

RÈGLES

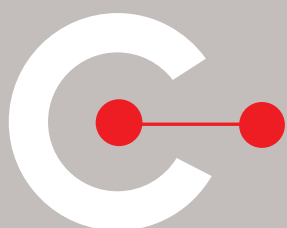
DE

RÈGLES DE CONFIDENTIALITÉ

CONFI-

DENTIA-

LITE



CASD

09/2020

TABLE DES MATIERES

1	DESCRIPTION DU DOCUMENT	2
2	LES TYPES DE SORTIES	2
2.1	LES PROGRAMMES	2
2.2	LES REGRESSIONS, MODELES ECONOMETRIQUES	2
2.3	LES GRAPHIQUES ET LES CARTES	2
2.4	LES TABLEAUX DE DONNEES AGREGES	3
2.5	LES ARTICLES « FINALISES »	3
3	REGLES GENERALES	3
3.1	LES DONNEES « MENAGES »	3
3.2	LES DONNEES « ENTREPRISES »	4
3.3	LES SOURCES MIXTES	4
3.4	SECRET PRIMAIRE	4
3.5	SECRET SECONDAIRE	5
3.5.1	MARGES ET VARIABLES HIERARCHISEES	5
3.5.2	1 SEULE MODALITE NON NULLE	6
3.6	FICHER DE CONTROLE	6
4	REGLES DETAILLEES POUR CHAQUE SOURCE	6
4.1	DONNEES DE L'INSEE	6
4.1.1	DADS (DECLARATION ANNUELLE DE DONNEES SOCIALES)	6
4.1.2	CLAP (CONNAISSANCE LOCALE DE L'APPAREIL PRODUCTIF)	6
4.1.3	FARE (FICHER APPROCHE DES RESULTATS D'ESANE) / FICUS (FICHER DE COMPTABILITE UNIFIE DANS SUSE)	6
4.1.4	SINE (SYSTEME D'INFORMATION SUR LES NOUVELLES ENTREPRISES)	7
4.1.5	RECENSEMENT DE LA POPULATION	7
4.1.6	SIASP (SYSTEME D'INFORMATION SUR LES AGENTS DES SERVICES PUBLICS)	8
4.1.7	FIDELI (FICHERS DÉMOGRAPHIQUES SUR LES LOGEMENTS ET LES INDIVIDUS)	8
4.2	DONNEES DE LA DGFIP – MINISTERE DES FINANCES	9
4.3	DONNEES DE LA DARES – MINISTERE DU TRAVAIL	9
4.4	DONNEES DE LA DEPP – MINISTERE DE L'ÉDUCATION	9
4.5	DONNEES DU SSP – MINISTERE DE L'AGRICULTURE	9
4.5.1	PKGC (PRATIQUES CULTURALES SUR LES GRANDES CULTURES)	9
4.6	DONNEES DE LA DPMA – MINISTERE DE L'AGRICULTURE	10
4.7	DONNEES DU SDSE – MINISTERE DE LA JUSTICE	10
4.7.1	FICHERS CPH, PACS ET FICHER STATISTIQUE DU CASIER JUDICIAIRE NATIONAL	10
4.7.2	FICHER STATISTIQUE DES TRIBUNAUX DE COMMERCE	10
4.8	DONNEES DU SDES – MINISTERE DE L'ENVIRONNEMENT	10
4.8.1	EPTB (ENQUETE SUR LES PRIX DES TERRAINS A BATIR)	10
4.9	DONNEES DU SIES – MINISTERE DE LA RECHERCHE	10
4.10	DONNEES CIFRE DU MESRI – MINISTERE DE LA RECHERCHE	10
4.11	DONNEES DE LA DGE	11
4.12	DONNEES DU CEREQ	11
4.13	DONNEES DE BPIFRANCE	11
4.14	DONNEES DE L'ANIL	11
4.15	DONNEES DE L'ODR	11
4.16	DONNEES DE LA MSA	11
4.17	DONNEES DE L'ACOSS	11
4.18	DONNEES DE LA CNAF	12

1 DESCRIPTION DU DOCUMENT

Suite aux traitements des données auxquelles vous avez été autorisés à accéder, vous serez amenés à opérer des exports de fichiers depuis votre environnement de travail sécurisé CASD. Ce processus est appelé « demande de sortie ». Vous trouverez dans ce document l'ensemble des règles de confidentialité à appliquer aux fichiers que vous souhaitez exporter. Il s'agit de vous aider à respecter les différents types de secret applicables (statistique, fiscal, etc.) et de s'assurer qu'aucune information permettant d'identifier une personne physique, un ménage ou une entreprise ne sera divulguée.

Si votre sortie est de nature à générer une rupture du secret applicable aux données concernées, le CASD vous en informera et vous aidera à masquer des cases concernées ou à agréger des variables différemment pour que le problème soit résolu (zone géographique plus grande, tranche d'âge regroupé au lieu de l'âge fin...). Cependant, **le CASD ne pourra en aucun cas modifier les fichiers que vous avez demandé à sortir.**

2 LES TYPES DE SORTIES

Tout d'abord, vous devez décrire de manière précise le contenu de votre sortie, soit dans votre mail de demande de sortie, soit dans un fichier texte inclus dans la sortie. Vous trouverez une description de la procédure pour effectuer une sortie dans le [guide utilisateur disponible sur le site web du CASD](#).

2.1 LES PROGRAMMES

Vous pouvez demander à sortir vos programmes pour les réutiliser en dehors du CASD. Le programme ne doit pas contenir de données confidentielles.

2.2 LES REGRESSIONS, MODELES ECONOMETRIQUES

Pour les résultats de régressions ou modèles économétriques, indiquez le nombre d'observations afin de permettre au CASD de s'assurer du respect des règles de confidentialité.

2.3 LES GRAPHIQUES ET LES CARTES

Pour les graphiques comme les courbes, les histogrammes, les nuages de points, les diagrammes, etc, vous devez fournir les informations qui ont permis leur construction (population et signification des variables utilisées).

D'autres types de graphiques sont plus complexes à traiter car ils peuvent contenir des informations individuelles. Par exemple, les box plot peuvent contenir, non seulement le maximum et le minimum, mais aussi des points extrêmes (outliers) qui ne doivent pas être identifiés.

Attention aux graphiques d'analyse factorielle représentant les individus s'ils contiennent des individus atypiques qui peuvent être identifiés (par exemple une ACP sur des entreprises où le SIRET est utilisé sur le graphique comme identifiant de chaque point).

Les graphiques **STATA** au format **LIVE** contiennent la base de données qui a servi à constituer le graphique, il est préférable de les convertir dans un autre format (AS-IS : format ne contenant pas les bases de données, pdf, jpg...).

Les cartes peuvent concerner des populations très fines (communales ou infra-communales). Il faut, pour vérifier que les règles de confidentialité sont bien respectées sur la carte, fournir les données qui ont permis de construire cette carte.

2.4 LES TABLEAUX DE DONNEES AGREGES

Vous devez fournir dans la description des sorties toutes les informations permettant de les comprendre et d'identifier les types de données utilisés : liste des variables, leur signification, les effectifs correspondant à chaque case et l'information portant sur la contribution maximale dans la case pour les variables de montants (fichier de contrôle non secrétisé).

De manière générale, les règles de confidentialité sont différentes et propres à chaque type de source.

2.5 LES ARTICLES « FINALISES »

Les articles finalisés que vous souhaitez sortir de votre espace projet CASD ne doivent pas contenir de données sur une population ne respectant pas les règles indiquées ci-après.

Par ailleurs, votre article doit indiquer que votre travail a été rendu possible grâce au dispositif CASD de la manière suivante :

« L'accès à certaines données utilisées dans le cadre de ce travail a été réalisé au sein d'environnements sécurisés du Centre d'accès sécurisé aux données – CASD (Réf. ANR-10-EQPX-17) » ou en anglais **« Access to some confidential data, on which is based this work, has been made possible within a secure environment offered by CASD - Centre d'accès sécurisé aux données (Ref. ANR-10-EQPX-17) »**.

3 REGLES GENERALES

La règle générale du secret statistique est la suivante : pas de diffusion, ni de données individuelles, ni de résultats permettant l'identification indirecte de personnes ou d'entreprises.

3.1 LES DONNEES « MENAGES »

Le secret statistique impose qu'une information n'est publiable que si elle a fait l'objet d'un traitement rendant impossible l'identification des personnes physiques. C'est une obligation qui est également applicable aux entrepreneurs individuels. En outre, et au titre de la protection de la vie privée (article L311-6 CRPA), sans qu'il y ait nécessairement identification de la personne, il ne doit pas être non plus possible de déduire ou d'inférer de façon quasi certaine de nouvelles informations sur un individu, notamment pour des données « sensibles » entendues au sens large (données sur les revenus par exemple).

Ainsi, la seule règle imposée pour les tableaux fournissant des données agrégées est que l'identification directe ou indirecte des individus soit impossible. Dans la pratique, on considère que le secret statistique est respecté si la connaissance d'une caractéristique pour un individu ne peut pas entraîner la connaissance d'une autre caractéristique avec laquelle elle est croisée dans un tableau.

Exemple :

Le tableau ci-dessous donne la répartition par âge et la situation matrimoniale et indique que les personnes entre 50 et 59 ans ont toutes le même état matrimonial « divorcé ». Le secret statistique n'est plus respecté dans ce tableau, et ce dernier n'est donc pas diffusable. En effet, si l'on sait par ailleurs qu'un individu donné se situe dans la tranche des 50-59 ans, le tableau nous informe que cette

personne est divorcée, et ceci même si la case qui croise les modalités « 50 à 59 ans » et « divorcé » comporte plusieurs individus.

Situation matrimoniale et classe d'âge	18-25 ans	26-49 ans	50-59 ans	60 ans et +
Marié	7	27	0	30
Divorcé	0	11	9	22
Autre	21	12	0	4

3.2 LES DONNEES « ENTREPRISES »

Pour les tableaux fournissant des données agrégées sur les entreprises, les règles à suivre s'appliquent aux unités dotées de la personnalité juridique (SIREN) et non aux établissements, et concernent également les entrepreneurs individuels. Elles sont les suivantes :

- Aucune case du tableau ne doit concerner moins de trois SIREN (décision du 13 juin 1980 du directeur général de l'Insee)
- Aucune case du tableau ne doit contenir des données pour lesquelles un SIREN représente plus de 85% du total (règle du CNIS, 7 juillet 1960)

Note : Pas de sortie de liste de SIREN/SIRET associée à d'autres informations.

3.3 LES SOURCES MIXTES

Les sources mixtes sont des sources provenant de combinaisons (appariements) d'enquêtes statistiques et de données administratives ou bien des sources comportant à la fois des informations d'ordre économique et financier (entreprises) et des informations relatives à des faits et comportements d'ordre privé (ménages).

La démarche à adopter face à de telles sources est, dans son principe, très simple : les règles à prendre en considération s'obtiennent par le cumul des règles applicables d'une part aux enquêtes statistiques, d'autre part aux fichiers administratifs. Par exemple, l'enquête revenus fiscaux et sociaux repose sur la combinaison d'enquêtes statistiques, des résultats de l'enquête emploi, de données fiscales et des données fournies par les Caisses d'allocations familiales, ou bien les enquêtes Esane (Élaboration des Statistiques Annuelles d'Entreprises).

3.4 SECRET PRIMAIRE

Les cases qui ne respectent pas les règles de confidentialité et qui sont donc masquées forment ce que l'on appelle le secret primaire. Il est simple à gérer, il faut appliquer les règles propres à chaque source.

Exemple :

Catégorie d'entreprise	Nombre d'entreprises
TPE	7
PME	2
Grandes entreprises	10

↓

Catégorie d'entreprise	Nombre d'entreprises
TPE	7
PME	s
Grandes entreprises	10

3.5 SECRET SECONDAIRE

Le secret secondaire est plus complexe à gérer, il est lié à la présence des marges dans un tableau et empêche la reconstitution, par somme ou par différence, des cases masquées au secret primaire

Les situations qui peuvent mener au secret secondaire sont les suivantes :

3.5.1 Marges et variables hiérarchisées

En raison des marges diffusées dans les tableaux ou publiées sur internet, il est parfois possible de retrouver les cases masquées du secret primaire. Pour y remédier, il suffit de masquer deux cases.

Exemple :

Catégorie d'entreprise	Nombre d'entreprises
TPE	7
PME	2
Grandes entreprises	10
Total	19

Catégorie d'entreprise	Nombre d'entreprises
TPE	s
PME	s
Grandes entreprises	10
Total	19

Les variables hiérarchisées concernent souvent des variables géographiques ou des nomenclatures, il faut les traiter comme des marges et les imbriquer dans les tableaux.

Exemple :

Ces deux tableaux ne peuvent pas être diffusés indépendamment car il y a des relations d'additivité. Si un utilisateur souhaite sortir le premier tableau donnant le nombre d'entreprises en Bretagne, il doit, selon les règles de confidentialité appliquées sur les données entreprises, masquer la case mentionnant le nombre d'entreprises dans le Finistère.

ATTENTION: dans cet exemple, il est impératif que le second tableau, qui donne le nombre d'entreprises au niveau régional et permet de retrouver la case masquée dans le premier, ne fasse jamais l'objet d'une demande de sortie ultérieure ou soit diffusé de quelque manière que ce soit. Vous pouvez par exemple anticiper tout risque potentiel de brèche de confidentialité en masquant deux cases du premier tableau.

Département	Nombre d'entreprises
Morbihan	8
Finistère	2
Côtes-D'Armor	9
Ille-et-Vilaine	6

Région	Nombre d'entreprises
Basse-Normandie	20
Bretagne	25
Pays de la Loire	28

Département	Nombre d'entreprises
Morbihan	8
Finistère	s
Côtes-D'Armor	9
Ille-et-Vilaine	s

3.5.2 1 seule modalité non nulle

Avoir une seule modalité non nulle dans un tableau permet de connaître la caractéristique d'un individu ou d'une entreprise à partir d'une autre. Pour l'éviter, il faut avoir au moins 2 modalités non nulles.

Exemple :

Si, après traitement, on obtient un tableau indiquant que les personnes ayant entre 20 et 34 ans ont toutes la même situation professionnelle (demandeurs d'emploi), le secret statistique n'est plus respecté dans ce tableau. En effet, si l'on sait par ailleurs qu'un individu donné se situe dans la tranche des 20-34 ans, ce tableau nous informe que cette personne est nécessairement demandeuse d'emploi, et ceci même si la case qui croise les modalités « 20-34 ans » et « demandeur d'emploi » comporte plusieurs individus (7 en l'occurrence).

Situation professionnelle et tranche d'âge	Salarié	Demandeur d'emploi	Non salarié	Autre
20-34 ans	0	7	0	0
35-49 ans	6	5	5	8
50-64 ans	4	5	6	5



Situation professionnelle et tranche d'âge	Salarié	Demandeur d'emploi	Non salarié	Autre
20-34 ans	0	s	s	0
35-49 ans	6	5	5	8
50-64 ans	4	5	6	5

3.6 FICHER DE CONTROLE

Afin de vérifier les résultats portant sur les variables de montant, vous devez également transmettre un fichier de contrôle contenant toutes les informations à exporter ainsi que des colonnes indiquant le maximum et le pourcentage du maximum des variables de montant.

Ce fichier de contrôle sera retiré de la sortie avant son envoi.

Exemple :

Fichier de contrôle

	Nombre entreprises	Montant du max	Montant du total	% du max
Bretagne	139	1 668	27 800	6%
Morbihan	82	1 590	19 882	8%
Finistère	19	590	3 476	17%
Côtes-d'Armor	36	1 103	2 567	43%
Ille-et-Vilaine	2	1 312	1 875	70%
Picardie	99	825	20 643	4%
Oise	67	825	13 750	6%
Aisne	5	722	821	88%
Somme	27	790	6072	13%



Fichier de sortie

	Nombre entreprises	Montant total
Bretagne	139	27 800
Morbihan	82	19 882
Finistère	S3	S4
Côtes-d'Armor	36	2 567
Ille-et-Vilaine	S1	S2
Picardie	99	20 643
Oise	67	13 750
Aisne	5	S9
Somme	27	S10

Valeur cachée parce que :

- S1 Nombre d'entreprises inférieur à 3
- S2 Informations relatives à S1
- S3 Pour ne pas recalculer S1 à l'aide du total
- S4 Pour ne pas recalculer S2 à l'aide du total
- S9 Pourcentage du maximum supérieur à 85%
- S10 Pour ne pas recalculer S9 à l'aide du total

4 REGLES DETAILLEES POUR CHAQUE SOURCE

4.1 DONNEES DE L'INSEE

Pour les sources produites par l'Insee, les règles du secret sont décrites dans le guide du secret statistique. Ce guide est disponible à l'adresse :

<https://www.insee.fr/fr/statistiques/fichier/1300624/guide-secret.pdf>

4.1.1 DADS (Déclaration annuelle de données sociales)

Tout tableau diffusé ne doit en aucun cas permettre l'identification directe ou indirecte d'un salarié ou d'un établissement.

- Tableaux au lieu de résidence (logique « ménages »)
 - Aucune case ne doit comporter moins de 5 salariés
 - Aucune case ne doit comporter un salarié représentant plus que 80% de la masse salariale
- Tableaux au lieu de travail (logique « entreprises » en plus), **en plus des deux règles précédentes qui s'appliquent**, s'ajoutent les règles concernant les données sur les entreprises qui sont les suivantes :
 - Aucune case ne doit comporter moins que 3 entreprises ou établissements
 - Aucune case ne doit comporter une entreprise ou un établissement représentant plus de 85% du total

4.1.2 CLAP (Connaissance locale de l'appareil productif)

CLAP fait partie des données « entreprises ». Les indicateurs qui sont soumis au secret statistique sont les effectifs et les rémunérations. Les règles à suivre sont les suivantes :

- Aucune case ne doit concerner moins de 3 unités (l'unité est un établissement ou une entreprise)
- Aucune case ne doit comporter une unité représentant plus que 85% du total
- Aucune case ne doit concerner moins de 5 salariés

4.1.3 FARE (Fichier approché des résultats d'Esane) / FICUS (Fichier de comptabilité unifié dans SUSE)

Les données FARE et FICUS sont des sources mixtes « fiscale et statistique ». Les règles de confidentialité applicables sont donc le cumul des règles applicables d'une part aux enquêtes statistiques, d'autre part aux données fiscales. Elles sont alors les suivantes :

- Aucune case ne doit concerner moins de 3 unités
- Aucune case ne doit comporter une unité représentant plus que 85% du total
- Pour les entreprises individuelles, aucune case ne doit concerner moins de 11 unités

4.1.4 SINE (Système d'information sur les nouvelles entreprises)

SINE est également une source « entreprises », les règles de confidentialité à appliquer sont les suivantes :

- Aucun résultat qui concerne moins de trois entreprises
- Aucune donnée pour laquelle une seule entreprise représente 85 % ou plus de la valeur obtenue

Par ailleurs, les taux de survie ne doivent pas être calculés pour des populations de moins de 20 entreprises. Ce seuil minimum de 20 entreprises est également requis pour les zonages et regroupements particuliers.

4.1.5 Recensement de la Population

Les principes de diffusion des données du RP ont évolué à partir de 2006 avec l'arrêté du 19 juillet 2007 relatif à la diffusion des résultats du recensement de la population. Celui-ci est venu en remplacement de l'arrêté du 22 mai 1998, modifié le 8 avril 2002, relatif à la diffusion des résultats du recensement de la population.

→ En raison des taux de sondage des différentes exploitations des recensements de la population, on retiendra les seuils minimums de diffusion suivants pour les tableaux de résultats :

- Pour le RP jusqu'en 1999 :
 - RP exploitation principale : au moins 4 unités
 - RP exploitations complémentaires :

Année du RP	Taux de sondage	Seuil minimal de diffusion (après pondération)
1962	1/20	80 unités
1968	1/4	16 unités
1975	1/5	20 unités
1982	1/4	16 unités
1990	1/4	16 unités
1999	1/4	16 unités

- Pour le RP rénové (après 2006) : au moins 10 unités après pondération

→ Respecter les seuils suivants de diffusion des variables dites « sensibles » : nationalité actuelle (ou à la naissance), lieu de naissance, lieu de résidence antérieur, année (ou ancienneté) d'arrivée en France. Les notions d'immigrés et de français par acquisition sont dans le champ des variables sensibles jusqu'au recensement de 1999.

Zones minimales de diffusion de ces variables, seuil après pondération :

1999 et avant	A partir de 2006 (RP annuel)
Communes de plus de 5 000 habitants	Communes de plus de 5 000 habitants
Seuil de <u>10 000 habitants</u> pour les arrondissements, zones d'emploi, aires urbaines, unités urbaines (ou leurs regroupements) et zones de la politique de la ville ou leurs regroupements	Seuil de <u>5.000 habitants</u> pour les arrondissements, zones d'emploi, aires urbaines, unités urbaines et zones de la politique de la ville

Zones infracommunales résultant du regroupement de 3 quartiers (zone fixe résultant du découpage de la commune en zones géographiques d'un seul tenant d'environ 2 000 habitants)	Zones infracommunales résultant du regroupement de 3 quartiers (zone fixe résultant du découpage de la commune en zones géographiques d'un seul tenant d'environ 2 000 habitants)
Département pour l'année (ou ancienneté) d'arrivée	Département pour l'année (ou ancienneté) d'arrivée

Pour les sorties qui concernent les variables sensibles du recensement, afin de vérifier les fichiers à sortir, vous devez préciser les échelles géographiques utilisées pour construire les données. Dans le cas où ces dernières ont été agrégées au niveau des communes, vous devez fournir le nombre d'habitants dans chaque commune ou groupe de communes pour toutes les variables concernées.

4.1.6 SIASP (Système d'information sur les agents des services publics)

La diffusion de résultats statistiques tirés des données issues de SIASP doit être conforme aux dispositions prévues par les textes relatifs au secret en matière de statistique et à la protection des données individuelles. En particulier, aucun tableau destiné à la diffusion ne doit permettre l'identification directe ou indirecte d'un salarié ou d'un établissement.

Pour les tableaux au lieu de résidence les critères sont les suivants :

- Aucune case ne doit comporter moins de 5 salariés
- Aucune case ne doit comporter un salarié représentant plus de 80% de la masse salariale de la case

Pour les tableaux au lieu de travail, **en plus des deux critères précédents**, les critères sont les suivants :

- Aucune case ne doit se rapporter à moins de 3 établissements
- Aucun établissement ne doit représenter plus de 85 % de la grandeur étudiée dans la case

Exception : pour le champ de la fonction publique d'État, aux niveaux national, régional ou départemental, il n'est pas toujours possible de trouver 3 établissements sirétisés. Par conséquent, cette règle ne sera pas appliquée, uniquement dans ce contexte.

4.1.7 FIDELI (Fichiers Démographiques sur les Logements et les Individus)

Les règles pour FIDELI sont les suivantes :

- Pour les communes de 5000 habitants ou plus, la maille géographique la plus fine pour produire des tableaux est la commune ou l'IRIS ;
- Pour les communes de moins de 5000 habitants, la maille géographique la plus fine pour produire des tableaux est l'Établissement public de coopération intercommunale (EPCI). Les EPCI, pour lesquels la population totale des communes de moins de 5000 habitants dans l'EPCI est inférieure à 2000 habitants, ne peuvent toutefois donner lieu à la production de tableaux ;
- Pour les résultats portant sur les quartiers prioritaires de la politique de la ville (QPV), la maille géographique la plus fine pour produire des tableaux est la région ;
- Pour tout tableau produit, chaque case doit comporter plus de 11 individus.

Pour les sorties qui concernent des données issues de FIDELI, afin de vérifier les fichiers à sortir, vous devez préciser les échelles géographiques utilisées pour construire les données de votre sortie. Dans

Le cas où les données ont été agrégées au niveau des communes ou des EPCI, vous devez fournir le nombre d'habitants dans chaque commune pour toutes les variables concernées.

4.2 DONNEES DE LA DGFIP – MINISTERE DES FINANCES

Pour les données sur les entreprises :

- Aucune case du tableau ne doit concerner moins de trois unités (décision du 13 juin 1980 du directeur général de l'Insee)
- Aucune case du tableau ne doit contenir des données pour lesquelles une entreprise représente plus de 85% du total (règle du CNIS, 7 juillet 1960)

Pour les données sur les ménages : aucune case ne doit comporter moins de 11 individus.

Pour les données de nature fiscale (règles avalisées par la CNIL dans un avis du 27 mai 1997 et énoncées au §30 du BOI-DJC-CADA-20 :

- Règle du nombre d'unités : une donnée agrégée n'est pas communiquée lorsqu'elle concerne moins de onze unités
- Règle du poids des unités : une donnée agrégée n'est pas communiquée lorsqu'elle comprend un élément dominant qui représente plus de 85 % du montant agrégé

4.3 DONNEES DE LA DARES – MINISTERE DU TRAVAIL

Les règles pour les données de la DARES sont les suivantes :

- Pour les données individus/ménages : respecter un effectif minimum de 5 individus par case
- Pour les données entreprises : respecter la règle de 3 entreprises au minimum par case et aucune case ne doit contenir des données pour lesquelles une entreprise représente plus de 85% du total

4.4 DONNEES DE LA DEPP – MINISTERE DE L'ÉDUCATION

Les règles pour les données de la DEPP sont les suivantes :

- Pour les données sur les individus : aucune case avec moins de 10 individus
- Pour les données agrégées au niveau établissement (collèges, lycée, etc.) : les cases ne doivent pas contenir moins de 10 établissements

4.5 DONNEES DU SSP – MINISTERE DE L'AGRICULTURE

Le secret s'applique sur les exploitations et non sur les parcelles (exploitation = entreprise et parcelle = établissement). Les règles sont les suivantes :

- Aucune case ne doit concerner moins de 3 exploitations
- Aucune case du tableau ne doit contenir de données pour lesquelles une exploitation représente plus de 85% du total

4.5.1 PKGC (Pratiques culturelles sur les grandes cultures)

- Lorsque l'effectif total d'un tableau porte sur moins de 30 parcelles au total, celui-ci ne peut pas être diffusé
- Aucune case d'un tableau ne doit concerner moins de trois parcelles agrégées

4.6 DONNEES DE LA DPMA – MINISTERE DE L’AGRICULTURE

Aucune information (case, item de graphique/carte, etc.) ne doit concerner strictement moins de 5 navires uniques (champ CFR_COD).

4.7 DONNEES DU SDSE – MINISTERE DE LA JUSTICE

4.7.1 Fichiers CPH, PACS et fichier statistique du casier judiciaire national

Au moins 5 unités statistiques pour chaque case diffusée ou calculable par recoupement d’informations à partir des données diffusées. Cela implique que, dans le cas où toutes les modalités d’une variable ne sont pas présentées, ce minimum de 5 unités s’applique également au complémentaire à l’effectif des modalités présentées sur l’effectif total de la population étudiée. Cela implique également qu’il ne doit pas être fait mention d’une distribution unanime d’une caractéristique (autre que l’application d’un critère légal) au sein d’une sous-population.

4.7.2 Fichier statistique des tribunaux de commerce

Deux conditions :

- Au moins 3 unités statistiques pour chaque case diffusée ou calculable par recoupement d’informations à partir des données diffusées
- Au sein de chaque case diffusée ou calculable, aucune unité statistique ne peut avoir de contribution supérieure à 85 % à l’agrégat présenté

Ces deux règles s’appliquent également, dans le cas où toutes les modalités d’une variable ne sont pas présentées, aux effectifs et contributions des modalités complémentaires à celles présentées.

4.8 DONNEES DU SDES – MINISTERE DE L’ENVIRONNEMENT

4.8.1 EPTB (Enquête sur les prix des terrains à bâtir)

A minima un effectif de 11 permis dans chaque case du tableau. Il faut se référer à l’effectif non pondéré.

4.9 DONNEES DU SIES – MINISTERE DE LA RECHERCHE

Les règles pour les données du SIES sont les suivantes :

- Pour les données sur les établissements éducatifs ou les entreprises :
 - Aucune case du tableau ne doit concerner moins de trois unités
 - Aucune case du tableau ne doit contenir des données pour lesquelles un établissement représente plus de 85% du total
- Pour les données sur les étudiants, l’identification directe ou indirecte des individus doit être impossible. Dans la pratique, on considèrera que le secret statistique est respecté si aucune case du tableau ne concerne moins de cinq individus.

4.10 DONNEES CIFRE DU MESRI – MINISTERE DE LA RECHERCHE

Les règles pour les données CIFRE du MESRI sont les suivantes :

- S’agissant des particuliers, il est interdit de publier des données qui permettraient une identification directe ou indirecte d’une personne ou même, sans pouvoir l’identifier, d’obtenir une information à son sujet. Ces règles limitent la finesse des informations disponibles en diffusion. Des règles très strictes sont fixées spécifiquement pour les recensements de la

population. Pas de sortie de données sur moins de 10 observations en ce qui concerne les individus.

- S'agissant des entreprises, aucun résultat n'est publié s'il concerne moins de trois entreprises, ni aucune donnée pour laquelle une seule entreprise représente 80% ou plus de la valeur obtenue. Cependant, il est admis que la diffusion de listes extraites du répertoire des entreprises ou des établissements peut mentionner l'activité économique, une classe d'effectifs et une tranche de chiffres d'affaires.
- Pas d'édition de listes donnant le nom, l'adresse ou autre caractéristique individuelle des étudiants bénéficiaires ou des entreprises.

4.11 DONNEES DE LA DGE

Aucune case du tableau ne doit :

- Concerner moins de trois unités
- Contenir de données pour lesquelles une entreprise représente plus de 85% du total

4.12 DONNEES DU CEREQ

A minima un effectif de 5 observations dans chaque case du tableau.

4.13 DONNEES DE BPIFRANCE

A minima un effectif de 10 observations dans chaque case du tableau.

4.14 DONNEES DE L'ANIL

Les règles pour les données de l'ANIL sont les suivantes :

- Pas de diffusion de données au logement
- A minima un effectif de 50 observations par résultat (case d'un tableau agrégé, données pour calibration de modèles)

4.15 DONNEES DE L'ODR

Ne pas sortir des informations des données à une échelle géographique inférieure au canton en respectant la règle de 3 observations minimum par case.

Pour les sorties qui concernent les données de l'ODR, afin de vérifier les fichiers à sortir, vous devez préciser les échelles géographiques utilisées pour construire les données de votre sortie.
--

4.16 DONNEES DE LA MSA

Les règles pour les données de la MSA sont les suivantes :

- Aucune case du tableau ne doit concerner moins de 5 unités
- Aucune case du tableau ne doit contenir des données pour lesquelles un cotisant non-salarié représente plus de 85% du total
- La caractéristique pour un individu ne peut pas entraîner la connaissance d'une autre caractéristique avec laquelle elle est croisée dans un tableau

4.17 DONNEES DE L'ACOSS

A minima un effectif de 10 observations dans chaque case du tableau.

4.18 DONNEES DE LA CNAF

A minima un effectif de 5 observations dans chaque case du tableau.

CASD 