

Exploiter les données pour l'analyse d'impact : Module 4

Shelley Jeglic, Division de l'accès aux données

21 mars 2024



Éclairer grâce aux données, pour bâtir un Canada meilleur



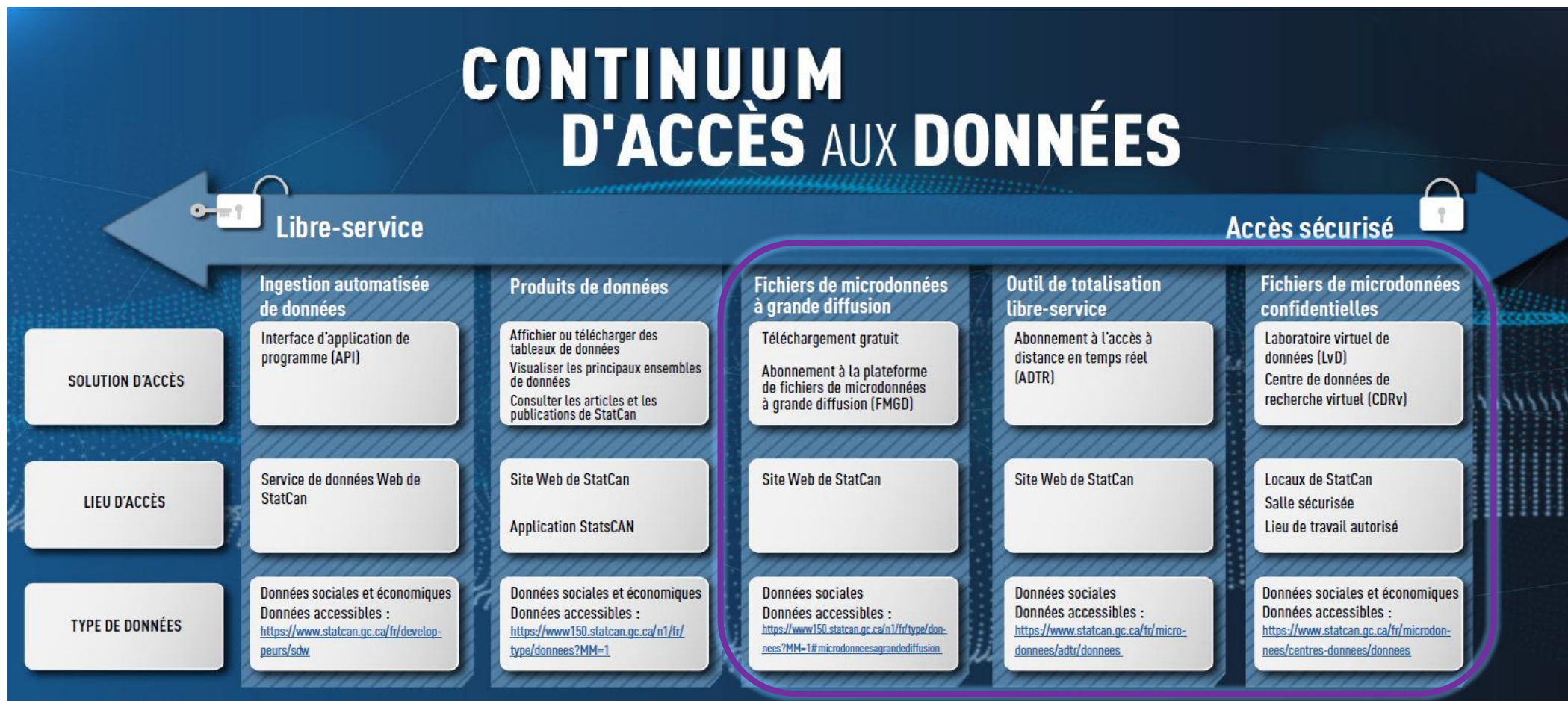
Statistique
Canada

Statistics
Canada

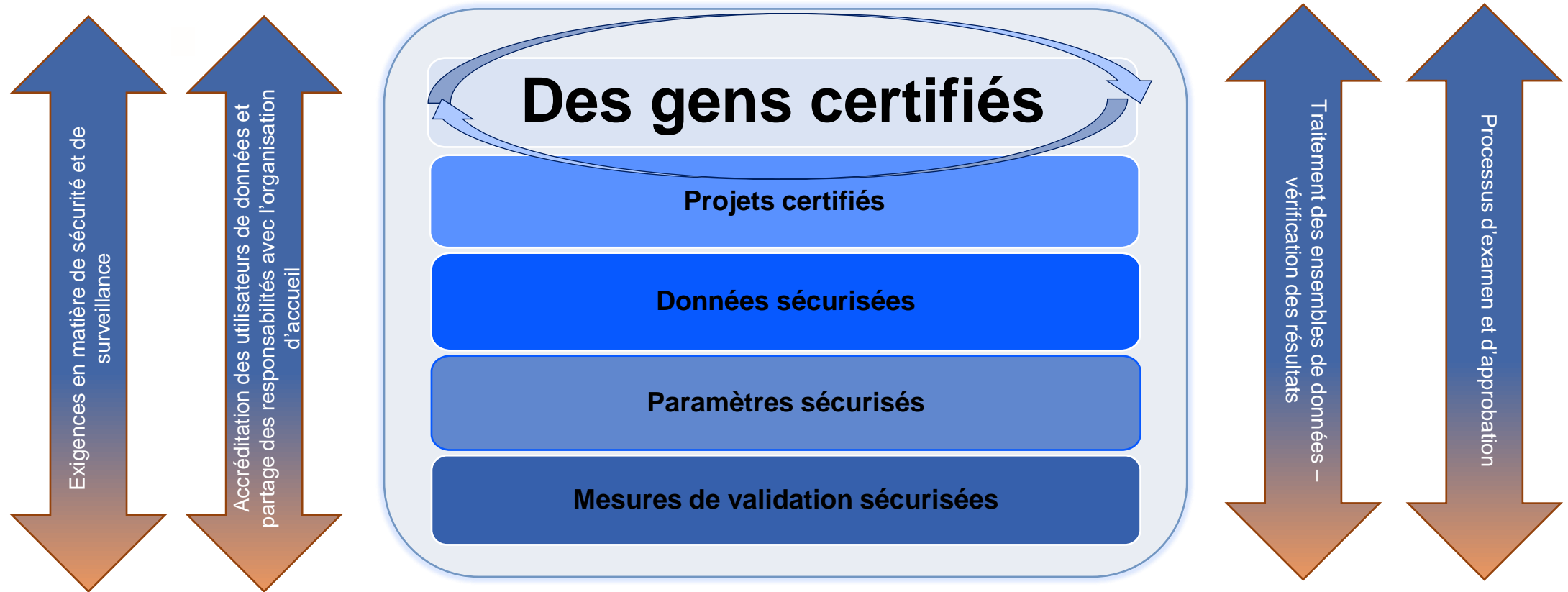
Canada

La Division de l'accès aux données de StatCan

Ce que nous offrons : des données accessibles, représentatives et utiles sur les Canadiens, pour les Canadiens



Utilisation du cadre des cinq éléments de la sécurité comme principe directeur



Les cinq éléments sont pris en compte ensemble – selon une échelle mobile – en fonction du type de chercheur, de l'accès et des données.

Fichiers de microdonnées à grande diffusion

Accès



- Les abonnés à la **collection de fichiers de microdonnées à grande diffusion (FMGD)** se voient accorder l'accès à toutes les données anonymisées et non agrégées de la collection de FMGD

Idéaux pour



- Exploration et extraction de données, outil analytique pour la recherche universitaire et stratégique :
 - calcul des fréquences, tableaux croisés, moyennes, centiles, répartition en pourcentage, proportions, ratios et parts
- Utilisation comme outil de formation analytique

Avantages



- Recherches approfondies avec des résultats sélectionnés et arrondis sous forme de fichier
- Utilisation d'une licence ouverte
- Principal produit utilisé pour enseigner la littératie des données
- Modification et examen prudents de toutes les données à des fins de recherche

Frais : Collection de FMGD



- Frais d'accès institutionnels annuels (accès autorisé par la reconnaissance de la plage IP)
- **10 000\$ par année**

Accès aux FMGD

Tous les nouveaux FMGD peuvent désormais être téléchargés gratuitement à partir du site web de StatCan.

- Les anciens FMGD seront mis à disposition progressivement.

Programme de collection de FMGD

- L'abonnement est disponible pour les ministères et les entreprises à l'intérieur et à l'extérieur du Canada.
- L'accès illimité à toutes les microdonnées et à la documentation est disponible par le biais du Service de transfert électronique de fichiers (STEF) de StatCan.
- Certains fichiers sont disponibles via la plateforme analytique *Rich Data Services (RDS)*.

Programme de l'Initiative de démocratisation des données (IDD)

- L'adhésion est offerte aux établissements d'enseignement postsecondaire.
- L'accès illimité à toutes les microdonnées et à la documentation est disponible par le biais du Service de transfert électronique de fichiers (STEF) de StatCan.
- Certains fichiers sont disponibles via la plateforme analytique *Rich Data Services (RDS)*.



Accès aux FMGD: *Rich Data Services (RDS)*

Particularités

- [RDS Landing \(statcan.gc.ca\)](https://statcan.gc.ca)
- Base de données en ligne libre-service restreinte aux IP
- Interface web conviviale, bilingue et accessible
- Fichiers téléchargeables dans différents formats
- Compatible avec l'interface de programme d'application (*API*)
- Conforme à l'initiative de documentation des données (*DDI*)

RDS Explorer

- Utilisé pour parcourir les enregistrements de données, créer des extraits personnalisés en filtrant les enregistrements et en sélectionnant des variables, et produire des ensembles personnalisés de données disponibles au téléchargement.

Moteur de tabulation *RDS*

- Utilisé pour agréger les données à travers différentes dimensions afin de créer rapidement des tableaux analytiques.

Accès à distance en temps réel

Accès



- L'outil d'accès à distance en temps réel (ADTR) est un outil de totalisation en ligne qui permet d'accéder rapidement aux microdonnées par l'intermédiaire des programmes du système d'analyse statistique
- Il s'agit d'un outil accessible par abonnement

Idéal pour



- Produire une gamme complète de statistiques descriptives qui peuvent être utilisées dans le cadre de recherches, de formations et de séances d'information sur les politiques
- Calculer les fréquences, les moyennes, les centiles, les proportions, les ratios et les parts

Avantages



- Permet d'effectuer des analyses de statistiques descriptives
- Permet d'accéder rapidement aux totalisations diffusables
- Ne permet pas aux utilisateurs d'accéder directement aux microdonnées ni de consulter le contenu des fichiers de microdonnées

Frais



- **10 000\$/année (maximum de 10 comptes d'utilisateurs)**
- **Comptes d'utilisateurs supplémentaires 2 000\$**

Fichiers de microdonnées confidentielles

Accès



- Accès à partir d'espaces de travail autorisés par l'entremise de l'infonuagique depuis un bureau ou une résidence personnelle
 - Centre de données de recherche virtuel pour les chercheurs universitaires
 - Laboratoire virtuel de données pour le gouvernement et d'autres organisations
- Comprend les centres de données de recherche et le Centre fédéral de données de recherche, ainsi que les salles sécurisées

Avantages



- Soutien spécialisé et accès direct à des fichiers de microdonnées confidentielles et détaillées
- Accès à distance sécurisé à des données anonymisées 24 heures sur 24, 7 jours sur 7 pour les chercheurs, sous réserve de certains critères, grâce à une infrastructure informatique virtuelle sécurisée
- Possibilités accrues de collaboration

Idéaux pour



- Répondre à des questions de recherche qui nécessitent l'utilisation de méthodes analytiques avancées comme l'analyse multivariée complexe ou la modélisation

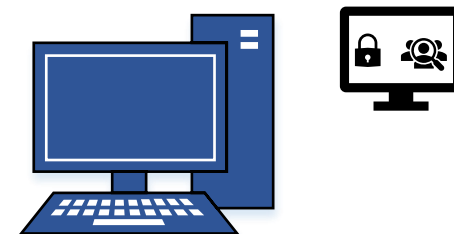
Frais



- Les frais d'accès varient en fonction des paramètres du projet

Accès aux microdonnées : Laboratoire virtuel de données

- Le laboratoire virtuel de données (LvD) est une infrastructure en nuage sécurisée utilisée pour stocker et accéder aux microdonnées des projets des chercheurs.
- Le laboratoire virtuel de données (LvD) offre aux chercheurs une approche plus souple de l'accès aux microdonnées de Statistique Canada, qui comprend des contrôles de surveillance et une gestion intégrée du filtrage.
- En utilisant des « **espaces de travail autorisés** » pour les projets admissibles, les chercheurs admissibles peuvent accéder à des projets de microdonnées tout en télétravaillant à domicile ou à partir de leur propre bureau.
- Le LvD a remplacé la nécessité d'une salle sécurisée sur place et/ou d'ordinateurs portables fournis par StatCan pour l'accès aux données.



Espace de travail autorisé

Un bureau fermé ou une salle de conférence dans les locaux d'une organisation, ou un espace privé dans la résidence personnelle de l'employé réputé.

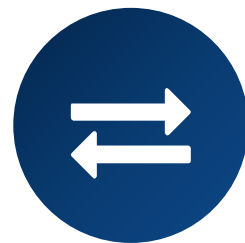
Nouveauté – Laboratoire de données virtuelles (LvD)



**Accès à distance
protégé 24 heures sur
24, 7 jours sur 7**



**Mesures de sécurité
informatique robustes**



**Approche à
risque partagé**



**Hébergé dans un
environnement
d'analyse collaborative**



**Importation et
exportation sécurisées
à destination et en
provenance de
l'environnement**



**Authentification à
deux facteurs**



**Capacité de
surveillance et
de production de
rapports**



**Remplace le
recours à une
salle sécurisée**



Accès aux microdonnées : Présentation de nouvelles propositions de projets de microdonnées

Les demandes d'accès aux microdonnées sont soumises à moyen du Portail d'accès aux microdonnées (PAM) : <https://www.statcan.gc.ca/rdc-cdr/fra/user/login>

Le processus de présentation d'une nouvelle demande comprend les étapes suivantes :

- Créer un compte dans le PAM;
- Déterminer le type de projet et fournir des détails sur le projet (notamment la proposition de projet, exposée conformément aux [lignes directrices sur le modèle de proposition de projet](#));
- Rédiger les profils des cochercheurs;
- Soumettre la demande.

Si vous avez besoin d'aide pour remplir une demande, veuillez écrire à l'adresse suivante : statcan.dad-apu-dad-uta.statcan@statcan.gc.ca



Exigences pour les organismes concernant l'accès au LVD

Organismes admissibles

- Doivent signer une entente
 - Entente en vertu de l'article 10 pour les ministères provinciaux
- Doivent être un organisme accrédité (connu et digne de confiance)
 - Documents d'accréditation d'organisme remplis
- Doivent nommer un agent de liaison pour l'accès aux données



Exigences pour les chercheurs concernant l'accès au LVD

Chercheurs admissibles

- Devenir une personne réputée être employée de Statistique Canada
 - Prêter ou affirmer le serment ou une affirmation d'office et de discrétion
 - Obtenir une cote de sécurité de fiabilité
- Suivre la formation obligatoire et réussir le test obligatoire
- Être agréé (l'affiliation à un organisme agréé augmente le niveau)
- Ne pas présenter de conflit d'intérêts lié à l'accès aux données
- Recevoir l'approbation de la proposition de projet de recherche sur les microdonnées
- Signer un Contrat de recherche pour l'utilisation de microdonnées avec Statistique Canada



Coûts du Laboratoire virtuel de données

Facturation

La facturation est effectuée par confirmation d'achat (moins de 20 000 dollars par) ou par lettre d'entente (20 000 dollars ou plus).

Coûts

Le Centre fédéral de données de recherche (CFDR) fonctionne selon le principe du recouvrement des coûts. Tous les frais comprennent la mise en place du projet, la création du compte, les coûts d'hébergement du projet dans le nuage, l'assistance de l'analyste et le contrôle des résultats pour 12 mois d'accès continu.

Projet de base : 13 500\$

- 1 à 3 chercheurs accédant à un "espace de travail autorisé" dans le laboratoire virtuel de données (LvD) ou sur place dans le Centre fédéral de données de recherche (CFDR)
- 12 mois d'accès continu
- Un maximum de 1 440 heures d'accès (à travers les machines virtuelles (MV) du LvD ou les chercheurs sur site)
- Suite logicielle standard
 - Project Libre (logiciel libre équivalent à la suite MS Office)
 - STATA MP64
 - IDLE (Python 3.9)
 - Python 3.9
 - SAS Enterprise Guide 8.1
 - R 64 4.1.0
 - IBM SPSS Statistics 24
 - Libre Office
 - R Studio 1.4.1717
 - SAS 9.4 (Fr+En)
 - Miniforge3 (pour remplacer Anaconda dans un futur proche)

Prolongation du projet : 13 500 \$ par an ou 1 125 \$ par mois

- S'applique uniquement aux projets de plus de 12 mois
- 1 à 3 chercheurs LvD ou chercheurs CFDR sur site
- Un maximum de 120 heures d'accès par mois (à travers les machines virtuelles (MV) LvD)
- Suite logicielle standard
- Des chercheurs supplémentaires peuvent également être prolongés à un coût par chercheur de 2 500 \$ (12 mois) ou 208 \$ (par mois).

Remarque : Les prolongations de projet de moins de 12 mois doivent se terminer à la fin de la période de prolongation ou des frais supplémentaires peuvent être encourus pour des prolongations multiples au cours de la même année fiscale.

Compléments de projet :

- 2 500 \$ par chercheur supplémentaire ayant accès au CFDR sur site ou à la LvD (12 mois d'accès)
- 500 \$ pour le remplacement ou le changement d'un chercheur MV (pour les projets ayant atteint leur nombre maximal de chercheurs)
- 500 \$ par participant au projet (chercheurs sans MV) pour les projets comportant 3 MV ou plus.
- Logiciels non standard - sur la base du coût des logiciels
- Puissance informatique supplémentaire - en fonction des besoins informatiques
- Services/fonctionnalités supplémentaires selon les besoins

Tarifs en vigueur à partir du 1er avril 2023. Les taxes applicables sont en sus.

Protéger la confidentialité des données

Contrôle de la confidentialité

Le processus de vérification de produits statistiques pour s'assurer qu'elles ne présentent pas de risque de divulgation pour la confidentialité des données.

Règles de contrôle

Ensemble de stratégies visant à minimiser les risques liés à la divulgation. Par exemple,

- *Observations contributives minimales
- *Application des pondérations de l'enquête
- *Arrondissement des effectifs et des valeurs en \$.
- *Limitation aux niveaux géographiques inférieurs
- *Tests de dominance

Délais de contrôle

De nombreux facteurs sont à prendre en compte :

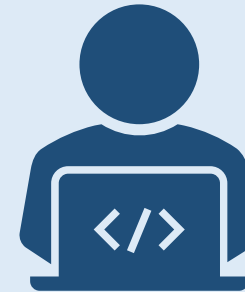
- *Documentation complète et précise
- *Taille et complexité de la demande
- *Version provisoire ou définitive
- *Demandes de produits statistiques antérieures

Secrétariat du Conseil du Trésor du Canada liaison pour l'accès aux données de StatCan

La liaison pour l'accès aux données du Secrétariat du Conseil du Trésor du Canada est le principal contact de Statistique Canada pour la communication d'informations relatives à l'accès aux microdonnées de leur ministère.

Diana Macecek

Diana.macecek@tbs-sct.gc.ca



Coordonnées des personnes-ressources

Si vous avez des questions au sujet de ces produits, veuillez consulter les sources appropriées ci-dessous :

Fichiers de microdonnées à grande diffusion

Collection de FMGD : <https://www150.statcan.gc.ca/n1/fr/type/donnees?MM=1#microdonneesagrandediffusion>

Demandes de renseignements ou d'assistance : statcan.fmgd-pumf.statcan@canada.ca

Accès à distance en temps réel

Liste des données disponibles : <https://www.statcan.gc.ca/fr/microdonnees/adtr/donnees>

Demandes de renseignements ou d'assistance : statcan.dad-rtra-dad-adtr.statcan@statcan.gc.ca

Centres de données de recherche

Liste des données disponibles : <https://www.statcan.gc.ca/fr/microdonnees/centres-donnees/donnees>

Demandes de renseignements ou d'assistance : statcan.dad-apu-dad-uta.statcan@statcan.gc.ca



Merci

Geneviève Jourdain

Directrice

Division de l'accès aux données

genevieve.jourdain@statcan.gc.ca

Shelley Jeglic

Directrice adjointe

Division de l'accès aux données

shelley.jeglic@statcan.gc.ca

