

Exigences relatives aux propositions de recherche

Lignes directrices générales pour les propositions

- Toute proposition soumise doit contenir suffisamment de détails pour permettre d'évaluer la faisabilité du projet.
- L'accès aux microdonnées confidentielles vise à appuyer la modélisation statistique inférentielle. Les propositions de recherche doivent être analytiques et axées sur la modélisation.
- Les chercheurs sont tenus d'utiliser, dans la mesure du possible, [des tableaux de données accessibles au public](#) ou [l'accès à distance en temps réel \(ADTR\)](#) pour appuyer leurs résultats de modélisation analytique.
- Cet environnement n'est pas destiné aux analyses purement descriptives ni à la publication de grands volumes de données tabulaires. Les statistiques descriptives (par exemple, les tableaux croisés, les statistiques descriptives/de synthèse) doivent se limiter à supporter la modélisation inférentielle.
- Pour obtenir des totalisations personnalisées, veuillez communiquer avec infostats@statcan.gc.ca.
- Tous les chercheurs qui prévoient inclure des membres des Premières Nations, des Métis ou des Inuits dans leur population d'étude sont tenus de collaborer avec les organisations, les communautés ou les chercheurs autochtones. Des ressources ont été élaborées par des organisations des Premières Nations, métisses et inuites qui expliquent les considérations éthiques, le rôle de la recherche dans l'autodétermination des peuples autochtones et l'impératif de pratiques de recherche respectueuses. Veuillez consulter le document suivant pour obtenir des conseils : [Working-with-FN-Inuit-Metis-data_FR.pdf](#)
- Veuillez-vous référer à ces directives supplémentaires pour les propositions qui requièrent les bases de données suivantes :
 - [Recensement et enquête nationale auprès des ménages \(ENM\)](#)
 - [Étude longitudinale et internationale des adultes \(ELIA\)](#)

- [Enquête sur les dépenses des ménages \(EDM\)](#)

Modèle de proposition de recherche

Il est fortement recommandé de faire examiner votre proposition de recherche par un(e) analyste de l'accès aux données de Statistique Canada avant de la soumettre.

Les coordonnées du CDR et CASD sont indiquées à l'emplacement suivant : [Coordonnées de la Division de l'accès aux données](#), sous la section « Centres de données de recherche ».

La proposition de projet ne doit pas dépasser dix pages et doit respecter les titres indiqués dans le modèle ci-dessous. Tout renseignement manquant ou tout détail insuffisant pourrait entraîner des délais dans le traitement de votre demande.

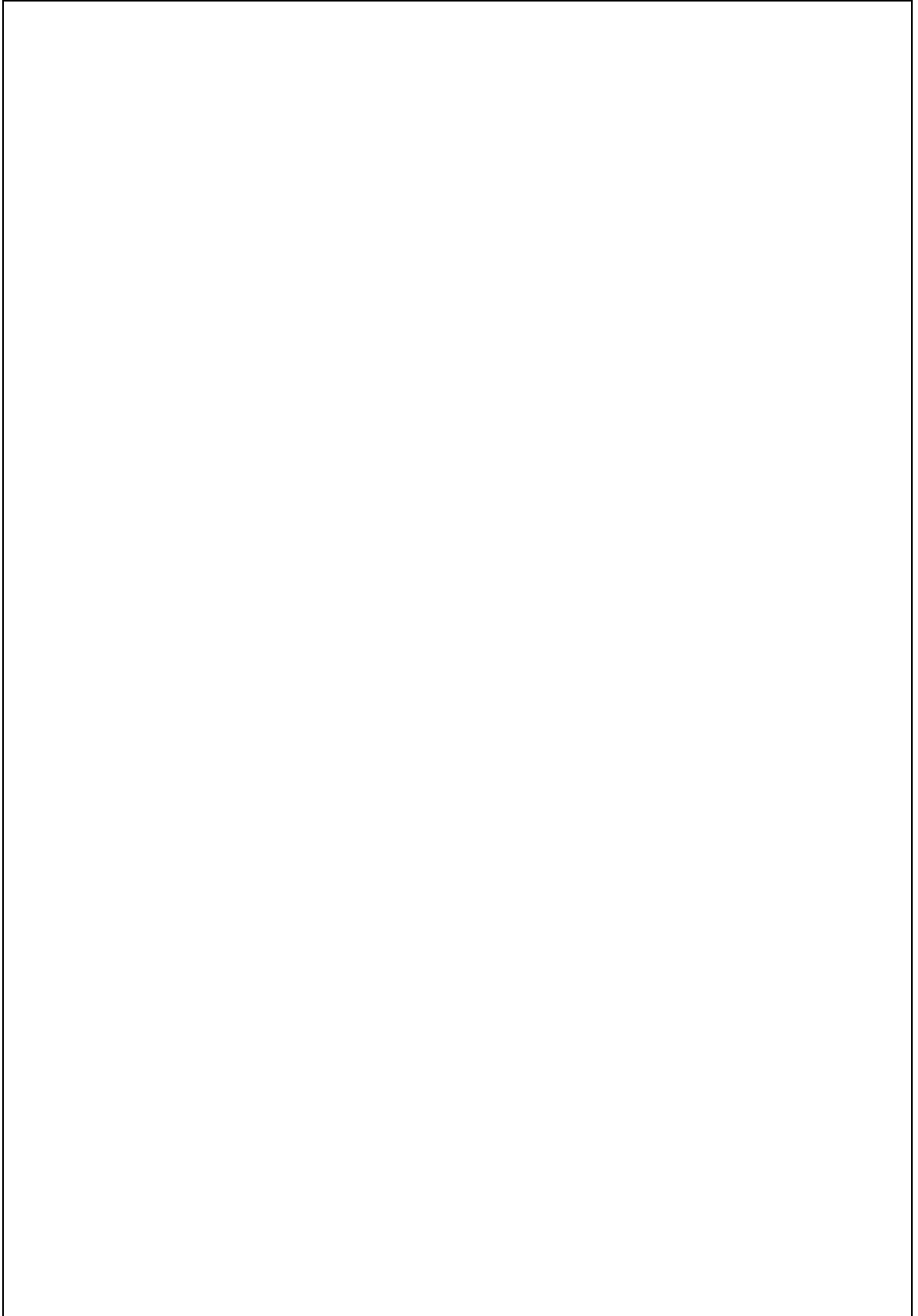
PROPOSITION DE RECHERCHE POUR L'ACCÈS AUX DONNÉES

Chercheur principal et affiliation	
Co-chercheur(s) et affiliation(s)	
Point d'accès*	

1. Titre du projet :

2. Description, justification et objectifs du projet :

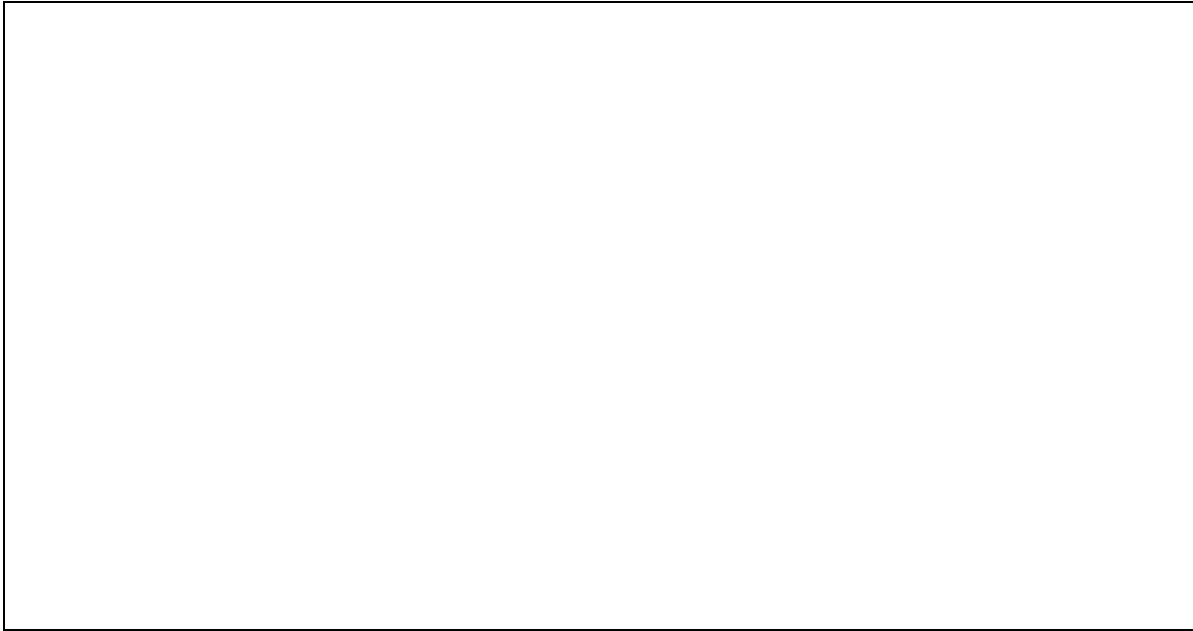
- a) Veuillez fournir un résumé de la documentation actuelle et identifier les lacunes dans les connaissances à combler (environ une page).



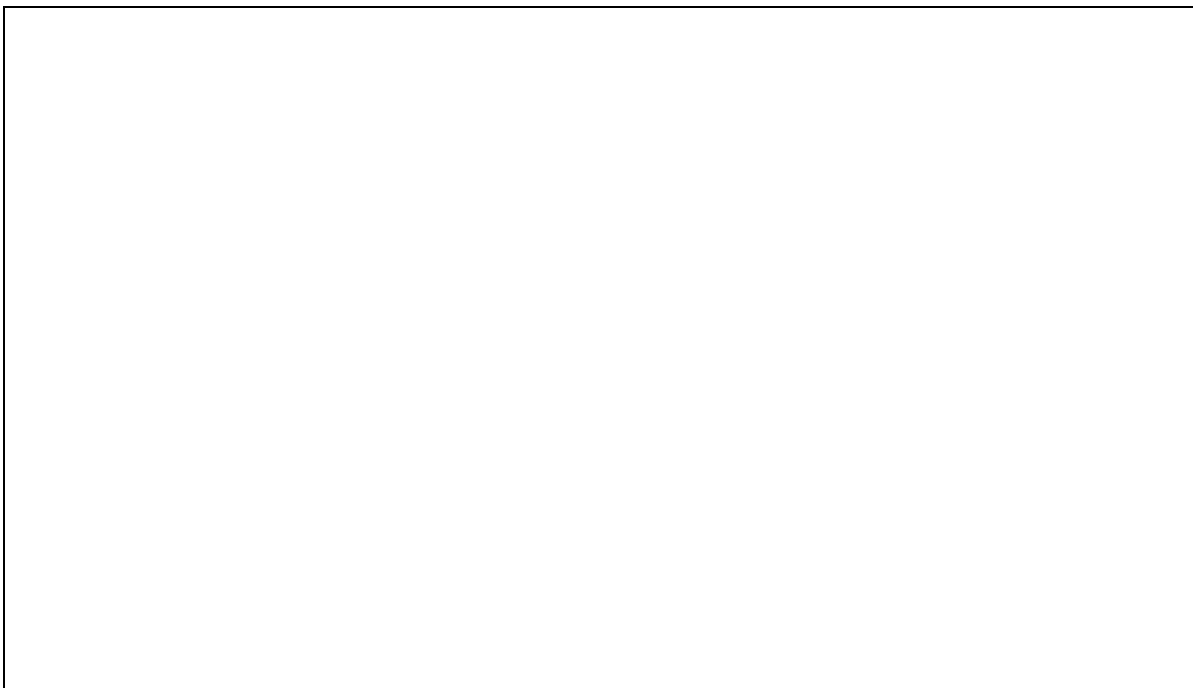
- b) Veuillez identifier clairement les questions ou les objectifs précis du projet :
- Veuillez énumérer les effets et/ou les objectifs précis qui sont évalués.
 - Veuillez décrire les efforts déployés dans le cadre de ce projet, y compris la participation de tout partenaire, chercheur ou collaborateur autochtone.
 - Si ce projet a des répercussions sur les Premières Nations, les Métis ou les Inuits, veuillez expliquer les avantages pour les peuples et les communautés autochtones et identifier tout préjudice possible, ainsi que les stratégies d'atténuation des risques.

3. Exigences relatives aux données :

- a) Justification de l'utilisation de microdonnées (par opposition à d'autres méthodes d'accès aux données comme les fichiers de microdonnées à grande diffusion (FGMD), l'accès à distance en temps réel (ADTR) ou les totalisations personnalisées).

A large, empty rectangular box with a thin black border, intended for providing a justification for the use of microdata compared to other methods like FGMD, ADTR, or personalized totals.

- b) Ensembles de données requis : Précisez comment chaque ensemble de données sera utilisé.

A large, empty rectangular box with a thin black border, intended for specifying the required data sets and how each set will be used.

c) Période : Veuillez indiquer les années/cycles requis.

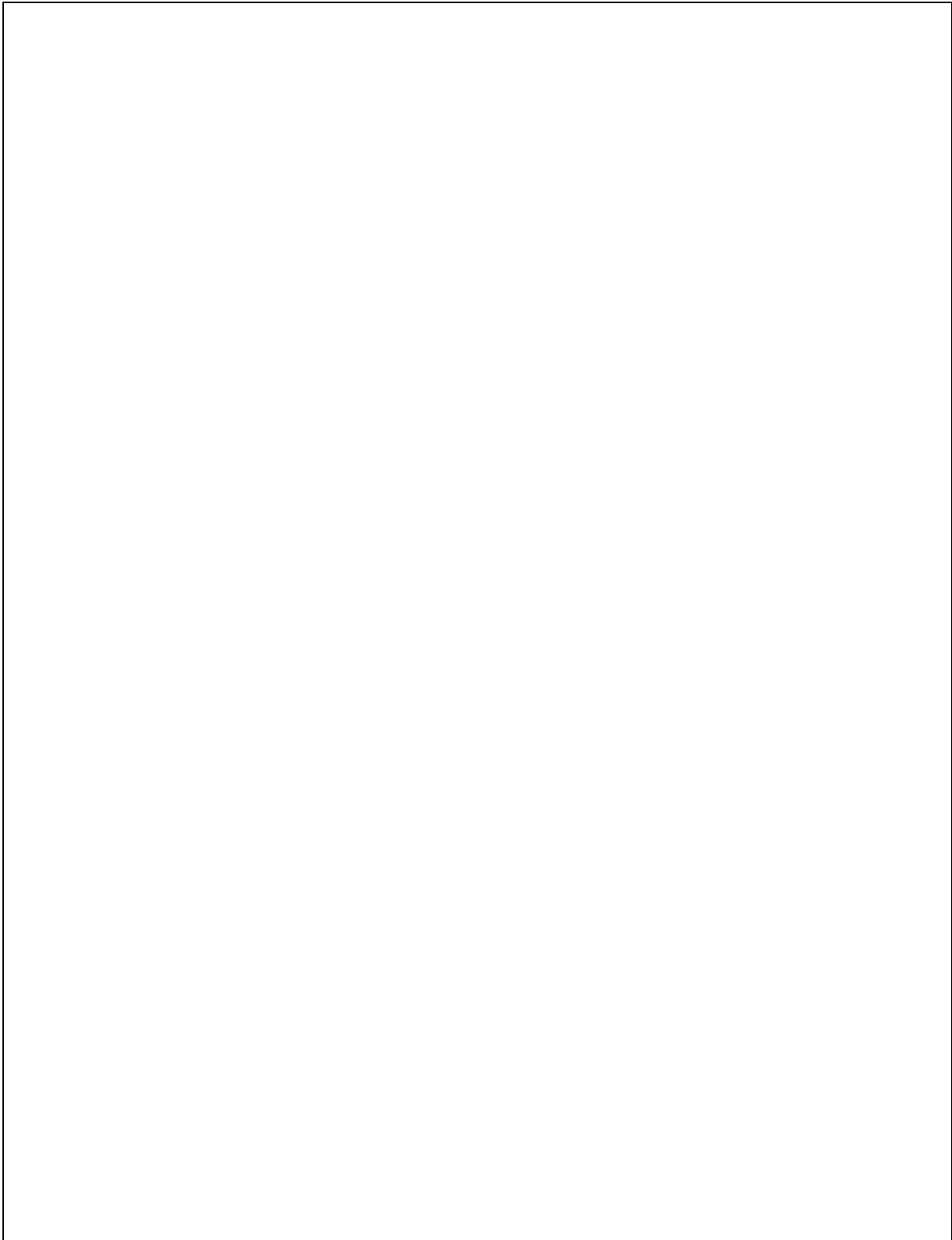
** Remarque : Seulement les bases de données disponibles devraient être énumérées. Veuillez consulter les [Données accessibles dans les Centres de données de recherche](#).*

d) Description de l'échantillon : Précisez l'unité d'analyse et la population visée.

e) Géographie : Veuillez préciser le ou les niveaux géographiques pour lesquels les estimations seront produites (par ex., national, provincial ou inférieur).

f) Liste des variables : Fournissez un tableau énumérant les variables qui devraient être utilisées dans l'analyse (ce tableau peut figurer en annexe). Des livres de codes

non confidentiels peuvent être disponibles au besoin, communiquer avec votre analyste local de l'accès aux données de Statistique Canada pour obtenir de plus amples renseignements.



g) Les ensembles de données seront-ils fusionnés ou regroupés ?

- Si oui, veuillez expliquer comment les données de chaque source seront combinées et utilisées dans cette analyse. Par exemple, vous pouvez regrouper des données provenant d'une même source pour augmenter la taille de votre échantillon ou fusionner des données contextuelles avec les micro-enregistrements.

- Le couplage d'enregistrements n'est pas autorisé dans les centres de données.
 - Si vous avez besoin d'un **nouveau** couplage de microdonnées pour des données sociales, veuillez ne pas soumettre votre demande par l'entremise du PAM. Consultez le site internet de [l'Environnement de couplage de données sociales](#) pour obtenir plus d'information sur le couplage de microdonnées sociales à Statistique Canada.
 - Si vous avez besoin d'un **nouveau** couplage de microdonnées pour des données économiques, veuillez ne pas soumettre votre demande par l'entremise du PAM. Consultez le site internet de [l'Environnement de fichiers couplables \(EFC\)](#) pour obtenir plus d'information sur le couplage de microdonnées économiques à Statistique Canada.

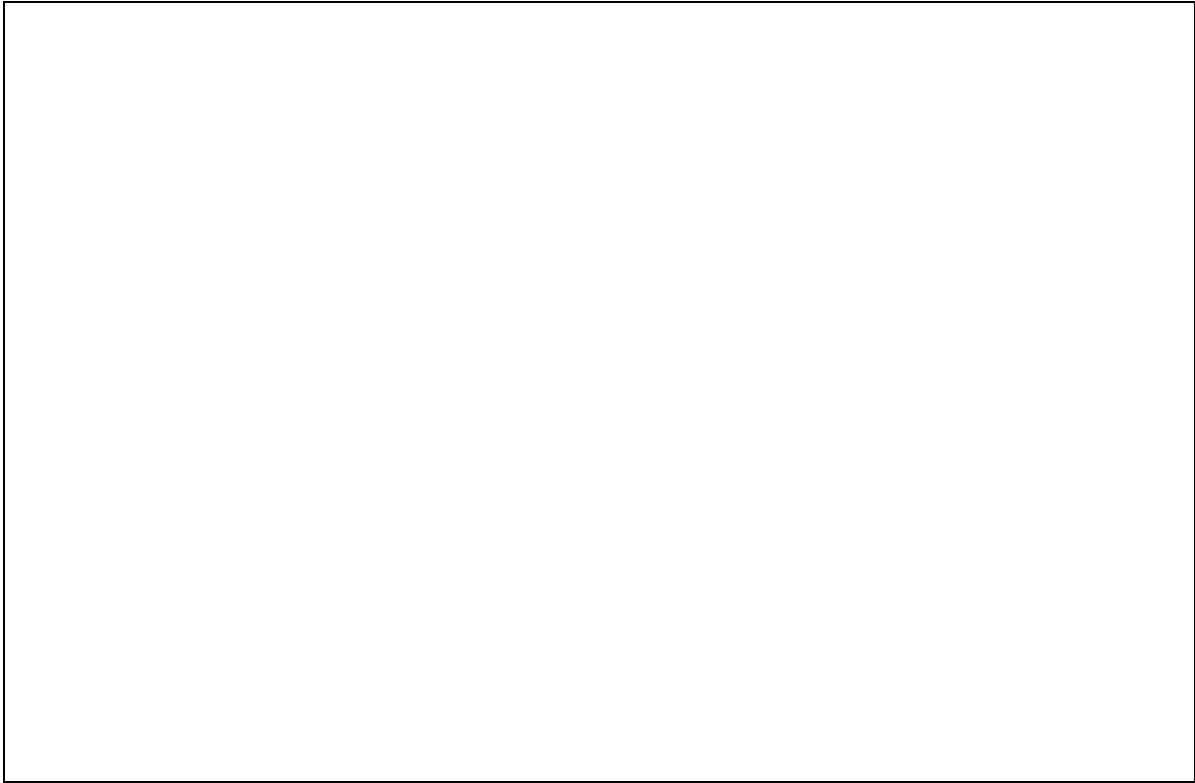
4. Méthodologie :

a) Les tableaux croisés et les statistiques descriptives doivent :

- Être réduits au minimum (justifier le nombre de tableaux et de lignes prévues) ;
- Appuyer les résultats analytiques ;
- Être publiés, dans la mesure du possible, à la fin du projet.

** Remarque : La publication de résultats intermédiaires est fortement déconseillée car elle peut entraîner des problèmes de divulgation résiduelle.*

- Indiquer : a) les variables concernées, b) le nombre de dimensions dans les tableaux et c) le nombre de tableaux.

A large, empty rectangular box with a thin black border, intended for the user to provide details about variables, dimensions, and the number of tables as requested in the instructions above.

b) Modélisation :

- Indiquer : a) la ou les variables dépendantes, les variables indépendantes et b) le type de modélisation utilisé (MCO, régression logistique, etc.).

A large, empty rectangular box with a thin black border, intended for the user to specify dependent and independent variables and the modeling technique used, as per the instructions above.

- c) Inclure une déclaration indiquant que les pondérations applicables seront appliquées lors de l'utilisation des données d'enquête.

- d) Le projet utilise-t-il des techniques d'apprentissage automatique ?

Si oui : Décrivez en détail les techniques d'apprentissage automatique utilisées (par ex., apprentissage supervisé/non supervisé, réseaux de neurones, machines à vecteurs de support (SVM) et méthodes arborescentes, notamment les forêts aléatoires, la sélection de variables (par ex., Lasso, Ridge, Elastic Net, etc.).

e) La taille de l'échantillon prévue est-elle suffisante pour garantir la confidentialité des répondants ?

5. Décrire les exigences relatives aux données externes :

- Statistique Canada reconnaît que, dans certains cas, la combinaison de données provenant de différentes sources peut renforcer la recherche en augmentant la taille de l'échantillon ou en fournissant des renseignements contextuels.

** Remarque : la fusion des données ne signifie pas la liaison des enregistrements ; voir la section 3g ci-dessus.*

- Les données externes désignent les renseignements recueillis par d'autres organisations, ministères et particuliers à leurs propres fins.
- Des tableaux de données agrégées non confidentiels de Statistique Canada (anciennement tableaux CANSIM) peuvent être ajoutés à votre projet de recherche une fois approuvés.

** Remarque : Ceci ne s'applique pas aux fichiers de microdonnées à grande diffusion (FMGD), qui ne sont pas autorisés.*

a) Veuillez décrire les ensembles de données externes que vous prévoyez utiliser et la façon dont vous les intégrerez à votre analyse.

6. Configurations logicielles requises : Quels logiciels seront nécessaires ?

- Veuillez noter que certains logiciels et progiciels ne sont pas disponibles pour une utilisation au CASD/CDR.
- L’approbation de votre proposition ne signifie pas l’approbation du logiciel demandé. Tous les logiciels et modules complémentaires non encore approuvés doivent faire l’objet d’une procédure d’approbation distincte, ce qui pourrait engendrer des délais.
- Les logiciels STATA, SPSS, SAS* et R sont disponibles dans tous les centres de données.
 - ***Veuillez noter que :** SAS est couramment disponible au CASD mais sera mis hors service en 2028, après quoi il ne sera plus accessible aux chercheurs du CASD.
 - Nous encourageons les chercheurs à utiliser des logiciels libres tels que R et Python.
- Pour obtenir une liste plus complète des logiciels disponibles, veuillez contacter votre centre de données local via [Coordonnées de la Division de l'accès aux données](#).



7. Résultats attendus : Décrivez les produits qui découleront de l'analyse proposée (par ex., document de travail, article de revue à comité de lecture, livre ou chapitre, thèse ou mémoire, présentation à une conférence ou rapport commandé (par ex., rapport gouvernemental)).

- Si ce projet comprend des données autochtones, veuillez décrire comment les résultats seront communiqués aux communautés et organisations autochtones.

8. Date de début et durée prévues du projet :

9. Source(s) de financement : Veuillez indiquer toutes les sources de financement ou préciser « S.O. » le cas échéant.

10. Références bibliographiques

