui sont en transit.

But du voyage

La raison principale (ou le but) d'un voyage est classée dans le FMGD de l'EDA de 2023 selon les catégories suivantes :

- Vacances, loisirs ou agrément
- Pour visiter des amis ou de la parenté
- Faire des études
- Pour une autre raison personnelle
- Pour des réunions, ventes ou des appels de service de routine
- Participer à un congrès, à une conférence ou à une foire commerciale
- Pour une autre raison reliée au travail
- En transit

Dans le passé, et pour d'autres produits de l'EDA ou de l'EVV, ces raisons ont été regroupées comme suit :

- Affaires (Comprend la participation à une réunion, à un appel de vente ou de service de routine, à une conférence, à un congrès, à une foire commerciale ou à un séminaire, ainsi que d'autres raisons professionnelles).
- Visite d'amis ou de la parenté.
- Agrément (Comprend les vacances, les visites à une résidence secondaire, une maison de campagne ou une copropriété et les voyages pour participer à des événements ou voir des attractions).
- Autres raisons (Comprend fréquenter l'école ou faire des études, assister à une conférence ou à un congrès, voyager pour des raisons médicales ou de santé, magasinage, transport en commun et autres raisons personnelles).

3.2 Visites pendant un voyage

Visite

Un voyageur non résident peut visiter plusieurs endroits au Canada pendant le même voyage. Dans le questionnaire de l'EDA, un répondant peut déclarer jusqu'à 10 endroits visités pendant son voyage, et ces lieux sont consignés au niveau de la subdivision de recensement. Veuillez noter que, aux fins de publication, un voyage comportant plusieurs visites dans le même lieu géographique sera compté comme une seule visite à cet endroit.

Codage des visites de lieux géographiques

Les endroits visités (p. ex. villes, régions) sont codés selon les classifications géographiques types suivantes :

Région métropolitaine de recensement et agglomération de recensement

Les régions métropolitaines de recensement (RMR) sont définies comme la principale zone du marché du travail d'un noyau urbain de 100 000 habitants ou plus. Les petites régions urbaines, dont le noyau urbain compte au moins 10 000 habitants, sont définies comme des agglomérations de recensement (AR). Pour de plus amples renseignements, consulter la Classification géographique type (CGT) de 2021.

Région touristique

Les régions touristiques sont définies par les ministères du tourisme de chaque province et territoire comme des secteurs d'intérêt distincts pour l'industrie touristique. Ils sont mis à jour chaque année.

Durée de la visite

Dans le cadre de l'EDA, les répondants déclarent le nombre de nuits passées à chaque endroit visité. Aux fins de publication, la durée des voyages est divisée en deux groupes mutuellement exclusifs : les visites de même jour et les visites d'une nuit ou plus.

Visite de même jour

Une visite de même jour s'entend d'une visite pour laquelle le répondant n'a pas déclaré de nuitées pour l'endroit visité. Le voyageur peut visiter plusieurs endroits au cours d'un voyage de même jour, mais un seul est enregistré pour un voyage de même jour. Pour les voyages d'une nuit ou plus, chaque endroit visité où le répondant n'a pas séjourné pour une nuitée est enregistré comme une visite de même jour.

Visite d'une nuit ou plus

Une visite d'une nuit ou plus s'entend d'une visite pour laquelle le répondant a déclaré avoir séjourné pendant une ou plusieurs nuits à l'endroit visité. Les visites d'une nuit ou plus ne peuvent être effectuées que pendant un voyage d'une nuit ou plus.

3.3 Autres caractéristiques des voyages et des visites

Moyen de transport

Dans le cadre du FMGD de l'EDA de 2023, tous les voyageurs entraient au Canada et en sortaient au moyen de vols commerciaux.

Voie d'entrée

Les voyageurs non résidents sont regroupés selon le type de voie d'entrée utilisé pour entrer au Canada. Il y a trois catégories : à partir des États-Unis seulement, directement d'un autre pays et à partir d'un autre pays via les États-Unis.

Hébergement utilisé

Les types d'hébergement utilisés à chaque endroit visité sont recueillis dans le questionnaire, et un répondant peut déclarer plusieurs types d'hébergement pour chaque visite.

Les catégories disponibles sont :

- Hôtel
- Motel
- Domicile d'amis ou de parents
- Terrain de camping
- Chalet
- Autre

Activités

Les répondants déclarent toutes les activités auxquelles ils ont participé pendant le voyage selon les catégories proposées. Toutes les catégories sont regroupées en deux grands groupes : participé à des activités sportives ou de plein air et autres activités.

Aucune information n'est recueillie sur les activités pendant des visites précises. Pour cette raison, une activité ne peut pas être associée à un endroit précis. Par exemple, une personne peut se rendre à Vancouver et à Whistler pendant le même voyage et déclarer qu'elle a fait du ski alpin. Lors de la compilation des données, l'activité « Fait du ski alpin » sera attribuée à Vancouver et à Whistler, puisque les renseignements sur l'activité ne sont recueillis qu'au niveau du voyage. Toutefois, cette activité n'a peut-être eu lieu qu'à Vancouver ou seulement à Whistler.

Dépenses

Les dépenses de voyage effectuées par tous les membres du groupe de voyage sont recueillies dans le questionnaire. Les dépenses ne sont pas déclarées par visite, mais plutôt par catégorie pour l'ensemble du voyage. Les dépenses peuvent avoir été payées par des particuliers, un gouvernement ou une entreprise du secteur privé et sont déclarées en dollars canadiens ou converties en dollars canadiens.

Les dépenses peuvent être classées dans les catégories suivantes :

- Hébergement
- Transport au Canada (ce qui comprend l'essence, la location de voiture, les transports interurbains en avion, en autobus, en train, en bateau, les transports locaux en autobus et en taxi)
- Repas et boissons (dont aux restaurants, bars et magasins)
- Loisirs et divertissements
- Vêtements et cadeaux
- Autres

Tarifs

Les tarifs correspondent aux sommes dépensées par tous les membres du groupe pour les vols commerciaux pour entrer au Canada et en sortir. Les tarifs sont déclarés en dollars canadiens ou convertis en dollars canadiens. Ils ne font pas partie des dépenses et sont donc déclarés séparément.

3.4 Voyageurs

Voyageur international

L'expression « voyageur international » englobe toutes les personnes arrivant au Canada qui passent par un point d'entrée de l'ASFC. Le dénombrement à la frontière énumère chaque entrée au Canada, ce qui comprend les entrées par des non-résidents et les rentrées par des résidents. Les entrées sont dénombrées plutôt que des voyageurs uniques, de sorte qu'un voyageur sera compté chaque fois qu'il traverse la frontière pendant la période de référence.

Le terme « voyageur international » est divisé en trois groupes mutuellement exclusifs : les voyageurs non résidents, les voyageurs résidents et les autres voyageurs. Le FMGD de l'EDA ne comprend que les voyageurs non résidents. Pour obtenir de la documentation sur les voyageurs résidents, consultez le Guide de l'utilisateur des microdonnées de l'ENV.

Voyageur non résident

L'EDA définit un voyageur non résident comme un voyageur international dont le pays de résidence est un pays autre que le Canada et qui passe par un point d'entrée de l'ASFC au cours d'une visite pour une période de moins de huit mois. Les citoyens canadiens qui ont résidé dans un autre pays pendant plus de 12 mois et qui visitent le Canada sont considérés comme des voyageurs non résidents.

Voyageur résident

Un voyageur résident est un résident canadien (citoyen canadien ou étranger) qui a voyagé à l'extérieur du Canada pendant moins de 12 mois et qui revient au Canada par un point d'entrée de l'ASFC. Ces voyageurs constituent la population cible de l'ENV et ne font pas partie du FMGD de l'EDA.

Autre voyageur

Les autres voyageurs sont exclus de la population cible de l'EDA puisqu'ils ne voyagent pas à des fins touristiques, et comprennent les types suivants de voyageurs internationaux :

- **Immigrants** : Citoyens d'autres pays entrant au Canada pour y résider de façon permanente pour une période de plus de 12 mois.
- **Anciens résidents** : Citoyens canadiens revenant au Canada pour y résider de façon permanente après avoir résidé à l'extérieur du pays pour une période de plus de 12 mois.
- **Personnel militaire, diplomates et personnes à charge** : Citoyens canadiens et les personnes à leur charge en poste à l'étranger depuis plus de 12 mois et qui reviennent de leur affectation. Citoyens étrangers et les personnes à leur charge entrant au Canada pour y occuper un poste militaire ou diplomatique.

- **Équipage** : Personnes participant à l'exploitation d'un aéronef commercial.

3.5 Groupes de voyage

On entend par « groupe de voyage » plusieurs personnes provenant du même pays qui voyagent ensemble (qu'elles fassent partie du même ménage ou non) et qui s'entendent pour déclarer collectivement les dépenses et les activités effectuées pendant leur voyage.

Composition du groupe de voyage

La composition du groupe de voyage désigne le nombre d'adultes ou d'enfants qui font partie du même groupe de voyage.

Lieu de résidence (lieu d'origine)

Le lieu de résidence (ou d'origine) est le dernier lieu de résidence permanente. Dans le FMGD de l'EDA de 2023, pour les résidents des États-Unis d'Amérique, il s'agit de l'état (ou regroupement d'état) de résidence, et pour les résidents des pays d'outre-mer, il s'agit du pays (ou regroupement de pays).

Voyage-personne

Un voyage-personne commence lorsqu'un voyageur non résident passe par un point d'entrée de l'ASFC, entre au Canada et se termine lorsque le voyageur quitte le Canada pour retourner à son lieu de résidence habituel. Un voyage-personne est différent d'un voyage parce que le voyage-personne est pondéré par le nombre de personnes faisant partie du groupe de voyage, alors que le voyage est entrepris par tout le groupe de voyage.

Nuitée-voyage-personne

Une nuitée-voyage-personne se définit comme la nuitée qu'un voyageur non résident passe au Canada pendant un voyage-personne.

Visite-personne

Une visite-personne est soit une visite de même jour ou une visite d'une nuit ou plus effectuée par une personne qui est en voyage. La visite-personne diffère d'une visite, puisque la visite-personne est pondérée par le nombre de personnes faisant partie du groupe de voyage, alors que la visite est prise par tout le groupe de voyage. Le nombre total de visites-personnes correspond au nombre total de visites effectuées par chaque personne, dans chaque groupe de voyage, pour chaque voyage effectué.

Nuitée-visite-personne

Une nuitée-visite-personne est une nuitée passée à l'endroit visité par une personne qui effectue un voyage. Par exemple, une personne qui passe trois nuits à l'endroit visité compte pour trois nuitées-visite-personne à cet endroit, au même titre que trois personnes qui passent une nuitée à l'endroit visité.

4.0 TRAITEMENT DES DONNÉES

Le traitement de l'enquête a lieu après que toutes les données ont été recueillies et avant l'estimation, et il comprend le contrôle, le codage, l'imputation et la vérification des données. L'objectif du traitement est de s'assurer que les données conviennent à l'estimation et à la totalisation.

4.1 Contrôle

Les données ne sont pas toujours déclarées sur un questionnaire de manière à ce qu'elles puissent être utilisées directement dans une analyse plus poussée. Le contrôle des données permet de reformater les données afin qu'elles puissent être utilisées pour la totalisation, l'analyse et l'estimation.

4.1.1 Vérifications de validité et de cohérence

Des erreurs peuvent survenir dans la saisie des données, par exemple une entrée de donnée erronée ou encore une donnée manquante dans le questionnaire. Ces erreurs doivent être repérées et corrigées. Les règles de vérification sont conçues pour s'assurer que les données sont cohérentes avec des caractéristiques connues, et que les données du questionnaire représentent des valeurs valides compte tenu des questions.

Voici quelques exemples de modifications apportées pendant le processus de saisie des données :

- Tous les répondants à l'EDA quittent le Canada au moyen de vols commerciaux. Donc, même si un répondant indique qu'il part en voiture, cela est corrigé pour indiquer qu'il s'agit d'un vol commercial.
- Si tous les coûts de transport sont inclus dans le coût d'un forfait, mais que des coûts sont aussi indiqués dans les tarifs pour quitter le Canada et y revenir, alors ces données sont corrigées afin d'éliminer la duplication des dépenses déclarées.

4.1.2 Création de variables dérivées

Les dépenses déclarées lors des voyages à forfait doivent être ventilées en composantes de dépenses, les données déclarées pour les voyages comportant plusieurs entrées au Canada doivent être séparées et des variables dérivées doivent être générées.

L'EDA demande aux répondants de l'information sur les voyages à forfait ou tout compris achetés. Ces achats ne sont pas déclarés dans la section des dépenses effectuées au Canada. Par conséquent, les voyages à forfait doivent être ventilés selon les composantes appropriées dans l'ensemble des coûts et ajoutés aux catégories de dépenses pertinentes. Les variables des dépenses qui comprennent les valeurs des voyages à forfait et non à forfait sont appelées variables des dépenses globales.

Selon la définition de « voyage » de l'EDA, toute sortie à l'extérieur du Canada marque la fin d'un voyage. Les répondants peuvent déclarer plusieurs voyages selon cette définition en remplissant leur questionnaire. Ces voyages doivent donc être ventilés afin de permettre une analyse cohérente.

Enfin, on crée toutes les variables requises pour une analyse approfondie.

4.2 Imputation

On effectue l'imputation de données dans l'EDA lorsque des champs sont manquants dans un questionnaire, ou lorsqu'un nombre insuffisant de questionnaires est reçu pour satisfaire aux objectifs de l'EDA.

Un minimum de renseignements est nécessaire pour qu'un questionnaire soit considéré comme fiable. Les renseignements suivants doivent être disponibles pour qu'un questionnaire puisse être conservé et utilisé pour les estimations de l'enquête :

- Point d'entrée
- Mode d'entrée
- Voie d'entrée
- Date d'entrée
- Date de départ
- But du voyage
- Nombre de personnes faisant partie du groupe de voyage
- Endroits visités
- Si un voyage à forfait a été acheté et les nuitées correspondantes qui sont comprises

Ces zones remplies permettent de lier un questionnaire à un groupe de calage¹ pour imputer les valeurs manquantes. Par exemple, les tarifs de transport manquants ou le total des dépenses engagées durant le voyage sont imputés à l'aide de modèles de régression fondés sur le nombre de personnes faisant partie du groupe de voyage, la durée du voyage, le but du voyage, l'âge des membres des groupes de voyage, l'information sur l'hébergement et le point d'entrée.

L'imputation totale (c.-à-d. l'imputation de questionnaires entiers) est effectuée pour tout groupe de calage pour lequel un nombre insuffisant de questionnaires ont été remplis pour un trimestre donné. Étant donné que les estimations sont générées par groupe de calage, un nombre minimum de questionnaires est requis pour fournir des estimations de qualité pour le groupe de calage. Dans les cas où le nombre minimal n'est pas atteint par les questionnaires reçus du trimestre en cours, les questionnaires sont ajoutés (imputés) au groupe de calage pour le trimestre de référence. Ces questionnaires imputés sont sélectionnés à partir de questionnaires du même trimestre des années d'enquête précédentes. Pour de l'EDA de 2023, les taux d'imputation totale sont faibles, à moins de 1 % des enregistrements.

Étant donné qu'il n'y a pas eu de collecte de données pour l'EDA au cours du premier trimestre de 2023 (janvier à mars), peu de réponses ont été recueillies auprès des voyageurs qui sont arrivés au cours de ce trimestre, et toutes provenaient de voyageurs qui ont quitté le pays au cours des trimestres suivants. L'imputation totale a été appliquée pour ajouter des questionnaires au premier trimestre de 2023, de sorte qu'une année complète de données serait disponible aux fins d'analyse. Les données de l'EDA pour le premier trimestre de 2024 ont servi de base à cette imputation totale pour le premier trimestre de 2023.

4.3 Contrôle de la divulgation

¹ Dans l'Enquête sur les départs aériens, les populations cibles (voyageurs internationaux) sont divisées en groupes de calage en fonction de certaines caractéristiques relatives aux voyageurs, comme le pays de résidence et la durée du séjour.

Tous les répondants au sondage sont assurés que leur identité et leurs réponses demeureront confidentielles et qu'aucun renseignement pouvant permettre de les identifier ne sera divulgué dans le FMGD. Il s'agit d'une exigence de la loi à laquelle se conforment toutes les données de Statistique Canada. Tout renseignement qui pourrait révéler l'identité d'un répondant, à son insu ou sans son consentement écrit, demeure confidentiel et ne peut être divulgué ni publié.

Certaines données sont donc supprimées, perturbées ou entièrement retirées du FMGD afin d'empêcher toute identification directe ou résiduelle des répondants à partir des données diffusées. Les variables ou catégories de variables qui peuvent lier une réponse à une personne sont supprimées, ce qui comprend les variables comportant des catégories choisies par un nombre restreint de répondants ou fournissant des renseignements permettant de les identifier, comme les visites effectuées durant un voyage, qui sont susceptibles d'être attribuées à un répondant en particulier.

Les programmes sont exécutés pour assurer la confidentialité de toutes les données diffusées à la suite des mesures de contrôle de la divulgation qui sont appliquées au FMGD.

5.0 PONDÉRATION

Pour le flux des voyageurs des États-Unis, les poids finaux sont calés afin de correspondre aux totaux communiqués par le programme de dénombrement à la frontière de Statistique Canada qui est fondé sur des données administratives issues des renseignements sur les passages aux postes-frontières fournis par l'ASFC. Les groupes de calage final sont définis par sept provinces canadiennes ou régions d'entrée (Atlantique, Québec, Ontario, Manitoba, Saskatchewan, Alberta et Colombie-Britannique/Nord) ainsi que deux catégories de durée du voyage (de même jour et d'une nuit ou plus), pour un total de 14 groupes.

Pour le flux de voyageurs des pays d'outre-mer, chaque voyageur est codé selon le pays de résidence (ou, dans certains cas, selon le regroupement de pays) afin d'obtenir un code de région géographique type (RGT). L'EDA de 2023 compte un total de 24 régions géographiques types. Les dénombrements de la durée (courte et longue) sont estimés en fonction des données des bornes d'inspection primaire de l'ASFC et des dénombrements à la frontière. Les définitions de « courte » et de « longue » diffèrent selon la RGT.

Les poids initiaux pour le flux de voyageurs des pays d'outre-mer sont calculés pour chaque groupe de RGT par durée (48 groupes de calage). Les poids finaux sont calés afin de correspondre à cinq régions canadiennes par total de durée (10) et total de RGT (24). Il y a 82 groupes de calage au total pour les voyageurs des pays d'outre-mer. Les poids finaux sont calés pour correspondre au total des dénombrements à la frontière pour les groupes de calage final. Il est à noter que certains des groupes de calage final se chevauchent.

6.0 ESTIMATION

Aux fins de l'estimation, les réponses obtenues à partir des questionnaires doivent être traitées comme un échantillon aléatoire simple de l'ensemble des voyageurs de chaque groupe de calage. Les données sont sujettes à un certain degré de « biais de distribution » puisque la répartition des répondants peut être différente de celle de la population. Par exemple, il est possible qu'un nombre important de voyageurs ayant une habitude particulière en matière de dépenses répondent au sondage. Si cette habitude est peu

courante au sein de la population, un biais de distribution des réponses apparaît. Les données sont également sujettes à un certain « biais de non-réponse » puisque les personnes qui choisissent de ne pas répondre peuvent avoir une caractéristique en commun. Il se peut donc que les réponses ne soient pas représentatives de la population de voyageurs.

Les techniques de pondération utilisées dans le processus d'estimation peuvent atténuer l'effet des biais. Les questionnaires sont réorganisés par caractéristiques connues en groupes homogènes. À partir des poids calculés, des estimations peuvent être obtenues pour une gamme variée de caractéristiques des voyageurs. Ces estimations prennent la forme de totaux, de moyennes et de pourcentages et peuvent être obtenues s'il existe un échantillon d'enregistrements suffisant pour le regroupement souhaité de caractéristiques des voyageurs et de la géographie.

7.0 VARIABILITÉ D'ÉCHANTILLONNAGE

L'EDA est fondée sur des données recueillies auprès d'une partie seulement (échantillon) de la population. Il est important de mentionner que les résultats ne sont que des estimations des valeurs « réelles », où les valeurs « réelles » pour la population de voyageurs n'ont pu être obtenues qu'au moyen d'un recensement de tous les voyageurs. Étant donné qu'il existe toujours des différences entre les données-échantillons et les données du recensement, des mesures statistiques ont été élaborées pour indiquer la taille prévue des différences. Ces mesures permettent aux analystes d'évaluer la fiabilité des statistiques produites à partir d'enquêtes-échantillons.

Divers facteurs contribuent à la qualité d'une estimation, comme le biais de distribution ou le biais de non-réponse. La taille de l'échantillon utilisé pour produire une estimation est importante pour les estimations de qualité; si le nombre de répondants présentant les caractéristiques d'intérêt est trop petit, particulièrement par rapport au volume du recensement, alors l'estimation sera moins fiable.

7.1 Qualité des données

La qualité des données est compromise à la fois par l'erreur d'échantillonnage et par l'erreur non liée à l'échantillonnage. Pour les estimations dérivées du FMGD de l'EDA, le niveau de qualité d'une estimation sera déterminé uniquement en fonction de l'erreur d'échantillonnage, puisqu'il est difficile de mesurer l'erreur non liée à l'échantillonnage. Avant de faire des inférences ou de diffuser ou publier toutes estimations établies à partir du FMGD de l'EDA, les utilisateurs devraient premièrement déterminer le niveau de qualité de ces estimations.

Il faut commencer par établir le nombre d'enregistrements qui servent à calculer l'estimation. Généralement, si ce nombre est inférieur à 10, il faudrait considérer l'estimation pondérée comme étant de qualité inacceptable.

Pour les estimations pondérées fondées sur les tailles d'échantillon de 10 ou plus, les utilisateurs devraient calculer une mesure de variabilité, telle que le coefficient de variation (c.v.) de l'estimation et la diffuser en même temps que l'estimation. Les lignes directrices pour la diffusion recommandées par Statistique Canada sont présentées dans la section suivante et devraient être appliquées aux estimations pondérées arrondies.

On peut considérer qu'il est possible de divulguer toutes les estimations. Celles d'un niveau de qualité médiocre ou inacceptable devraient cependant être accompagnées d'une mise en garde pour avertir les futurs utilisateurs. Idéalement, les estimations de faible qualité ou de qualité inacceptable ne devraient pas être publiées. Lorsqu'elles le sont toutefois, elles doivent être accompagnées d'un avis de non-responsabilité.

7.2 Lignes directrices pour la diffusion de coefficients de variation

Les catégories de qualité recommandées en fonction du c.v. de l'estimation sont Excellente (A), Très bonne (B), Bonne (C), Acceptable (D) et À utiliser avec prudence (E).

Catégorie A : Excellente

Si le c.v. d'une estimation se situe entre 0,0 % et 5,0 %, il est considéré comme étant d'excellente qualité. Aucun avertissement ni aucune restriction ne seront associés à l'estimation dans les données.

Catégorie B : Très bonne

Si le c.v. d'une estimation se situe entre 5,01 % et 15,0 %, il est considéré comme étant de très bonne qualité. Aucun avertissement ni aucune restriction ne seront associés à l'estimation dans les données.

Catégorie C : Bonne

Si le c.v. d'une estimation se situe entre 15,01 % et 25,0 %, il est considéré comme étant de bonne qualité. L'estimation est diffusée, car elle peut être utile à certaines fins, mais elle est affichée accompagnée d'un avertissement concernant l'exactitude de l'estimation.

Catégorie D : Acceptable

Si le c.v. d'une estimation se situe entre 25,01 % et 35,0 %, il est considéré comme étant de qualité acceptable. L'estimation est diffusée, car elle peut être utile à certaines fins, mais elle est affichée accompagnée d'un avertissement concernant l'exactitude de l'estimation.

Catégorie E : À utiliser avec prudence

Si le c.v. d'une estimation se situe à 35,01 % ou plus, il est considéré comme étant de qualité inacceptable. L'estimation ne devrait pas être diffusée, car elle ne respecte pas les normes de qualité de Statistique Canada. Toute conclusion tirée à partir d'une estimation de qualité inacceptable est peu fiable et probablement invalide.

Toutefois, si un utilisateur insiste pour inclure des estimations jugées de qualité inacceptable, les estimations doivent être affichées accompagnées d'un avis de non-responsabilité. L'utilisateur doit également toujours inclure l'avis de non-responsabilité si ces estimations sont incluses (directement ou indirectement) dans les données diffusées, déclarées ou présentées. Toute estimation calculée à partir de moins de 10 enregistrements est également considérée comme d'une qualité inacceptable, quel que soit son c.v. correspondant.

7.2 Méthode bootstrap pour estimer la variance

En raison de la complexité du plan et du traitement de l'EDA, les calculs simples de variance réalisés à l'aide de méthodes normalisées ont tendance à sous-estimer la variance, ce qui conduit à des estimations présentant une précision surestimée. Pour l'EDA, Statistique Canada utilise la méthode bootstrap pour estimer la variance. Grâce à cette méthode, les répondants sont sous-échantillonnés et pondérés à nouveau de nombreuses fois (l'EDA utilise actuellement 500 sous-échantillons bootstrap). À partir de ces sous-échantillons, une variance estimée est obtenue pour calculer le c.v.

Les fichiers de pondération bootstrap sont fournis avec le FMGD en format CSV, afin que les utilisateurs puissent calculer des mesures de variabilité pour les estimations qu'ils dérivent du FMGD. L'annexe C donne un aperçu de la façon d'utiliser les poids bootstrap pour estimer la variance.

8.0 CONTENU DU FICHIER DE MICRODONNÉES À GRANDE DIFFUSION (FMGD)

Le FMGD de l'EDA contient des renseignements sur les habitudes de voyage des visiteurs (résidents des États-Unis et des pays d'outre-mer) qui entrent au Canada ou en sortent au moyen de vols commerciaux. Les caractéristiques suivantes sont disponibles dans le FMGD de l'EDA :

Caractéristiques du voyage :

- Trimestre d'entrée
- Province ou région d'entrée
- Mode de transport utilisé pour entrer au Canada et en sortir, et modes utilisés pendant le séjour au Canada
- But principal du voyage
- Activités
- Frais de transport international (tarifs payés à tous les transporteurs)
- Total des dépenses engagées durant le voyage (excluant les tarifs)
- Dépenses par catégorie
- Nombre de nuits passées au Canada
- Nombre de nuits passées selon la province
- Nombre de visites-personnes selon la province
- Lieu de résidence (État américain ou groupement d'États, et pays ou groupement de pays)
- Nombre de personnes faisant partie du groupe de voyage
- Poids des enregistrements

Caractéristiques démographiques :

- Composition des groupes de voyageurs selon l'âge et le sexe

Endroits visités :

- Endroit visité (province, région métropolitaine de recensement, région touristique) (maximum de 10)
- Durée du séjour sur place visité (nombre de nuits passées)
- Hébergement utilisé à l'endroit visité

Une liste complète des variables du FMGD et de leurs fréquences se trouve dans le recueil de codes.

Il est à noter que, dans le cadre des mesures de contrôle de la divulgation dont il est question à la section 4.3, l'ordre des visites de chaque voyage est habituellement rajusté; les utilisateurs devraient donc considérer l'ordre des visites 1 à 10 dans le fichier comme ayant peu de valeur analytique. Les utilisateurs doivent concentrer leur analyse sur les estimations globales des endroits visités, indépendamment de leur ordre durant le voyage, pour lesquelles un effort notable est déployé dans le processus de création du FMGD afin de préserver l'utilité analytique.

Annexe A : Liste des régions géographiques types pour le lieu de résidence et leur constitution pour les résidents des pays d'outre-mer

Consultez le fichier EDA_FMGD_2023_GuideUtilisateur_Annexe.pdf pour obtenir la liste des pays inclus dans chaque région géographique type.

Annexe B : Groupements d'États de résidence américains utilisés dans le FMGD de l'Enquête sur les départs aériens de 2023

Reportez-vous au fichier EDA_FMGD_2023_GuideUtilisateur_Annexe.pdf pour connaître les regroupements d'États de résidence américains qui ont été utilisés dans le FMGD de l'EDA de 2023, dont une liste des États qui constituent les catégories régionales « autres », s'il y a lieu.

Annexe C : Estimation de la variance bootstrap pour les FMGD

Pour déterminer la qualité des estimations et calculer le coefficient de variation (c.v.), on doit d'abord établir l'écart-type. On a aussi besoin de l'écart-type d'estimations pour les intervalles de confiance. L'Enquête sur les départs aériens (EDA) utilise des plans d'échantillonnage et d'étalonnage complexes et, par conséquent, il n'y a pas de formule simple d'estimation de variance qu'on puisse appliquer. Par conséquent, la méthode bootstrap d'estimation de la variance, une technique de pseudo-réplication, a été utilisée pour mesurer la qualité des estimations de l'EDA. Grâce aux poids bootstrap et à l'un des ensembles statistiques disponibles pour l'estimation de la variance, les c.v. et d'autres estimations de la variance peuvent être calculés avec exactitude.

Les poids bootstrap sont fournis avec le FMGD en format CSV.

Méthode bootstrap pour estimer la variance

La méthode bootstrap moyenne a été utilisée pour l'EDA et son FMGD. Pour chaque unité de l'échantillon, un ensemble de 500 poids bootstrap finaux est généré à l'aide de méthodes de réplication, chacun de ces poids bootstrap étant le résultat d'une moyenne de 20 sous-répliques bootstrap.

La variation entre les 500 estimations possibles fondées sur chacune des répliques est liée à la variance de l'estimateur en fonction des poids d'enquête et peut être utilisée pour l'estimer.

Les utilisateurs doivent noter que les estimations de la variance provenant des bootstraps du FMGD peuvent être légèrement différentes des estimations de la variance diffusées par Statistique Canada pour de l'EDA, car certains petits ajustements ont été apportés pour protéger la confidentialité des répondants.

Progiciels statistiques pour l'analyse statistique et l'estimation de la variance

Il existe divers progiciels commerciaux ou de source ouverte qui permettent aux utilisateurs d'utiliser des poids de réplication bootstrap pour l'estimation de la variance.

Utilisation des poids bootstrap

Par exemple : Estimation des dépenses

Supposons qu'un utilisateur souhaite estimer le montant que les résidents du Royaume-Uni voyageant par vol commercial ont dépensé pour le voyage au Canada en 2023 (β).

L'utilisateur peut estimer β en multipliant les dépenses totales (VGLTOTSP) pour chaque voyage effectué par des résidents du Royaume-Uni (VSGRCDEP = 101) par le poids voyage correspondant (VWEIGHTP), puis en additionnant ces dépenses pondérées. Le résultat est le suivant :

$$\hat{\beta} = 1\,278\,892\,283 \$.$$

Ce processus peut être répété de la même manière, sauf en substituant le poids voyage (VWEIGHTP) par le $b^{\text{ième}}$ poids bootstrap, où $b=1, \dots, 500$ (BSW1, ..., BSW500).

Il en résulte les estimations bootstrap (1), ..., $\hat{\beta}(500)$ de β qui sont présentées dans le tableau suivant :

Poids bootstrap	Estimation bootstrap	Somme pondérée des dépenses
BSW1	$\hat{\beta}(1)$	1 279 417 196 \$
BSW2	$\hat{\beta}(2)$	1 282 117 869 \$
BSW3	$\hat{\beta}(3)$	1 271 792 287 \$
BSW4	$\hat{\beta}(4)$	1 298 285 134 \$
BSW5	$\hat{\beta}(5)$	1 282 345 260 \$
...		...
BSW500	$\hat{\beta}(500)$	1 285 229 906 \$

Ensuite, l'estimation bootstrap habituelle de la variance de $\hat{\beta}$ est donnée par :

$$\hat{V}_B(\hat{\beta}) = \frac{1}{500} \sum_{b=1}^{500} (\hat{\beta}_{(b)} - \hat{\beta})^2 = 9,0713506 \times 10^{13}$$

De là, l'écart-type peut être estimé comme suit : $\hat{S}_B(\hat{\beta}) = \sqrt{\hat{V}_B(\hat{\beta})} = 9\,524\,364$

et le coefficient de variation par $CV_B(\hat{\beta}) = \frac{\hat{S}_B(\hat{\beta})}{\hat{\beta}} \times 100 = 0,745\%$.