

Bien préparer sa demande de divulgation : réduisez les retards et les erreurs

Par Anne-Charlotte Latour, Analyste,
Centre de données de recherche de
l'université de Montréal, Statistique
Canada



Statistique
Canada

Statistics
Canada

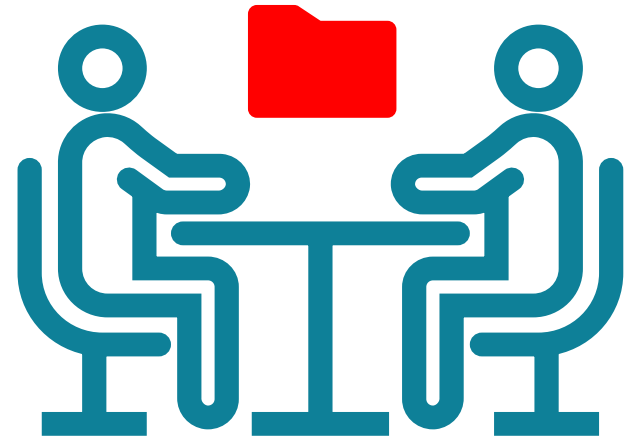
Canada

NOTRE ENGAGEMENT ENVERS LA CONFIDENTIALITÉ

- ❑ Le maintien de la confidentialité des personnes et des entreprises est notre priorité absolue
- ❑ Une personne ou une entreprise ne doit pas être identifiable lorsque des renseignements statistiques sont diffusés

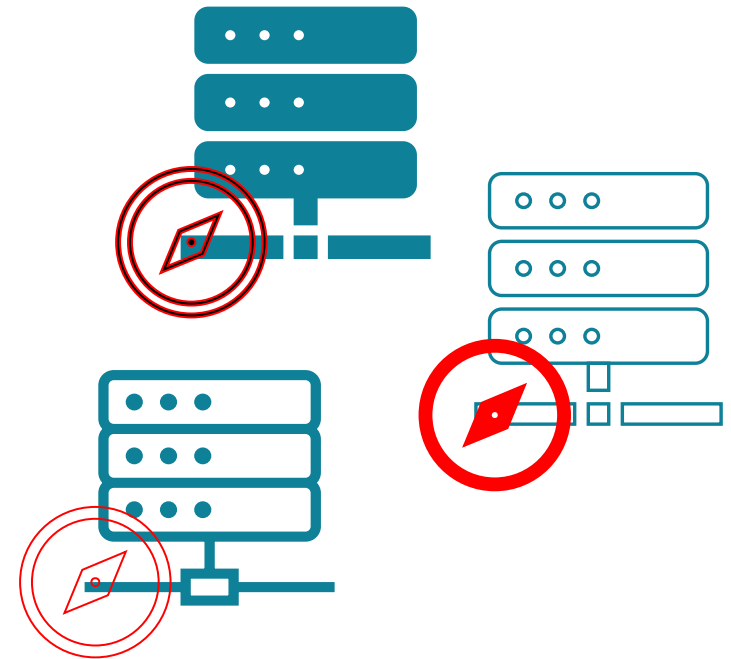


La vérification de la confidentialité est le processus par lequel un analyste – en collaboration avec l'utilisateur des données – examine un résultat, en veillant à ce que les risques de confidentialité soient minimisés comme le précisent les règles de confidentialité



QUI ÉTABLIT LES RÈGLES DE CONFIDENTIALITÉ?

- ❑ Chaque source de données comporte un ensemble de règles de confidentialité qui ont été élaborées par le Groupe spécialisé et méthodologique de Statistique Canada
- ❑ Les règles peuvent varier d'une source de données à l'autre



RÈGLES DE VÉRIFICATION DE LA CONFIDENTIALITÉ

- ❑ Les règles sont reçues avec les données dans un point d'accès sécurisé et servent de guide, précisant les conditions qui doivent être remplies pour diffuser une sortie
- ❑ La sortie ne peut être publiée que lorsque **toutes** les règles sont respectées

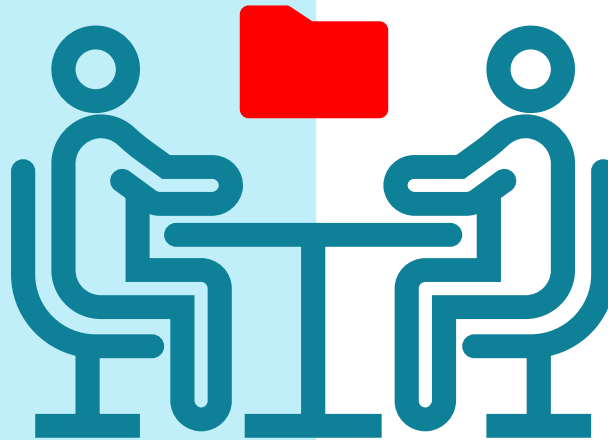


Chacun d'entre nous joue un rôle dans le maintien de la confidentialité des données !

Rôles des acteurs

Chercheurs

- ❑ Connaître, comprendre et appliquer les règles de confidentialité
- ❑ Fournir les documents justificatifs appropriés pour diffuser les résultats



Analyste

- ❑ Assure l'accès aux règles appropriées
- ❑ Aider les chercheurs à comprendre et à appliquer les règles
- ❑ S'assurer que les règles sont respectées avant de envoyer les résultats

ASSURER UN PROCESSUS DE VÉRIFICATION DE LA CONFIDENTIALITÉ FLUIDE

- ❑ Parlez à votre analyste si vous avez des questions sur les règles et la façon de les appliquer, et sur la préparation des résultats à l'appui appropriés requis
- ❑ Guidez votre analyste dans votre demande de vérification, en vous assurant qu'il peut trouver les documents justificatifs nécessaires pour chaque extrant demandé

Environnement de travail

- Ce PC > Equipe > Projet XXXX
 - _Demande À Sortie
 -  • Nouveau formulaire de demande
 - _Demande Archives
 -  • Demande X YY-MM-DD
 - Vetting / Contrôle de confidentialités
 -  • DAD_Vetting_handbook_FR
 -  • Règles de contrôle de la confidentialité

1. Prendre connaissance des règles de contrôle

- **DAD_Vetting_Handbook_FR** donne les outils nécessaires pour la vérification des produits selon les directives en matière de confidentialité. Autrement dit, il fournit la documentation requise pour que les produits soient correctement contrôlés => Maintenir la confidentialité
- **Règles de contrôle de la confidentialité** fournissent des indicateurs sur les seuils minimums des statistiques descriptives
- Tout produit qui ne respecte pas ces lignes directrices n'est pas diffusé

2. Faire sa demande de divulgation

- **Étape 1 : Confidentiality Vetting Request Form-FR** : ce document est à fournir obligatoirement avec sa demande de sortie
 - **Section A** : Permet de vérifier si vous avez pris connaissance des règles de divulgation dans votre demande de sortie. *Vous devez répondre aux questions par Oui/Non/NA*
 - **Section B** : Permet la vérification avec vos anciennes sorties, soit la divulgation par recoupement. **Tout changement échantillonnal doit être avisé à l'analyste**
 - **Section C** : Liste des fichiers que vous voulez sortir (à suivre)
 - **Section D** : Liste des fichiers justificatifs pour la sortie (à suivre)
 - **Section E** : Commentaires permettant la justification des sorties comme l'explication de la méthodologie utilisée (ex. : modification de la pondération)

Section C : Produit faisant l'objet de la demande de diffusion

EX 1 : Sortie d'un seul fichier Excel comprenant 3 Onglet

Nom du fichier (indiquez chaque feuille pour les fichiers de tableur)	Nom de l'enquête ou de l'ensemble de données et cycle(s)	Le cas échéant, nom de la variable de pondération (indiquez s'il s'agit d'une variable échelonnée ou normalisée)	Précisez la méthode utilisée en choisissant un numéro dans la liste ci-dessous	Description de l'échantillon (p. ex. femmes employées, âgées de 21 à 45 ans, en Ontario)	Niveau de géographie du produit demandé (c.-à-d. national, provincial, etc.)
1.Sortie_données : Onglet 1	EPA + IPC	WEIGHT (IPW)	1	Homme de 25 à 55 ans	QC et ON
1.Sortie_données : Onglet 2	Idem	Idem	4	Idem	Idem
1.Sortie_données : Onglet 3	Idem	Idem	3	Homme vs Femme de 25 à 55 ans	QC et ON vs Reste du Canada

Types de méthodes

1. Méthodes descriptives (p. ex. modèles de régressions avec une seule variable, fréquences, analyse par recoupement, moyennes et répartitions, matrice des corrélations, analyse de la variance)
2. Méthodes de mise à l'échelle (p. ex. analyse factorielle)
3. Graphiques (p. ex. histogrammes) – n'oubliez pas d'inclure des tableaux à l'appui.
4. Analyse de régression multivariée simple (p. ex. MCO, Logit, Probit, Tobit)
5. Méthodes de modélisation complexes (p. ex. modélisation par équations structurelles, modélisation linéaire hiérarchique, analyse de croissance, analyse de survie, analyse de données longitudinales, modèles d'équations simultanées, modèles à effets fixes, modèle à effets aléatoires)
6. Autres – veuillez décrire (p. ex. notes et fichiers de syntaxe)

Section D : Fichiers justificatifs

EX 1 : Sortie d'un seul fichier Excel comprenant 3 Onglets

Placez ces fichiers dans votre dossier Documents justificatifs

- 1) Fichiers de syntaxe (ou fichiers journaux) à analyser
- 2) Fréquences justificatives (p. ex. non pondérées/pondérées/pondérées, arrondies)
- 3) Documents justificatifs pour les variables dérivées ou recodées (p. ex. syntaxe, livre des codes, description)
- 4) Autres fichiers requis selon les règles de contrôle applicables (p. ex. résultats de test pour les valeurs en dollars, etc.)
- 5) Le cas échéant, documents justificatifs relativement au risque de divulgation par recoupement
 - a. Décrivez comment le produit se rapporte à ceux qui ont été diffusés précédemment.
 - b. Indiquez la ou les dates des demandes de contrôle précédentes liées à cette demande.
 - c. Fournissez les tableaux résiduels en tant que fichiers justificatifs. Veuillez vous reporter à l'orientation en matière de contrôle de la confidentialité.
 - d. Fournissez les deux ensembles de syntaxe et mettez en surbrillance ou précisez les modifications.

Nom du fichier	Notes
1.a_Sortie_données_Onglet 1	Contient les PJ justificatifs pour la sortie de mon tableau de l'onglet 1 soit : le tableau non pondéré, le tableau croisé, la matrice de covariance
1.b_Sortie_données_Onglet 2	Idem mais pour l'onglet 2
1.c_Sortie_données_Onglet 3	Idem mais pour l'onglet 3
Dossier do-file	Contient toute la syntaxe pour obtenir l'échantillon et l'obtention de mes sorties. Lire le README inclus dans le dossier pour savoir à quoi correspond chaque do-file
Liste_variable	Fichier Excel qui nomme toutes mes variables ainsi que les variables utilisées dans EPA pour les construire

Section C : Produit faisant l'objet de la demande de diffusion

EX 2 : Sortie de trois fichiers

Nom du fichier (indiquez chaque feuille pour les fichiers de tableur)	Nom de l'enquête ou de l'ensemble de données et cycle(s)	Le cas échéant, nom de la variable de pondération (indiquez s'il s'agit d'une variable échelonnée ou normalisée)	Précisez la méthode utilisée en choisissant un numéro dans la liste ci-dessous	Description de l'échantillon (p. ex. femmes employées, âgées de 21 à 45 ans, en Ontario)	Niveau de géographie du produit demandé (c.-à-d. national, provincial, etc.)
1.Stat_descrip	ELNEJ (Cycle 1 à 4)	WT68	1	Garçons de 6 à 11 ans	QC
2.MCO	Idem	Idem	4	Idem	Idem
3.Graph	Idem	Idem	3	Garçons vs Filles de 6 à 11 ans	QC vs ON

Types de méthodes

1. Méthodes descriptives (p. ex. modèles de régressions avec une seule variable, fréquences, analyse par recoupement, moyennes et répartitions, matrice des corrélations, analyse de la variance)
2. Méthodes de mise à l'échelle (p. ex. analyse factorielle)
3. Graphiques (p. ex. histogrammes) – n'oubliez pas d'inclure des tableaux à l'appui.
4. Analyse de régression multivariée simple (p. ex. MCO, Logit, Probit, Tobit)
5. Méthodes de modélisation complexes (p. ex. modélisation par équations structurelles, modélisation linéaire hiérarchique, analyse de croissance, analyse de survie, analyse de données longitudinales, modèles d'équations simultanées, modèles à effets fixes, modèle à effets aléatoires)
6. Autres – veuillez décrire (p. ex. notes et fichiers de syntaxe)

Section D : Fichiers justificatifs

EX 2 : Sortie de trois fichiers

Nom du fichier	Notes
1a.Stat_descrip	Contient les PJ justificatifs pour la sortie de mon tableau de l'onglet 1 soit : le tableau non pondéré, le tableau croisé, la matrice de covariance
2a.MCO	Contient les résultats non pondérés, pondérés, pondérés et arrondis
2b.MCO	Contient le R2 et tous autres justificatifs pour la sortie des régressions
3a.Graph	Contient les tableaux pondérés et non pondérés pour l'obtention des graphiques
Dossier do-file	Contient toute la syntaxe pour obtenir l'échantillon et l'obtention de mes sorties. Lire le README inclus dans le dossier pour savoir à quoi correspond chaque do-file
Liste_variable	Fichier Excel qui nomme toutes mes variables ainsi que les variables utilisées dans EPA pour les construire

1. Ne nommez pas les fichiers justificatifs de la même manière que les fichiers que vous voulez sortir
2. Si vous voulez également sortir vos do-files, vous devez les inclure dans le tableau de la section C (vous pouvez dans cette situation, mettre le dossier au complet). Attention, les règles de divulgation doivent toujours être respectées.
3. La liste de la construction de vos variables est obligatoire (nom, label, format, variable d'origine)

2. Faire sa demande de divulgation

- **Étape 2 : Informer l'analyste**

- Par mail à l'adresse du CDR, en mettant le PI en CC
 - statcan.mtl-rdc-mtl-cdr.statcan@statcan.gc.ca (UdeM)
 - statcan.uqam_rdc-uqam_cdr.statcan@statcan.gc.ca (UQAM-INRS)
 - statcan.mcg-rdc-mcg-cdr.statcan@statcan.gc.ca (McGill)
 - statcan.laval_rdc-laval_cdr.statcan@statcan.gc.ca (LAVAL)
 - statcan.sher_rdc-sher_cdr.statcan@statcan.gc.ca (SHERBROOKE)

- **Informations :**

- Le temps de traitement est de 10 jours ouvrables
- Laisser savoir si vous avez une échéance (être disponible s'il y a des modifications)
- Si votre projet contient plusieurs bases de données, vous devez utiliser les règles de diffusion les plus strictes.

Erreurs courantes

- Formulaire de demande erroné ou manquant
- Ne fournis pas les résultats non pondérés et non arrondis, ou ils sont mal indiqués
- Ne fournis pas les tests de contrôle
- Manque de clarté dans les pièces justificatives
- Manque de clarté dans la méthodologie, notamment avec la pondération et l'échantillon
- Non-vérification du recoupement avec les anciennes sorties
- Ne pas sortir les résultats en CSV pour faciliter la vérification de l'analyste

Conseils pratiques

- Lire toute la documentation
- Faire une rencontre avec l'analyste
- Reprendre le formulaire de demande vierge pour les nouvelles demandes
- Éviter de sortir les statistiques descriptives des résultats préliminaires
- Avoir une version « sortable » de vos résultats (enlever les informations justificatives)
- Les résultats ou les chiffres sont-ils vraiment nécessaires ?
 - Possibilité de sortir juste les tendances ou les graphiques sans les coordonnées (plus rapide et moins de risque de recoupement)

Merci!

RESTEZ BRANCHÉS

Balado Hé-coutez bien!

StatsCAN Plus

Le Quotidien

Mon StatCan

Site Web

Enquêtes et programmes statistiques

Centre de service de données



Vous avez des questions? Contactez-nous : infostats@statcan.gc.ca



Statistique
Canada

Statistics
Canada

statcan.gc.ca

Canada

STATISTIQUE CANADA

Éclairer grâce aux données, pour bâtir un Canada meilleur



Statistique
Canada

Statistics
Canada

statcan.gc.ca

Canada 