

L'inclusion numérique : Comprendre la cybervulnérabilité au Canada

Firouz Fallahi

Direction de la politique et de la stratégie de service
Emploi et Développement social Canada

Atelier sur l'évaluation quantitative d'impact - Fev 2026



Employment and
Social Development Canada

Emploi et
Développement social Canada

Canada

Aperçu

- Contexte
- Définition de l'indice de cybervulnérabilité
- Sources de données et défis associés
- Méthodologie
 - Analyse en composantes principales (ACP)
 - Estimation sur petits domaines (EPD)
- Modélisation et calcul des sous-composantes de l'ICV
- Résultats
- Résumé
- Implications de politiques



Contexte – Transformation numérique de la prestation de services et disparités relatives à l'accès aux services

- Le gouvernement du Canada améliore constamment la prestation de ses services et la diffusion d'information en ligne
- Notre transformation pour devenir un gouvernement numérique vise à moderniser et à adapter la prestation de services afin de mieux répondre aux besoins des Canadiens partout au pays
- L'attention croissante accordée à la communication et à l'interaction numériques souligne la nécessité pour le gouvernement de déterminer de manière proactive qui sont les citoyens les plus susceptibles d'être confrontés à des obstacles pour accéder aux services en ligne et les utiliser. Nous faisons référence à ce groupe en tant que « **personnes cybervulnérables** »



Justification du projet

- Les personnes cybervulnérables courent un plus grand risque d'être désavantagées en matière d'accès aux services, et elles sont plus susceptibles d'être confrontées à des obstacles pour obtenir les services en ligne dont elles ont besoin.
- De plus, le fossé numérique exacerbe d'autres disparités socioéconomiques, creusant encore davantage ces disparités. Par conséquent, il est essentiel de combler les lacunes numériques pour atteindre les objectifs de développement durable des Nations Unies.

Cette étude vise à déterminer les sous-groupes qui composent la population cybervulnérable et leur emplacement géographique.

- Quels sont les Canadiens les plus susceptibles d'être cybervulnérables?
- De quelle manière la cybervulnérabilité varie-t-elle selon les différents sous-groupes de la population et dans l'ensemble des diverses régions géographiques?



Qu'est-ce que l'indice de cybervulnérabilité?

- Emploi et Développement social Canada (EDSC) a élaboré son indice de cybervulnérabilité (ICV)* afin de mesurer la vulnérabilité numérique pour la première fois en 2015 et a procédé à plusieurs mises à jour depuis**
- L'ICV examine trois sous-dimensions de la cybervulnérabilité :
 1. **L'accès** : les personnes disposent-elles des moyens nécessaires pour profiter d'Internet?
 2. **Les compétences** : les personnes détiennent-elles les compétences et les connaissances nécessaires pour tirer pleinement profit d'Internet?
 3. **Le confort à l'utilisation** : les personnes désirent-elles utiliser Internet (notion de préférence)?

* e-vulnerability Index (EVI) en anglais

**Veuillez noter qu'en raison des changements de méthodologie et des différentes étapes de normalisation des données dans le calcul des scores, ceux-ci ne sont pas directement comparables entre différentes périodes. Les scores ICV peuvent uniquement être utilisés pour comparer la cybervulnérabilité relative des différents sous-groupes socioéconomiques/démographiques à un moment donné.

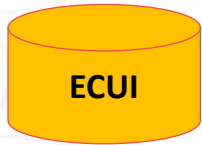


Création de l'ICV : Sources de données

Les données sur l'utilisation d'Internet par la population canadienne proviennent principalement des jeux de données de Statistique Canada :

➤ Enquête canadienne sur l'utilisation d'Internet (ECUI)

- La dernière enquête a été réalisée en 2022 dans dix provinces canadiennes;
- Taille de l'échantillon : environ 25 000 personnes âgées de 15 ans et plus;
- Elle comprend une série de questions sur :
 - l'accès à Internet et son utilisation;
 - les activités réalisées en ligne, de l'envoi et de la réception de courriels à l'utilisation de services gouvernementaux en ligne;
 - les difficultés, les questions de sécurité ou les préoccupations en matière de protection des renseignements personnels liées à l'utilisation d'Internet.
- L'ECUI a été utilisée pour mesurer les trois sous-dimensions de l'ICV, c'est-à-dire **l'accès, le confort et les compétences***.



* Dans les versions précédentes de l'ICV, les compétences étaient évaluées à partir des données du Programme pour l'évaluation internationale des compétences des adultes (PEICA). Cependant, les données de l'ECUI ont été utilisées en raison de la suppression du module « Résolution de problèmes dans des environnements technologiques » dans le dernier cycle du PEICA.

Création de l'ICV : La méthode (I)

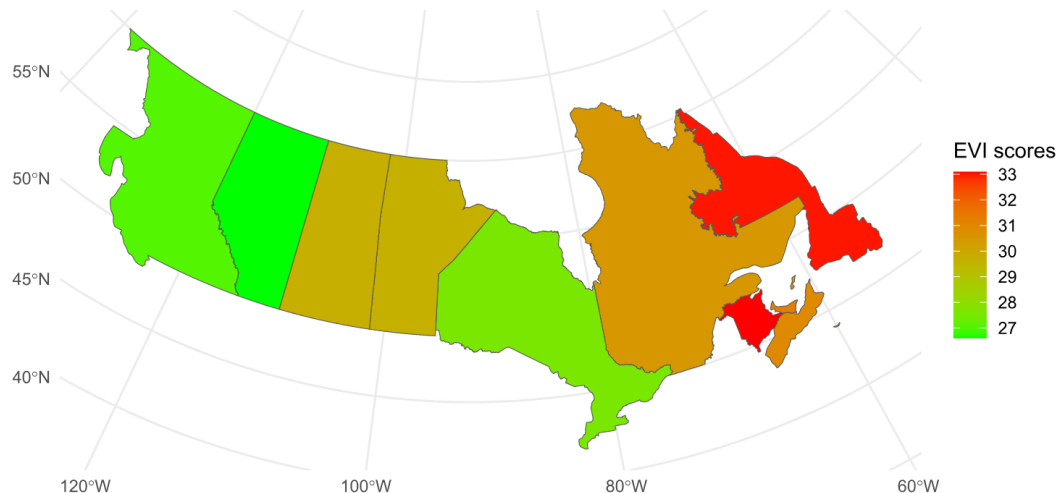
Analyse en composantes principales

- L'ECUI comprend plusieurs indicateurs pertinents pour l'une ou l'autre dimensions de l'ICV (accès, confort ou compétence)
- L'analyse en composantes principales est utilisée pour réduire l'ensemble des indicateurs liés à Internet présents dans l'ECUI à un sous-ensemble gérable et pour mesurer les trois sous-dimensions de l'ICV
- Grâce à cette approche, nous obtenons des scores de l'ICV fiables, mais seulement pour un ensemble limité de géographies, comme illustré sur la diapositive suivante



Création de l'ICV : Le défi

L'ICV dans les provinces canadiennes : Ensemble de la Population



- La taille de l'échantillon l'ECUI est suffisante pour produire des estimations fiables au niveau national et provincial ; cependant, leur précision diminue lorsqu'on analyse des communautés ou des groupes démographiques plus restreints.



Création de l'ICV : Le défi (suite)

Une précision moindre à des niveaux d'agrégation plus petits ou plus fins rend son utilisation inefficace pour l'élaboration de politiques basées sur des preuves et pour faire des interventions ciblées

Solution :

Recensement de la population

- Le Recensement de la population de 2021 recueille des données démographiques et socioéconomiques exhaustives sur des millions de personnes, **mais il ne comporte aucune question sur l'utilisation, les compétences et l'accès liés à l'utilisation d'Internet.**



Création de l'ICV : La méthode (II)

- Pour combler cette lacune, nous avons utilisé les **techniques d'estimation pour les petits domaines** afin de réaliser l'estimation des trois sous-dimensions de la cybervulnérabilité pour chaque personne recensée
- Chaque personne ayant répondu au Recensement se voit attribuer un score de cybervulnérabilité, ce qui permet d'obtenir des résultats adaptés à diverses régions géographiques et à différents segments démographiques, par exemple les aînés et les immigrants
- Cette approche fournit des renseignements précis pour l'élaboration de politiques et d'interventions ciblées, entre autres applications de politiques publiques



Qu'est-ce que l'estimation pour petits domaines (EPD)

- Les techniques d'estimation pour les petites régions se penchent sur l'élaboration de procédures statistiques permettant d'obtenir des estimations précises sur de petites régions, c'est-à-dire des régions où la taille des échantillons est faible ou nulle (Pratesi et Giusti, 2015)
- Rao et Molina (2015) définissent une petite zone comme "tout domaine pour lequel des estimations directes et d'une précision adéquate ne peuvent pas être produites"
 - Elle combine des données provenant de deux sources ou plus pour fournir des estimations fiables
 - Elle est largement utilisée dans les statistiques officielles lorsque les tailles d'échantillon sont trop petites pour fournir des estimations fiables à des niveaux détaillés
 - Elle **emprunte des forces** à des zones connexes et à des ensembles de données auxiliaires
 - Elle fournit des estimations indirectes lorsque des estimations directes ne sont pas réalisables



Approches de l'EPD

Il existe deux approches principales utilisées en EPD :

1) Modèles au niveau de du domaine :

- Utilisation de l'estimation directe entre la variable d'intérêt et les variables auxiliaires au niveau du domaine
- Le modèle le plus populaire est celui de Fay et Herriot (1979)

$$y_i^{Dir} = x_i^T \beta + u_i + e_i$$

- Où y_i est le paramètre d'intérêt dans le domaine i et Dir indique les estimations directes
- e_i est l'erreur d'échantillonnage
- u_i est l'effet aléatoire et est supposé indépendant de l'erreur d'échantillonnage



Approches de l'EPD (suite)

- Cette approche présente quelques avantages et inconvénients, dont certains sont les suivants :
 - Avantages : Elle nécessite des données au niveau de la zone plutôt que des données individuelles, qui sont plus largement disponibles et ne sont pas soumises à des règles strictes de confidentialité ou de protection de la vie privée
 - Inconvénients : Les résultats au niveau de la zone sont généralement moins efficaces et ne peuvent pas être subdivisés en unités plus petites une fois générés



Approches de l'EPD (suite)

2) Modèles au niveau des unités

- Cette approche utilise des données au niveau unitaire (par exemple, individu ou ménage), qui sont beaucoup plus riches que les données au niveau zonal, mais nécessitent l'accès aux microdonnées
- Nous utilisons des modèles au niveau des unités pour estimer les sous-dimensions de l'ICV:

$$y_{i,j} = x_{i,j}^T \cdot \beta + \delta_i + e_{ij}$$

Où :

$y_{i,j}$ représente la sous-dimension de l'ICV (Accès, Confort ou Compétence)

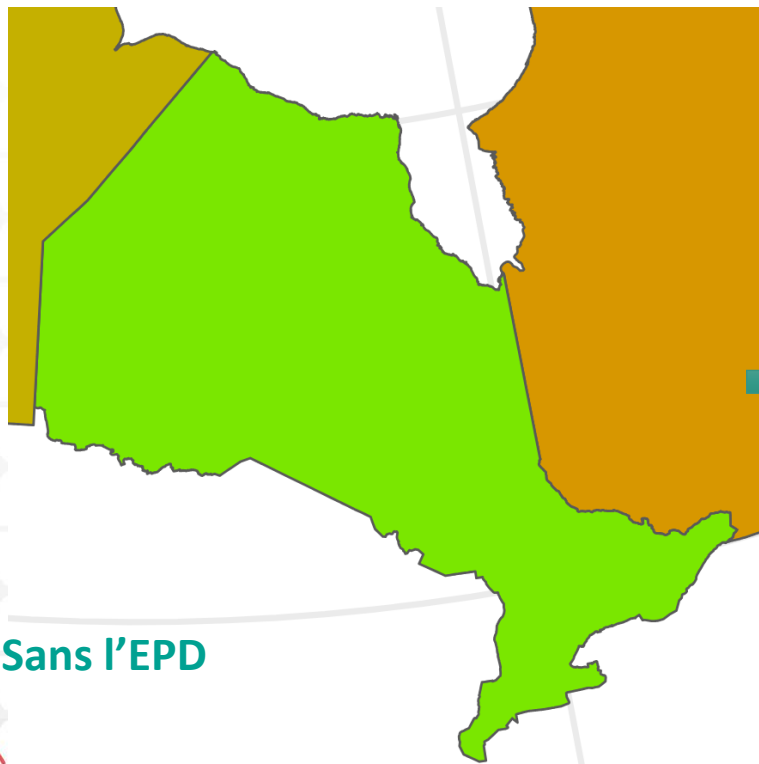
$x_{i,j}^T$ désigne un vecteur de covariables auxiliaires explicatives pour le répondant j dans le domaine i ,

e_{ij} sont supposés être indépendamment distribués et indépendants de δ_i

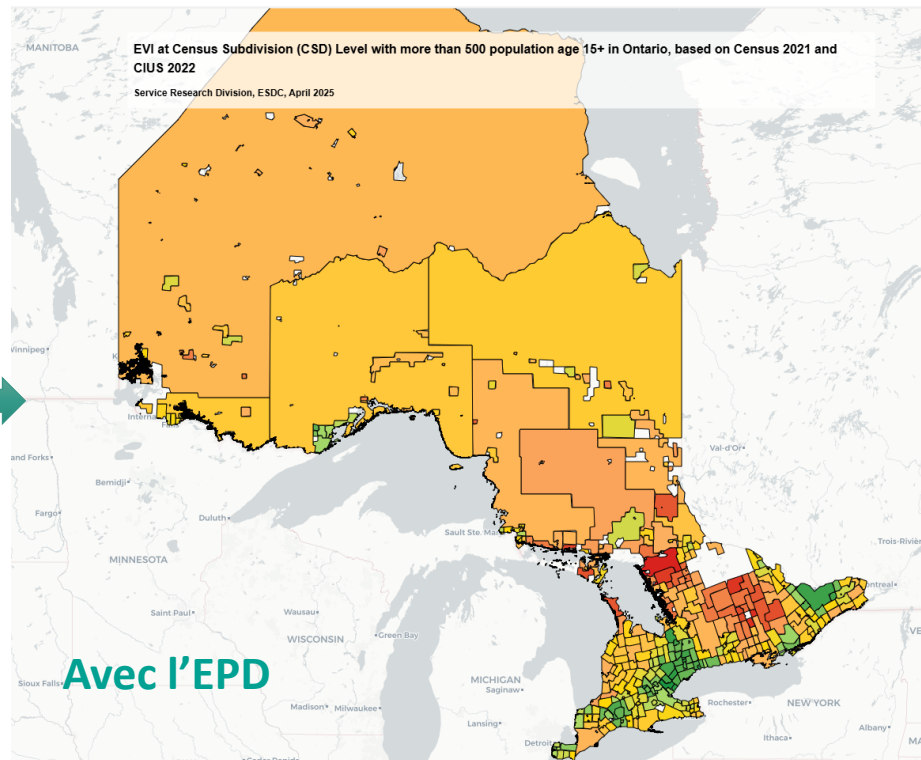
δ_i est l'effet aléatoire spécifique à la zone qui n'est pas expliquée par $x_{i,j}^T$



Comparaison des Estimations : Sans l'EPD vs. Avec l'EPD

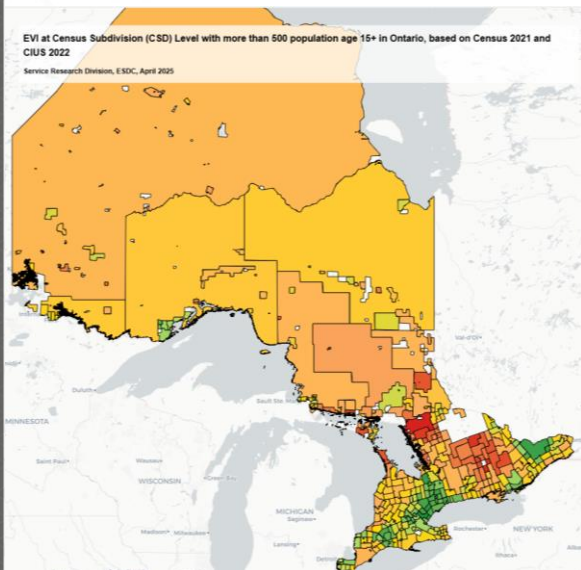


Sans l'EPD

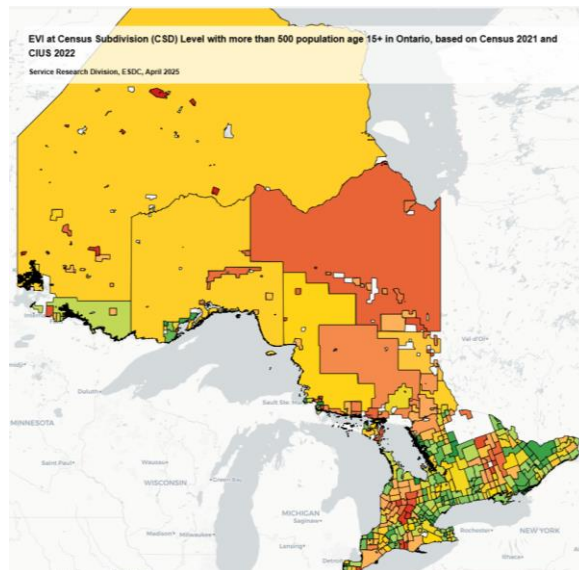


Avec l'EPD

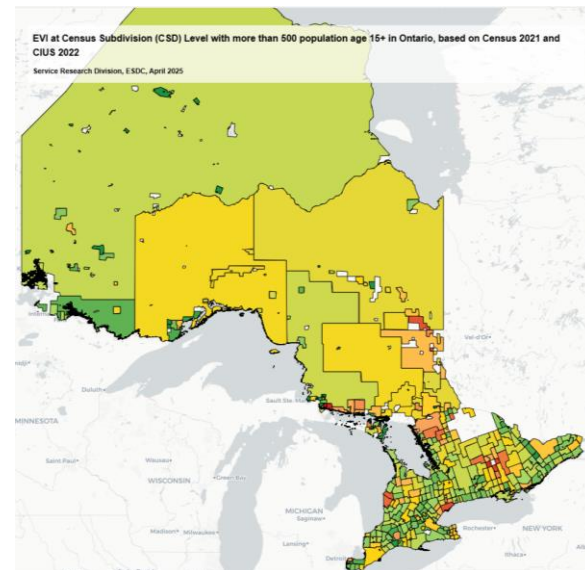
Générer de l'information grâce aux capacités de l'EPD



Scores d'ICV



Sans diplôme d'études
secondaires, 15 ans et plus



En situation de faible
revenue (Selon la
Mesure du Panier de
Consommation (MPC))

Un aperçu de la création de l'indice de cybervulnérabilité

Définition du cadre conceptuel : Identifier les indicateurs reliés à l'ICV en utilisant les techniques de l'analyse en composantes principales

Accessibilité (ECUI 2022)

- Accès à l'internet depuis la maison
- Accès depuis le lieu de travail
- Accès depuis l'école
- Accès depuis la bibliothèque publique
- etc

Confort (ECUI 2022)

- Avoir des préoccupations en matière de sécurité lors de l'utilisation d'Internet
- Pas d'accès à Internet à domicile en raison de préoccupations liées à la sécurité
- Ne pas faire d'achats en ligne en raison de préoccupations liées à la sécurité
- etc

Compétences (ECUI 2022)

- Utilisation d'Internet pour :
- Les communications
 - l'accès à l'information
 - les services bancaires en ligne
 - la suppression de l'historique
 - la modification des paramètres de confidentialité
 - le blocage des courriels ou des messages
 - l'utilisation de Word et de feuilles de calcul
 - etc

Modéliser chaque sous-dimension à l'aide de techniques d'estimation pour petites régions

Mise en œuvre des coefficients/paramètres estimés dans le **Recensement de la population** de 2021 pour estimer les scores individuels de l'Indice de cybervulnérabilité (ICV)

Sous-Indice: Accessibilité

Sous-Indice: Confort

Sous-Indice: Compétences

Normaliser chaque sous-indice pour qu'il varie entre 0 et 100

Moyenne pondérée des sous-indices

Normaliser pour obtenir une plage de valeurs entre 0 et 100

Indice de cybervulnérabilité (ICV)

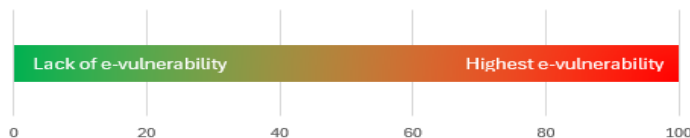
Modélisation et estimation des sous-dimensions de l'ICV (I)

- Le **nombre de composantes principales** dans l'ECUI a été établi en analysant les valeurs propres et la variance cumulative expliquée
 - Les pondérations attribuées à chaque composante ont été calculées en fonction du pourcentage de variance totale expliquée par chaque composante
- **Des modèles au niveau de l'unité** ont été élaborés à l'aide d'une méthode bayésienne de Monte Carlo par chaînes de Markov
 - Cette méthode comprend des procédures de sélection de données et de rodage.
 - La convergence et le manque d'autocorrélation dans les échantillons de la méthode bayésienne de Monte Carlo par chaînes de Markov ont été confirmés
- **Schématisation dans le Recensement** et application des coefficients estimés à partir de l'analyse de la méthode bayésienne de Monte Carlo par chaînes de Markov pour calculer les scores d'accès, de confort et de compétences pour **chaque personne** dans le Recensement de 2021



Modélisation et estimation des sous-dimensions de l'ICV (II)

- Les **sous-indices de cybervulnérabilité ont été normalisés** afin d'être contraint à l'échelle entre 0 et 100
- Des **pondérations égales** ont été attribuées à chaque sous-dimension de la cybervulnérabilité et agrégées pour obtenir le score de l'ICV pour chaque personne
 - Des analyses de sensibilité ont été effectuées en faisant varier les pondérations; toutefois, les conclusions générales sont demeurées les mêmes dans l'ensemble
- Les résultats des **scores de l'ICV ont été normalisés** à nouveau afin d'obtenir les scores finaux de l'ICV compris entre 0 et 100

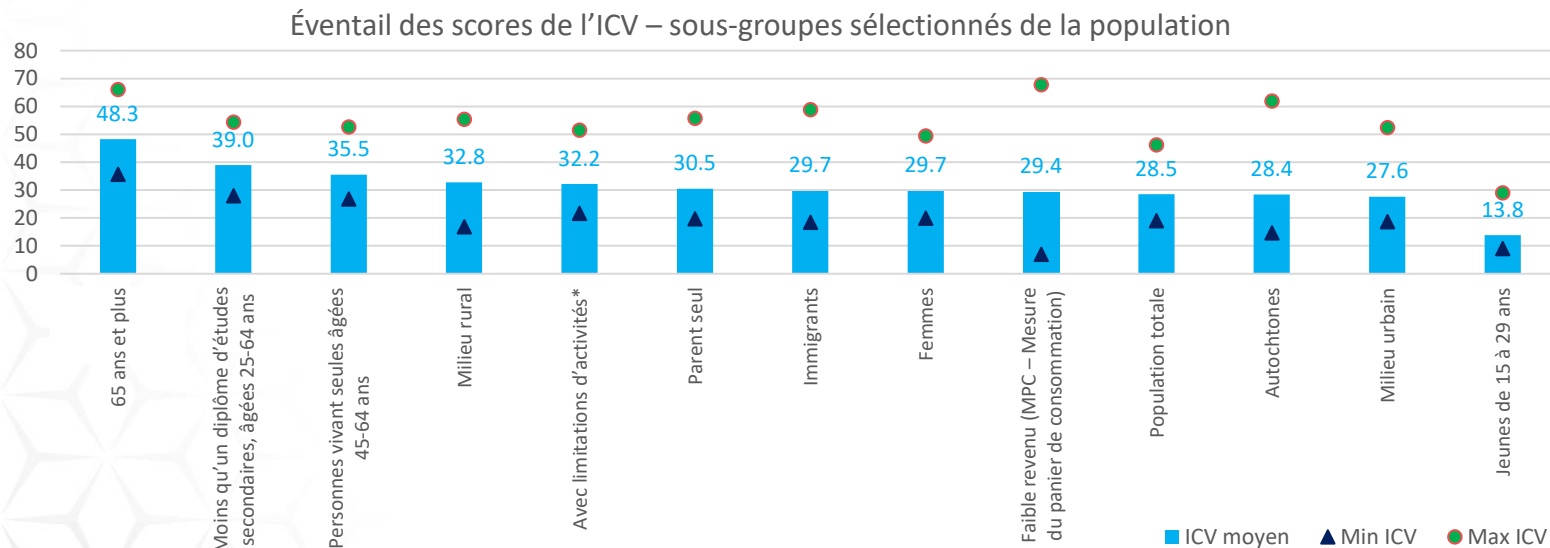


- À la fin, les scores de l'ICV étaient disponibles pour plus de sept millions de répondants au Recensement (non pondérés)
 - Ces scores ont été obtenus au niveau individuel, ce qui permet de les agréger à des niveaux géographiques plus précis pour des sous-groupes de la population canadienne



Indice de cybervulnérabilité : Résultats

- La cybervulnérabilité varie en fonction des divers aspects de la vie d'une personne. Par exemple, l'âge, l'origine ethnique, le statut d'immigrant, la situation géographique et les limitations d'activités peuvent tous influencer sur le degré de cybervulnérabilité d'une personne



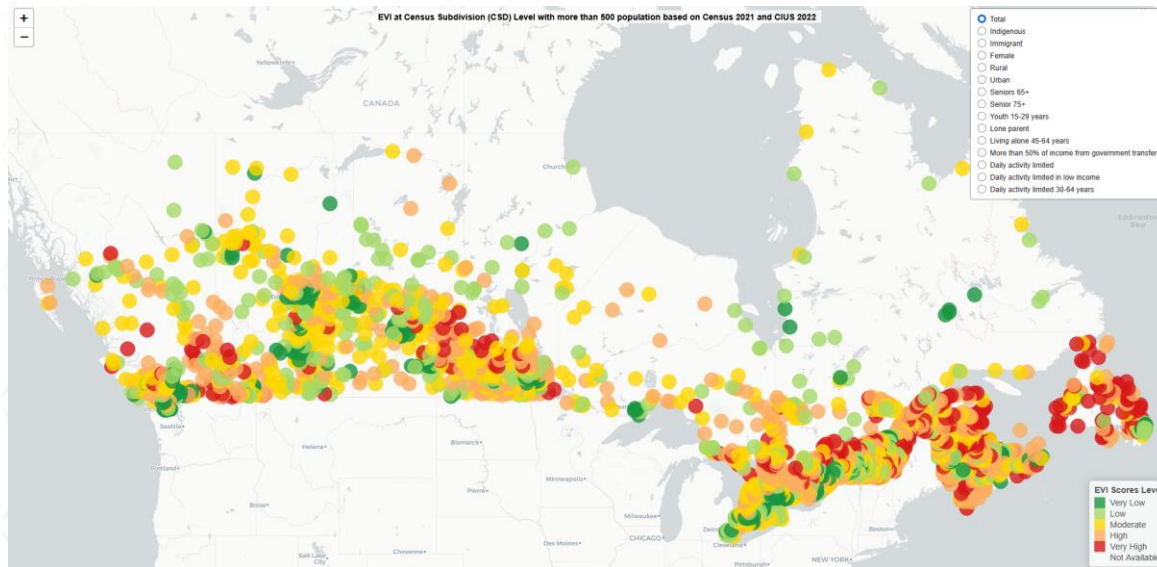
* Tel que défini dans le Recensement de la population. Les personnes ayant des limitations d'activités représentent un groupe plus étendu que les personnes en situation de handicap telles que définies par EDSC selon l'Enquête canadienne sur l'incapacité.

Indice de cybervulnérabilité : Résultats (suite)

- Au Canada, le score de l'ICV global est de 28,5, ce qui sert de point de référence pour comparer la cybervulnérabilité des sous-groupes de la population par rapport à la moyenne
- Parmi les sous-groupes particulièrement pertinents pour EDSC :
 - **Les aînés, les personnes ayant des limitations d'activités** et les personnes vivant dans des régions **rurales** présentent un niveau de cybervulnérabilité plus élevé que la moyenne
 - À l'inverse, les **jeunes** affichent le score de l'ICV le plus bas
 - Il existe des variations significatives dans les scores de l'ICV au sein de chaque sous-groupe :
 - Pour les aînés, les scores moyens peuvent varier entre 36 et 66, ce qui indique qu'ils sont généralement plus cybervulnérables, indépendamment de l'endroit où ils résident ou de leur statut socioéconomique
- Les scores de l'ICV dépendent de l'endroit où les personnes résident, comme le montrent les diapositives suivantes

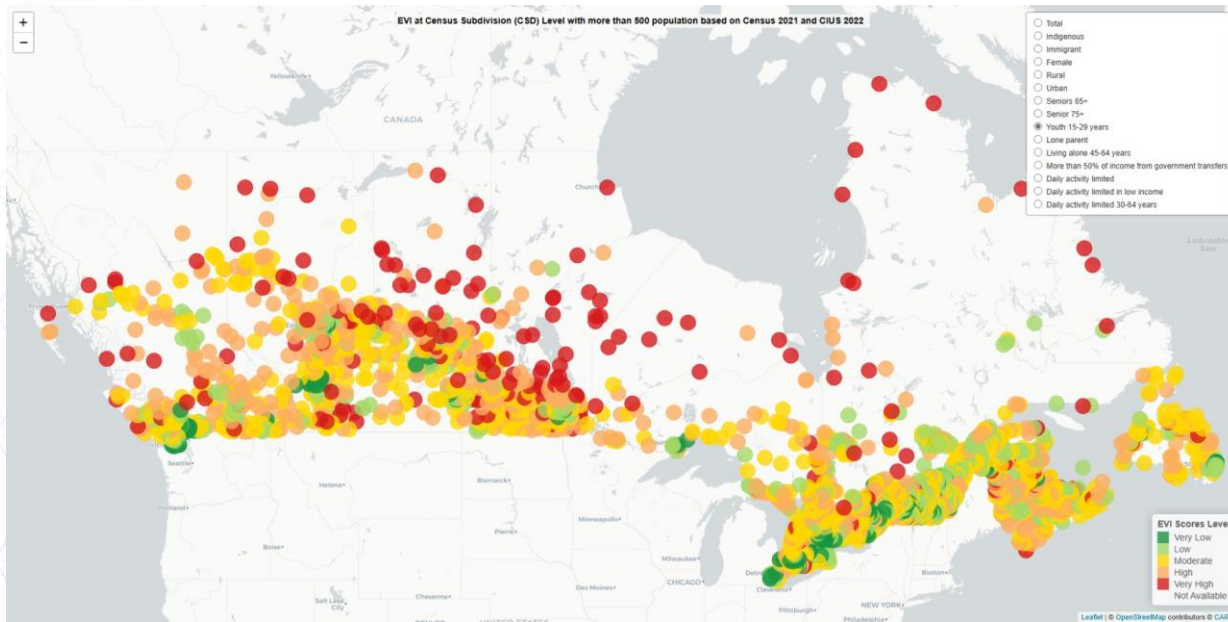


Ensemble de la population au niveau des subdivisions de recensement



- La répartition des scores de l'ICV entre les régions géographiques varie considérablement selon le sous-groupe observé
- Pour l'ensemble de la population, la plupart des subdivisions de recensement (SDR) ayant un score de l'ICV élevé (points rouges) sont situées dans les provinces de l'Est

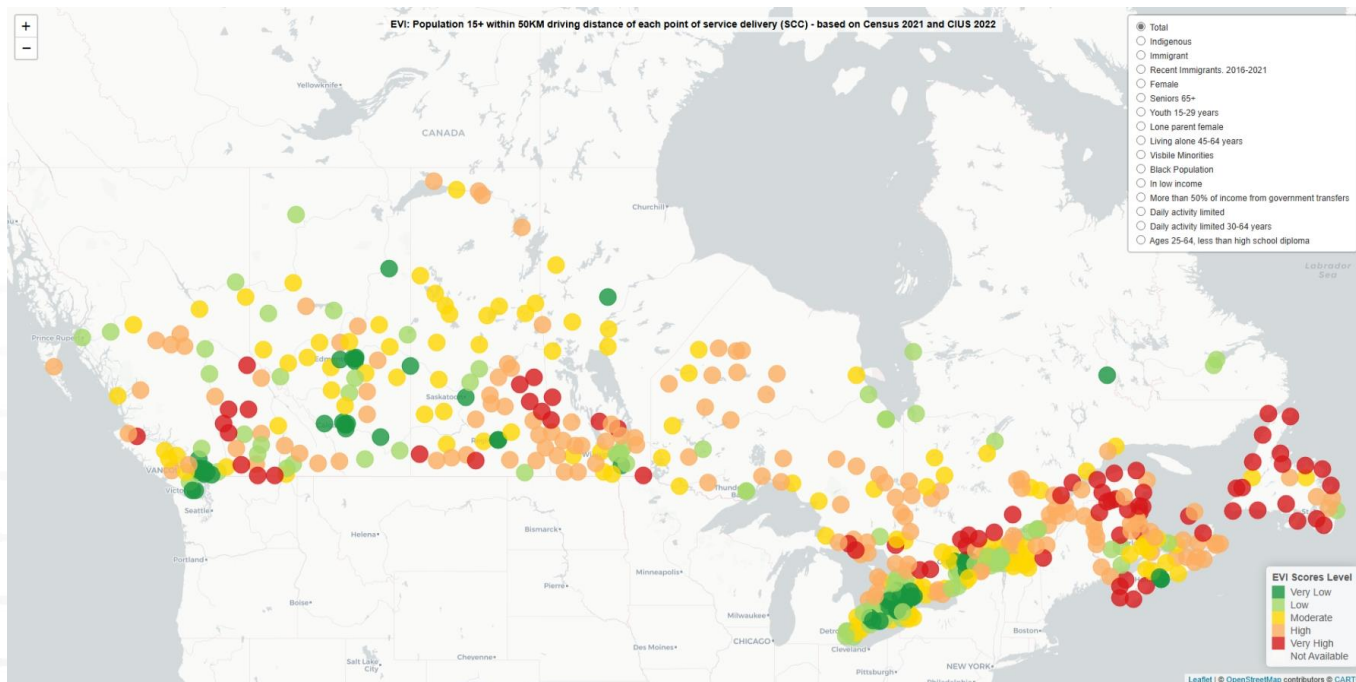
Les jeunes âgés de 15 à 29 ans au niveau des subdivisions de recensement



- Les jeunes obtiennent les scores de l'ICV les plus bas (13,8) parmi les groupes démographiques
- Leurs scores de l'ICV varient entre 9 et 29
- Les scores de l'ICV les plus élevés se trouvent principalement au Manitoba et en Saskatchewan, en particulier dans leurs régions au nord

Centres de prestation de services de Service Canada

Distance de 50 km en automobile



Le score de l'ICV pour un Centre Service Canada (CSC) donné dépend des caractéristiques de sa population (La prochaine diapositive présente cette information pour les CSC affichant les scores l'ICV les plus élevés et les plus faibles)

Subdivisions de recensement (SDR) où se trouvent les Centres Service Canada ayant les scores de l'ICV les plus élevés et les plus bas : caractéristiques choisies

SDR : Springdale (T.-N.-L.)	Caractéristiques sociodémographiques de la population dans la SDR : <ul style="list-style-type: none"> ▪ Proportion élevée d'aînés dans la population (31,7 %) ▪ Taux de prévalence élevé du faible revenu (évaluation de l'impact sur le marché du travail) (21,5 %) ▪ Faible taux d'emploi (37,3 %) ▪ Taux de chômage élevé (15 %) ▪ Proportion élevée de la population n'ayant pas obtenu de diplôme d'études secondaires (33,5 %) ▪ Personnes vivant seules (12,3 %)
Population 2 965	
Score de l'ICV moyen de la population 40,2	
SDR : Canmore (Alberta)	Caractéristiques sociodémographiques de la population dans la SDR : <ul style="list-style-type: none"> ▪ Faible proportion d'aînés dans la population (17,4 %) ▪ Faible taux de prévalence du faible revenu (6,5 %) ▪ Taux d'emploi élevé (65 %) ▪ Faible taux de chômage (9,7 %) ▪ Faible proportion de la population n'ayant pas obtenu de diplôme d'études secondaires (6,8 %) ▪ Personnes vivant seules (11,6 %)
Population 15 990	
Score de l'ICV moyen de la population 24,0	



Résumé

- L'évolution constante de nos contextes socioéconomique et technologique (entre autres) rend d'autant plus important de veiller à ce que les activités de collecte de données reflètent mieux nos besoins changeants relatifs aux données
- La collecte de données détaillées étant coûteuse, il est plus efficace d'intégrer des données provenant de plusieurs aspects de la vie et des activités des individus
- Nous avons utilisé la technique de l'estimation pour petits domaines (EPD), qui a permis de pallier les limitations de la taille de l'échantillon l'ECUI de manière rentable en combinant les données du Recensement de la population de 2021
 - Les scores de l'ICV sont calculés à titre individuel, ce qui permet de personnaliser et d'agréger les résultats en fonction de régions géographiques et de caractéristiques démographiques précises



Effets de l'ICV sur les politiques publiques

- Il détermine les secteurs précis et les groupes démographiques où d'autres modes de prestation de services (tels que le courrier, le téléphone ou en personne) ou de diffusion de l'information (notamment la radio ou les journaux) sont particulièrement nécessaires
- Faciliter la sensibilisation en informant les agents de terrain de la vulnérabilité numérique d'une zone avant de réaliser des visites en personne
- Il contribue à l'élaboration de services numériques, de manière à atténuer les inconvénients potentiels pouvant découler de la mise en place de services en ligne et à y remédier.
- Une politique universelle uniforme ne sera pas efficace





Merci!

Division de la recherche sur les services,
Direction de la stratégie et des politiques de service, EDSC

NC-DPSS-Recherche_Services-Service_Research-SPSD@hrsdc-rhdcc.gc.ca