



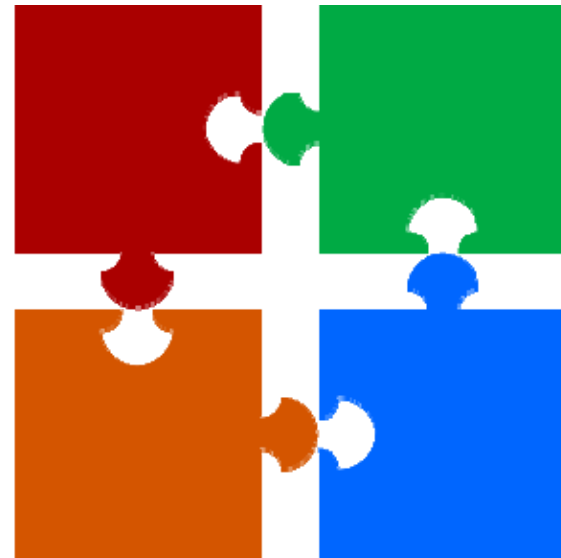
Statistique Canada  
[www.statcan.gc.ca](http://www.statcan.gc.ca)

# Intégration des données sociales

## Mise en valeur des fonds de données

**Lynn Barr-Telford**

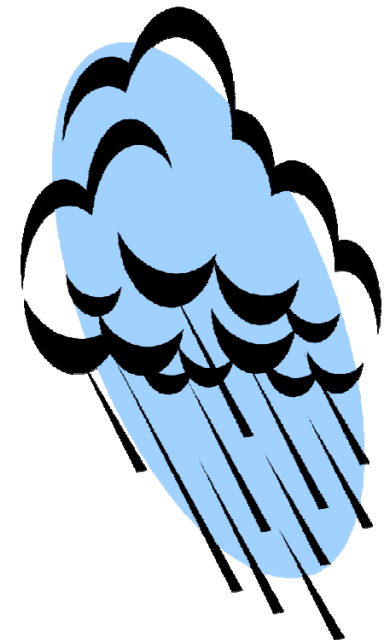
30 novembre, 2016





## Environnement actuel

- Budgets en baisse
- Taux de réponse aux enquêtes qui diminuent
- Fardeau du répondant
- Préoccupations concernant la protection des renseignements personnels
- Nécessité de considérer des solutions de rechange





## Comment répondre à l'environnement actuel

- Maximiser l'utilisation des données administratives
- Établir les priorités en ce qui concerne le contenu des enquêtes
- Intégration des données





## Stratégies à utiliser pour mettre en valeur les fonds de données

- Accroître l'utilisation des données administratives**
  - Pour compléter les activités de collecte de données actuelles
  - Pour réduire le fardeau du répondant
  
- Favoriser la création de partenariats afin d'éviter le chevauchement des efforts**
  
- Intégration des données**
  - Accroître le nombre de projets de couplage d'enregistrements dans divers domaines sociaux (santé, éducation, justice, revenu, etc.)
  - Produire de nouveaux renseignements sans collecte de données additionnelle
  - Améliorer la compréhension des déterminants sociaux et des résultats à long terme
  
- Accès aux fichiers couplés**
  - Réseau d'accès (CDR, CANSIM, Accès à distance, etc.)



## Possibilités en matière d'intégration des données

- ❑ Les fichiers longitudinaux existants demeurent pertinents :
  - Enquête nationale sur la santé de la population
- ❑ Comblir les lacunes statistiques actuelles :
  - Ajout des données sur le traitement au Registre canadien du cancer
- ❑ Examiner les possibilités d'élaboration de données:
  - Analyse des transitions du domicile aux établissements de soins (données de l'Enquête sur la santé dans les collectivités canadiennes et du Système d'information sur les soins de longue durée
  - Étudier le lien entre la santé mentale et le système de justice
- ❑ Améliorer la compréhension des sous-groupes de population :
  - Recensement : Autochtones, immigrants, familles à faible revenu



# Couplage de données : « Ce que Statistique Canada a à apporter »

## Expertise

- Chef de file dans l'élaboration de méthodes et de logiciels de couplage d'enregistrements (SGCE, G-Coup)

## Cadre législatif

- Chaque couplage d'enregistrements doit se faire conformément à la *Loi sur la statistique* et à la *Loi sur la protection des renseignements personnels*
- Les avantages doivent nettement dépasser l'atteinte à la vie privée

## Transparence

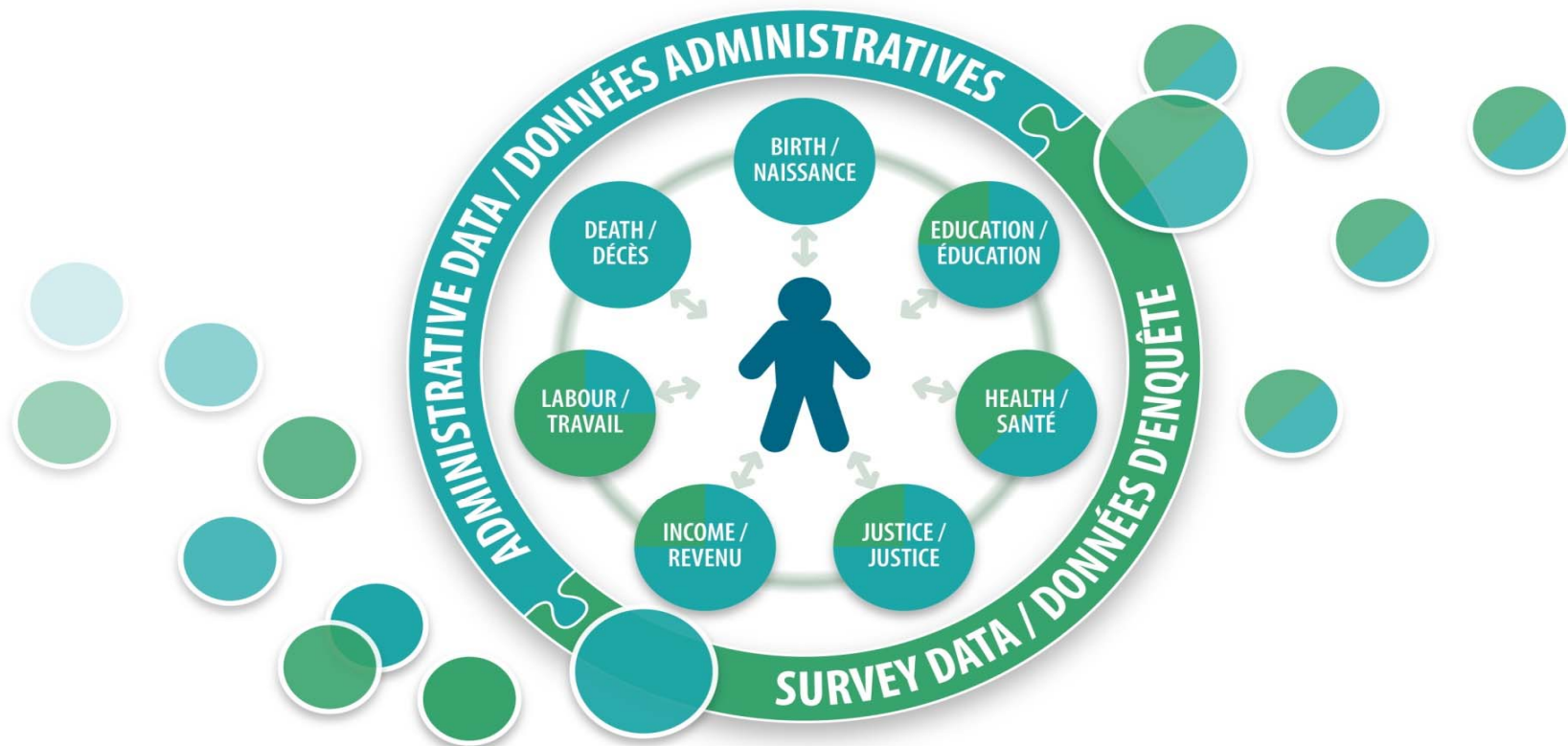
- Rapport annuel au Parlement
- Résumé de tous les couplages d'enregistrements approuvés accessible sur le site Web

## Accessibilité

- Aux données couplées (la plupart) dans les centres de données de recherche
- Sans frais... juste quelques formalités administratives!



# Environnement de couplage de données sociales





## Prolongation du projet de pertinence

Améliorer la valeur analytique de cinq fichiers longitudinaux terminés :

- Enquête auprès des jeunes en transition (EJET)
- Enquête nationale sur la santé de la population (ENSP)
- Enquête sur la dynamique du travail et du revenu (EDTR)
- Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ)
- Enquête longitudinale auprès des immigrants du Canada (ELIC)

Au moyen du couplage d'enregistrements à un ensemble d'indicateurs de sortie sélectionnés à partir d'une série de sources de données de Statistique Canada.





## Fichiers à coupler:

Fichier T1 sur les familles de 1993 à 2011 (34 éléments de données)

- Renseignements détaillés sur le revenu
- Mesure de l'incapacité
- Déductions pour l'éducation
- Constitution d'une famille, mobilité, état matrimonial

Registre canadien du cancer de 2011 (4 éléments de données)

- État de santé

Statistique de l'état civil – Décès de 1993 à 2011 (2 éléments de données)

- État de santé



## Fichiers à coupler:

### Recensement 2006 et Recensement 2006 2B (15 éléments de données)

- Situation d'emploi
- Indicateur d'incapacité
- Plus haut niveau de scolarité
- Renseignements démographiques sur l'état matrimonial, la constitution d'une famille et la situation relative au logement
- Mobilité

### Recensement de 2011 et Enquête nationale auprès des ménages (15 éléments de données)

- Situation d'emploi
- Indicateur d'incapacité
- Plus haut niveau de scolarité
- Renseignements démographiques sur l'état matrimonial, la constitution d'une famille et la situation relative au logement
- Mobilité



Un éventail de plus en plus large de données administratives sont intégrées aux fins de la recherche longitudinale. Par exemple...

**Emplois**

Feuillets T4  
Relevés d'emploi  
Données sur les prestataires d'assurance-emploi  
Données des régimes de pension

**Revenu**

Déclarations de revenus des particuliers T1  
Fichier sur la famille T1  
Données sur l'aide sociale

**Entreprises commerciales**

Déclarations de revenus des sociétés T2  
Registre des importations et exportations  
etc.

**Santé**

Statistiques de l'état civil  
Registre du cancer  
Données sur les hospitalisations  
Données sur les personnes âgées vivant dans des établissements de soins  
Données sur la mortalité

**Immigration**

Données sur les résidents temporaires  
Fichier des immigrants reçus

**Éducation**

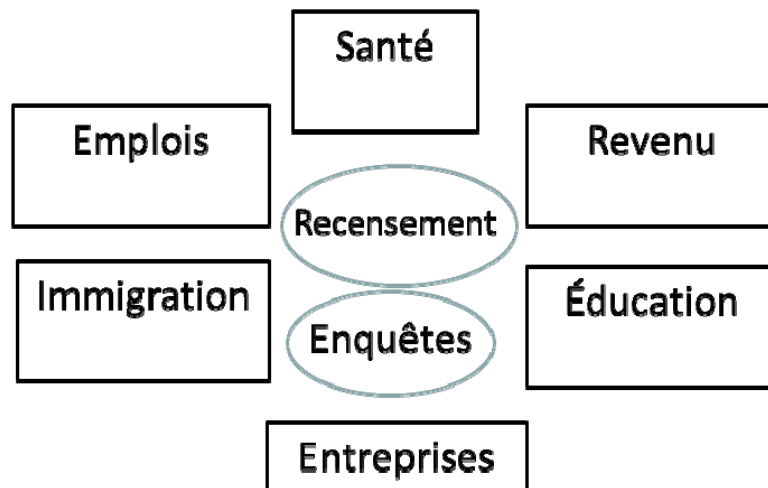
Données sur les programmes d'apprentissage  
Données sur les inscriptions aux études postsecondaires  
Données sur les prêts aux étudiants

- Couplages à l'intérieur des domaines et entre domaines
- Populations quasi universelles de travailleurs, de patients, de déclarants fiscaux, etc.
- Possibilité de recherches de plus en plus granulaires, dans les limites imposées par les variables disponibles
- Transition paradigmatique de « fichiers couplés » vers des « environnements couplés »
- La disponibilité des données continue de s'accroître

+ autres fichiers et autres domaines



## Les données du Recensement et les données d'enquête sont intégrées à l'environnement de données administratives



Les fichiers du recensement ajoutent des données transversales portant sur une année unique, qui n'ont que des répercussions modérées sur la taille de l'échantillon

- Ces données servent généralement à l'analyse des cohortes prospectives à l'aide de caractéristiques invariables (ou quasi invariables) dans le temps à l'année de recensement T et des résultats fondés sur des données administratives à l'année T+n.
- Le chevauchement entre le questionnaire 2B du Recensement de 2006 et celui de l'ENM de 2011 donne un échantillon d'environ 600 000 répondants ayant donné leur consentement.
- L'intégration planifiée des données transversales des questionnaires complets des recensements de 1991 à 2016 et des données administratives à l'aide d'une approche cohérente augmentera les possibilités de recherche longitudinale entre cohortes.

Les enquêtes auprès des ménages permettent d'ajouter des caractéristiques transversales et/ou longitudinales aux fichiers administratifs couplés, mais réduisent de beaucoup la taille de l'échantillon.



## Exemples de fichiers couplés et recherches longitudinales qui les utilisent

### Santé

Fichier: Données sur les hospitalisations + Données sur la mortalité + T1/T1FF + T4

Recherche: Répercussions d'un AVC, d'un IAM ou d'un arrêt cardiaque sur l'emploi et les revenus personnels et sur ceux du conjoint

Fichier: ESCC de 2005 à 2009 + Recensement de 2011 + Données sur la mortalité

Recherche: Transition des personnes âgées vers des établissements de soins

### Immigration

Fichier: Fichier des immigrants reçus + T1 + T2 et fichiers connexes (p. ex. Annexe 50)

Recherche: Création d'entreprises par des immigrants; survie et résultats des entreprises appartenant à des immigrants; effets de l'entreprise sur les revenus des immigrants

Fichier: Fichier des résidents temporaires + Fichier des immigrants reçus + T1/T1FF + T4

Recherche: Transitions d'immigrant temporaire à immigrant permanent; caractéristiques avant l'établissement et résultats économiques après l'établissement

### Éducation

Fichier: Apprentis inscrits + T1 + T4 + Relevés d'emploi + Données au niveau de l'entreprise

Recherche: Achèvement de la formation d'apprenti, mobilité géographique, retours aux études, emploi au niveau de l'entreprise

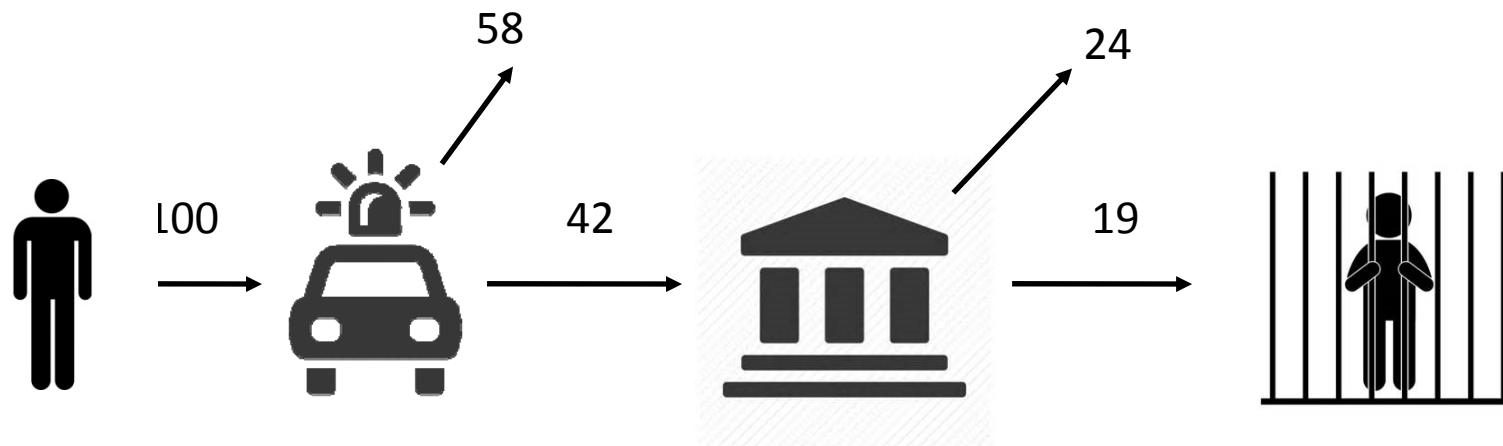
### Emploi

Fichier: Données au niveau de l'entreprise + Données sur les cessations d'emploi + Recensement de 1991 + Données sur la mortalité

Recherche : Cessations d'emploi et mortalité



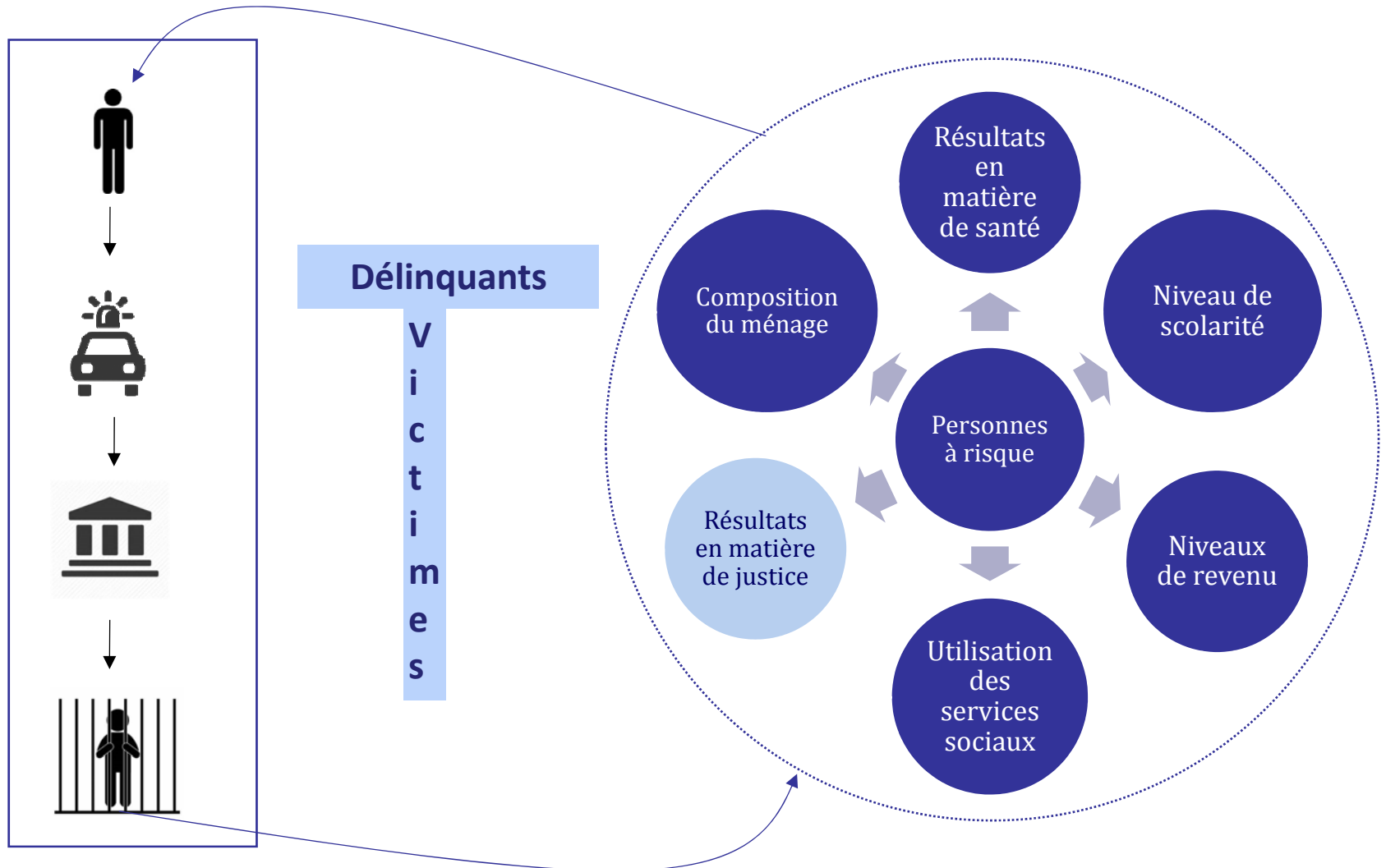
## Progrès récents : cheminements dans le système de justice pénale



Parmi les quelques 37 000 personnes qui sont entrées en contact avec les services policiers de la Saskatchewan en 2009-2010, 42 % ou un peu plus de 15 700 personnes ont fini par comparaître devant un tribunal de la Saskatchewan. De ce nombre, environ 6 900 personnes (ou 19 %) ont cheminé du contact avec la police jusqu'aux services correctionnels.



## Les prochaines étapes, au-delà du système de justice pénale





## Vision pour la plateforme de couplage longitudinal en éducation (PCLE)

- Le SIEP ne fournit, pour l'instant, qu'un portrait transversal du monde de l'enseignement postsecondaire.
- La PCLE est créée à partir du **couplage des données administratives** provenant du SIEP et du Système d'information sur les apprentis inscrits (SIAI).
- Ces données sont couplées longitudinalement entre elles, de même qu'avec d'autres sources de données et seront mises à jour annuellement.
- Le SIEP, dans le contexte du PCLE, permettra d'obtenir un **portrait plus dynamique de l'enseignement PS** et de **faire le lien entre l'éducation et le marché du travail**.
- Un système d'information sur les étudiants postsecondaires complet qui permettra **de suivre les étudiants dans le temps**, tout au long de leurs programmes universitaire, collégial et d'apprentissage et sur le marché du travail.





## Ajout d'information sur le traitement chirurgical par couplage

### ■ **Registre canadien du cancer (RCC)**

- Contient des données sur tous les résidents du Canada qui ont reçu un diagnostic de cancer entre 1992 et 2012
- Ne renferme à l'heure actuelle aucune données sur le traitement

### ■ **Base de données sur les congés des patients (BDCP)**

- Données sur les hospitalisations
- Interventions chirurgicales – Codes de la CCI
- 2010-2013

### ■ **Système national d'information sur les soins ambulatoires (SNISA)**

- Données sur les soins ambulatoires
- Interventions chirurgicales – Code de la CCI
- 2010-2013

**Couplage avec le RCC effectué dans l'*Environnement de couplage de données sociales (ECDS)* de Statistique Canada**



# L'accès aux données à Statistique Canada

