

GUIDE DU SECRET STATISTIQUE

Le secret statistique est une forme particulière du **secret professionnel** qui s'applique aux organismes qui relèvent de [la statistique publique](#).

Le secret statistique garantit aux personnes qui fournissent des informations utilisées pour l'établissement de statistiques l'assurance que ces informations ne seront pas utilisées d'une façon susceptible de leur porter tort.

En effet, le secret statistique interdit au dépositaire des informations :

- de communiquer à des tiers des renseignements individuels recueillis par voie d'enquête statistique ;
- de divulguer des informations qui lui ont été transmises par des tiers à des fins exclusives d'établissement de statistiques.

Le secret statistique garantit ainsi le respect :

- de la **confidentialité due à la vie privée, personnelle et familiale**, pour les personnes physiques ;
- du **secret commercial et des affaires**, pour les entreprises.

Les obligations relatives au secret statistique relèvent de textes de niveaux national et européen.

À l'échelon national, la [loi n° 51-711 du 7 juin 1951](#) modifiée sur l'obligation, la coordination et le secret en matière de statistiques détermine ce qu'est le secret statistique, ses limites et ses conditions d'application. La loi traite précisément de *l'obligation et du secret statistique*. Le service statistique public effectue un certain nombre d'enquêtes qui n'ont pas de caractère obligatoire, néanmoins les règles relatives au secret statistique s'appliquent aussi bien à ces enquêtes qu'aux enquêtes obligatoires.

À l'échelon européen, la confidentialité des informations statistiques est affirmée par [l'article 338](#) du traité de l'Union. Le secret statistique fait l'objet du chapitre V du règlement [n° 223/2009](#) modifié et du règlement d'application [n° 557/2013](#) en ce qui concerne l'accès aux données confidentielles à des fins statistiques.



Sommaire

Table des matières

Règles pratiques pour le respect du secret statistique	3
A. Informations tirées d'enquêtes statistiques	3
1. Tableaux agrégés	3
2. Fichiers de données individuelles.....	3
B. Informations tirées de sources administratives	6
C. Le cas des sources mixtes	6
1. Fichiers combinant des données statistiques et des données administratives.....	6
2. Fichiers combinant des données sur les entreprises et des données sur les ménages.....	7
Définitions	8

Règles pratiques pour le respect du secret statistique

Il faut distinguer les informations collectées au moyen d'enquêtes statistiques et celles qui ont été transmises par des tiers au service statistique public à des fins d'établissement des statistiques (données administratives).

Les enquêtes statistiques sont celles qui ont reçu le visa prévu par l'article 2 de la loi [n° 51-711 du 7 juin 1951](#) sur l'obligation, la coordination et le secret en matière de statistiques et qui figurent dans la liste des enquêtes de la statistique publique publiée chaque année au *Journal Officiel*.

A. Informations tirées d'enquêtes statistiques

Le secret doit être respecté dans les tableaux agrégés et les fichiers de données individuelles.

1. Tableaux agrégés

- a) Pour les tableaux fournissant des données agrégées sur les **entreprises**, la règle est la suivante :
 - 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 de données pour lesquelles une entreprise représente plus de 85% du total (règle de diffusion définie le 7 juillet 1960 par le Comité de coordination des enquêtes statistiques, prédécesseur du CNIS, Conseil National de l'Information Statistique).
- b) Pour les tableaux fournissant des données agrégées sur les **personnes physiques**, le secret statistique impose qu'on ne puisse pas connaître ou déduire des informations les concernant. Par exemple, si un tableau donne pour une commune la répartition par âge et situation matrimoniale, et que toutes les personnes âgées de 50 à 59 ans ont toutes pour état matrimonial « divorcé », le secret statistique n'est plus respecté dans ce tableau. En effet, si l'on sait par ailleurs que quelqu'un a entre 50 et 59 ans, le tableau nous informe que cette personne est divorcée.

Des règles spécifiques de diffusion concernant le recensement de la population ont été édictées, compte tenu de son exhaustivité dans les petites communes. Ces règles de diffusion sont regroupées dans [l'arrêté du 19 juillet 2007](#) relatif à la diffusion des résultats du recensement de la population.

2. Fichiers de données individuelles

- a) Il n'est pas possible de fournir des **données individuelles sur les entreprises** en respectant strictement l'anonymat, sauf à supprimer du fichier des variables

considérées comme essentielles pour l'analyse de tels fichiers : activité économique, localisation, indicateur de la taille.

Les données d'ordre économique et financier collectées par voie d'enquêtes statistiques sont librement accessibles à l'issue d'un délai de vingt-cinq ans suivant la date de réalisation de l'enquête. Pour répondre aux besoins d'informations ponctuelles sur les entreprises avant l'expiration de ce délai, la loi a créé en 1984 un [comité du secret statistique](#). Celui-ci peut proposer des dérogations d'accès aux données individuelles sur les entreprises.

- b) Les **enquêtes auprès des ménages**¹ réalisées par l'Insee donnent en général naissance à deux types de fichiers de données individuelles :
- Des fichiers « grand public » mis à disposition sur le site internet de l'Insee. Ces fichiers sont anonymisés, autrement dit construits de telle sorte qu'il n'est pas possible d'identifier, directement ou indirectement, un individu. Pour cela, un certain nombre de variables sont retirées du fichier d'origine : commune de résidence, profession détaillée, voire certaines autres variables propres à l'enquête qui permettraient de reconnaître une personne.
 - Des fichiers dits de « production et de recherche » (FPR). Ces fichiers sont également anonymisés, mais conservent toutefois des informations plus fines que celles qui figurent dans le fichier « grand public », pour répondre aux besoins spécifiques des chercheurs et d'organismes publics comme les services statistiques ministériels. Pour éviter tout risque de rupture de confidentialité par croisement avec d'autres sources de données, la mise à disposition à ces fichiers n'est possible que par l'intermédiaire du comité du secret statistique, selon une procédure similaire, quoique simplifiée, à celle mise en œuvre pour les données individuelles non anonymisées (voir ci-dessous).

Les données relatives à la vie personnelle et familiale, et plus généralement aux faits et comportements d'ordre privé sont librement accessibles à l'issue d'un délai de soixante-quinze ans, ou de vingt-cinq ans après le décès de la personne concernée, si ce délai est plus bref. Ainsi, les bulletins du recensement de la population de 1931 sont librement accessibles (sauf pour les personnes mineures à cette date), lorsqu'ils sont présents dans les centres des archives.

Avant l'expiration de ces délais, il est possible d'accéder aux données individuelles non anonymisées obtenues à partir d'enquêtes auprès des ménages, après avis du

¹ Hormis les recensements et les enquêtes annuelles de recensement, qui font l'objet de dispositions spéciales

[comité du secret statistique](#) et sur décision de l'administration des archives. La procédure est la même que pour l'accès aux informations individuelles sur les entreprises (voir plus haut). Dans le cas d'informations relatives aux faits et comportements d'ordre privé, la dérogation aux règles du secret statistique n'est accordée qu'à des demandes effectuées à des fins de statistique publique ou de recherche scientifique ou historique.

L'Insee a fait connaître au comité du secret statistique, lors de sa réunion du 6 octobre 2009, les conditions mises pour son accord² à la communication de données individuelles directement ou indirectement nominatives sur les ménages :

- Pour les services statistiques ministériels (SSM), l'accord est donné sous réserve que le SSM s'engage sur des mesures qui garantissent, au sein de leur ministère, le contrôle de l'accès à ces fichiers et leur utilisation à des fins exclusivement statistiques.
- Pour les autres demandeurs, l'accès aux informations indirectement nominatives s'effectue par le biais du Centre d'accès sécurisé à distance (CASD). Le CASD, via une infrastructure dédiée hautement sécurisée, permet de mettre à disposition, via Internet, des fichiers de données individuelles en empêchant toute extraction ou édition des données. L'accès à ces postes n'est possible que pour les personnes à qui une autorisation nominative a été accordée. Ces personnes signent toutes obligatoirement un engagement de confidentialité, qui les contraint au respect des règles de sécurité et du secret et les expose à de lourdes sanctions en cas de manquement.
- Le centre d'accès sécurisé permet aux chercheurs d'accéder aussi bien à des données d'enquêtes ménages que d'enquêtes entreprises, ainsi qu'à des informations d'origine administrative, indirectement nominatives, cédées à l'Insee ou à un SSM au titre de l'article 7 bis de la loi n° 51-711 du 7 juin 1951, et dans les conditions prévues à son article 7 ter.
- L'accès à des informations directement nominatives est examiné au cas par cas par le comité de direction de l'Insee. L'accès n'est en principe accordé que pour répondre à des demandes de constitution d'échantillon, pour réaliser une enquête ayant reçu le visa du ministre prévu par l'article 2 de la loi de 1951.

² Accord requis pour que l'administration des archives puisse délivrer l'autorisation de communication (article [L.213-3](#) du code du patrimoine).

B. Informations tirées de sources administratives

Les informations transmises à l'Insee ou aux services statistiques ministériels à des fins d'établissement des statistiques, en dehors des enquêtes statistiques, sont également couvertes par le secret statistique. Ceci est explicitement prévu dans l'article 7 bis de la [loi n° 51-711 du 7 juin 1951](#) sur l'obligation, la coordination et le secret en matière de statistiques.

Les règles du secret qui s'attachent à la diffusion de résultats issus de l'exploitation de ces fichiers peuvent être différentes d'une source à l'autre. Certains producteurs de données définissent des règles qui vont au-delà de ce que demandent les règles du secret statistique (voir plus haut). Ainsi, pour la diffusion de tableaux tirés d'informations fiscales, la règle est qu'aucune case ne doit comporter moins de onze individus. Pour les tableaux tirés des déclarations annuelles de données sociales (DADS), aucune case ne doit concerner moins de cinq individus.

Il convient donc se renseigner, avant toute publication, sur les règles de diffusion fixées par l'organisme qui a communiqué les informations. En général, ces règles sont inscrites dans la convention qui a permis la transmission des données.

L'accès aux données aux données individuelles issues de sources administratives est interdit sauf dérogations accordées pour les besoins de la recherche ou la réalisation d'études selon des procédures comparables à celles qui permettent l'accès aux données individuelles issues d'enquêtes statistiques. L'avis de l'administration ou de la personne morale ayant procédé à la collecte des données concernées peut être recueilli avant cette communication.

Suite à la [loi n° 2016-1321 du 7 octobre 2016](#) pour une République numérique en son article 36, l'article L.311-8 du code des relations publiques entre le public et l'administration ouvre désormais, en réponse à une demande d'accès à une base de données, à l'administration détenant la base ou à l'administration des Archives, la possibilité de solliciter l'avis du comité du secret. Le demandeur s'engage à respecter la confidentialité des données communiquées et les règles de secret applicables à la source. Le comité du secret peut recommander la mise en place d'un dispositif d'accès sécurisé, à distance ou sur place.

C. Le cas des sources mixtes

Par sources mixtes, on entend :

- Les sources provenant de combinaisons d'enquêtes statistiques et de données administratives ;
- Les 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)

1. Fichiers combinant des données statistiques et des données administratives

Les règles à prendre en considération s'obtiennent par la combinaison des règles applicables à chacun des deux types de sources.

On peut citer parmi ces sources l'enquête Revenus fiscaux et sociaux, qui apparie les résultats de l'enquête emploi, des données fiscales et des données fournies par les Caisses d'allocations familiales, ou bien le dispositif É sane (Élaboration des statistiques annuelles d'entreprises), qui repose sur la combinaison d'enquêtes statistiques et de données fiscales.

Il est à noter que la diffusion de données individuelles d'origine fiscale est interdite par le livre des procédures fiscales ([article L.103](#)). Des exceptions néanmoins sont prévues. L'[article L.135 D](#) du livre des procédures fiscales prévoit ainsi la possibilité pour les agents du fisc et des douanes de communiquer :

- Toute information aux agents de l'Insee et des SSM, à des fins exclusives d'établissement de statistiques ;
- Les données issues des procédures de contrôle et de recouvrement à des fins de recherche scientifique, dans les conditions de la loi de 1951, c'est-à-dire après avis du comité du secret statistique ;
- Les données d'entreprises, à des fins exclusives de réalisation d'études économiques, aux agents de services de l'État chargés de la réalisation d'études économiques.

2. Fichiers combinant des données sur les entreprises et des données sur les ménages

Ces fichiers peuvent provenir d'enquêtes statistiques (ECMOSS, enquête sur le coût de la main d'œuvre et la structure des salaires) ou de sources administratives (Déclarations annuelles de données sociales).

La demande de dérogation au secret statistique se fait auprès de ce comité en une seule fois. Les conditions d'accès aux données associent les conditions décrites d'une part pour les données entreprises, d'autre part pour les données ménages.

Définitions

Renseignement individuel

Est considéré comme renseignement individuel toute information par laquelle il est possible d'identifier, directement ou indirectement, l'unité à laquelle l'information se rapporte (voir plus bas). Cette définition vaut pour toutes les unités statistiques, qu'il s'agisse de personnes morales ou de personnes physiques.

Données à caractère personnel

Ces termes sont définis à l'article 4 du règlement général sur la protection des données ([règlement 2016/679 du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données](#)). On entend par « données à caractère personnel » toute information se rapportant à une personne physique identifiée ou identifiable. Est réputée être une « personne physique identifiable » une personne physique qui peut être identifiée, directement ou indirectement, notamment par référence à un identifiant, tel qu'un nom, un numéro d'identification, des données de localisation, un identifiant en ligne, ou à un ou plusieurs éléments spécifiques propres à son identité physique, physiologique, génétique, psychique, économique, culturelle ou sociale.

Identification directe ou indirecte

Ces termes sont définis à l'article 3 du règlement [n° 223/2009](#) modifié :

- Identification directe : identification d'une unité statistique à partir de son nom ou de son adresse, ou d'un numéro d'identification accessible au public ;
- Identification indirecte : identification d'une unité statistique par tout autre moyen que l'identification directe.

« Données sensibles »

Les termes de « données sensibles » réfèrent, selon l'article 9 du [règlement général sur la protection des données](#) :

- Aux données à caractère personnel qui révèlent l'origine raciale ou ethnique, les opinions politiques, les convictions religieuses ou philosophiques ou l'appartenance syndicale ;
- Aux données génétiques ou biométriques aux fins d'identification unique ;
- Aux données concernant la santé, la vie sexuelle ou l'orientation sexuelle d'une personne physique.



Par principe, le traitement de telles données est interdit. Des dérogations sont néanmoins possibles, notamment lorsque ce traitement est nécessaire à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques.