

La statistique publique entre code de bonnes pratiques et promesses de nouvelles données massives

Dominique Bureau



« La statistique publique a pour mission de fournir à tout un chacun des informations de qualité, élaborées en toute indépendance, pour éclairer le débat public et les choix publics et privés ».

« L'Autorité de la statistique publique ...émet tout avis qu'elle estime utile pour garantir le respect du principe d'indépendance professionnelle dans la conception, la production et la diffusion de statistiques publiques ainsi que des principes d'objectivité, d'impartialité, de pertinence et de qualité des données produites. Elle s'assure du respect, par le service statistique public, des principes du code de bonnes pratiques de la statistique »

Le Code de bonnes pratiques de la statistique

Le code de bonnes pratiques de la statistique européenne est la pierre angulaire du cadre qualité commun du système statistique européen. Il s'agit d'un instrument d'autorégulation fondé sur des principes couvrant l'environnement institutionnel, les processus statistiques et les résultats statistiques.

Nous considérons que la qualité est notre principal atout dans un monde où l'instantanéité de l'information va croissant et où la preuve de la qualité fait souvent défaut. Nos travaux se caractérisent par l'indépendance professionnelle, l'impartialité du traitement réservé à tous nos utilisateurs, l'objectivité, la fiabilité, la confidentialité statistique et l'efficacité. L'élaboration, la production et la diffusion de nos statistiques reposent sur des méthodologies éprouvées, sur des normes internationales d'excellence et sur des procédures appropriées, bien documentées et transparentes. En matière de qualité, nous appliquons les principes suivants: la pertinence, l'exactitude, l'actualité et la ponctualité, l'accessibilité et la clarté, ainsi que la comparabilité et la cohérence.

Nous nous efforçons constamment de minimiser la charge pesant sur les répondants, d'entretenir une bonne collaboration avec les fournisseurs de données et de travailler en étroite coopération avec les différentes parties prenantes, y compris avec les communautés scientifiques.

Principes

Principe 1: Indépendance professionnelle.

Principe 2: Mandat pour la collecte de données et l'accès aux données.

Principe 3: Adéquation des ressources.

Principe 4: Engagement sur la qualité.

Principe 5: Secret statistique et protection des données.

Principe 6: Impartialité et objectivité.

Principe 7: Méthodologie solide.

Principe 8: Procédures statistiques adaptées.

Principe 9: Charge non excessive pour les déclarants.

Principe 10: Rapport coût-efficacité.

Principe 11: Pertinence.

Principe 12: Exactitude et fiabilité.

Principe 13: Actualité et ponctualité.

Principe 14: Cohérence et comparabilité.

Principe 15: Accessibilité et clarté.

Les données massives: de nombreuses opportunités potentielles

- *...pour améliorer la connaissance statistique:*

- *Consommation (prix, conjoncture), recensement, activités de services (transports, culture, santé, tourisme...)*
- *Et réduire les délais, affiner les échelles...*

- *...grâce à la mobilisation de données telles que:*

- *Caisses des commerces, téléphonie, cartes bancaires, requêtes internautes, plateformes de réservations, satellites*

- *...mais aussi en termes de méthodes:*

- *Traitement-redressement des données, utilisation de données structurées de manière complexe, visualisation...*

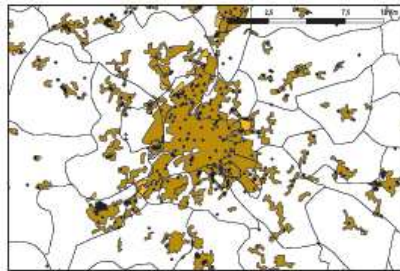
Exemple 1: l'utilisation des données de caisse pour les indices de prix

- *Phase expérimentale, en production à l'horizon 2020*
- *1° application des cadres de la loi numérique pour l'accès aux données (conventions avec les enseignes)*
- *pas de changement des concepts (indice de Laspeyres)*
- *...mais (outre l'enjeu en termes de coût des relevés), améliorations en perspective (intégration des remises, nouveaux produits...)*

Exemple 2: les espaces consommés par le bâti (1)

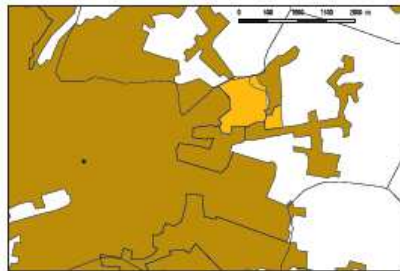
1. Catégoriser les permis de construire par type de construction

Tous les permis de construire de la zone déposés entre 2005 et 2013



Limites de communes
 Tampons de 300 mètres
• Permis
 Zone d'artificialisation de masse
 Zone bâtie en 2005

Type 1 – Densification ou renouvellement du bâti



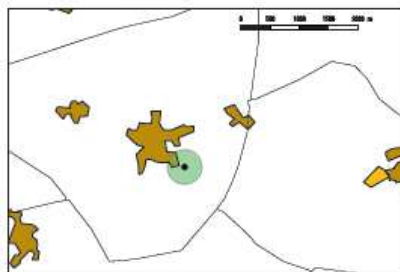
Lecture : construction dans une zone déjà bâtie. Permis de construire situé dans une zone déjà artificialisée en 2005.

Type 2 – Artificialisation de masse



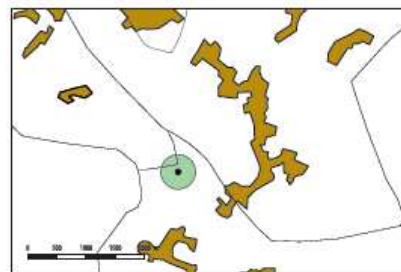
Lecture : construction appartenant à une artificialisation de masse. Permis situé dans un polygone de changement de CLC dont le code d'arrivée correspond à du bâti.

Type 3 – Continuité du bâti



Lecture : construction qui se trouve à proximité d'une zone déjà bâtie. Permis qui ne correspond pas au type 2 mais qui est à moins de 300 mètres d'une zone bâtie en 2005.

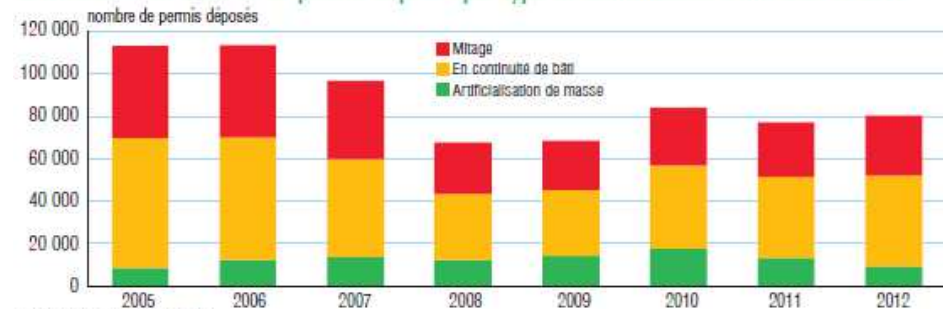
Type 4 – Mitage



Lecture : construction isolée par rapport aux autres constructions. Permis qui ne correspond pas au type 2 et qui est à plus de 300 mètres d'une zone bâtie en 2005.

Sources : SDES, Sit@del, CORINE Land Cover ; IGN, BD CARTO®.

5. Évolution du nombre de permis déposés par type de construction entre 2005 et 2012

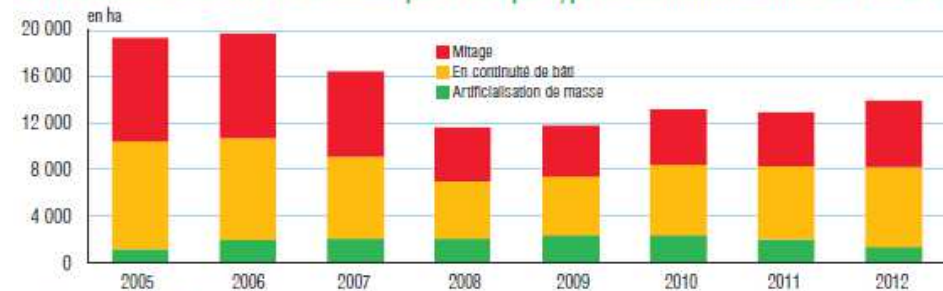


Champ : France métropolitaine.

Lecture : en 2005 plus de 110 000 permis de construire pour construction nouvelle ont été déposés dans des zones non encore bâties.

Sources : SDES, Sit@del, CORINE Land Cover ; IGN, BD CARTO® ; DGFIP, Fichiers fonciers.

6. Évolution de la surface consommée par le bâti par type de construction entre 2005 et 2012



Champ : France métropolitaine.

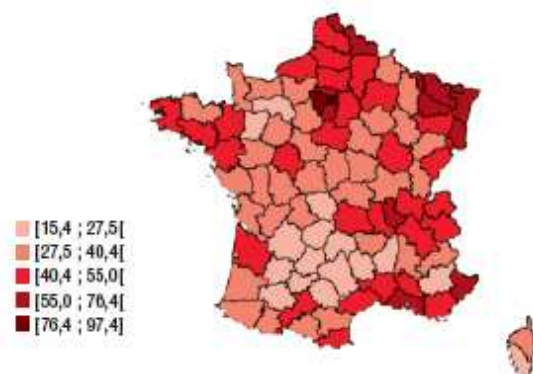
Lecture : en 2005, près de 20 000 hectares ont été consommés par le bâti, 46 % de cette surface correspond à du mitage, 48 % à des constructions en continuité de bâti et 6 % à de l'artificialisation de masse.

Sources : SDES, Sit@del, CORINE Land Cover ; IGN, BD CARTO® ; DGFIP, Fichiers fonciers.

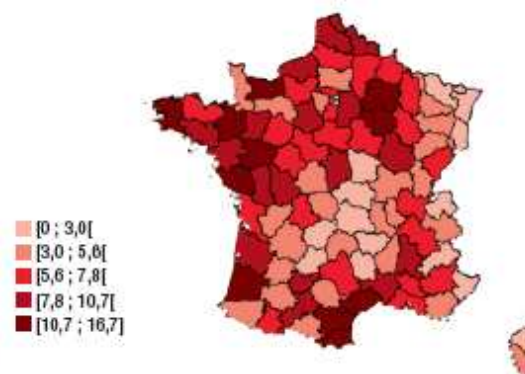
Exemple 2: les espaces consommés par le bâti (2)

3. Répartition géographique des types de construction entre 2005 et 2013 (en % de surface consacrée au bâti)

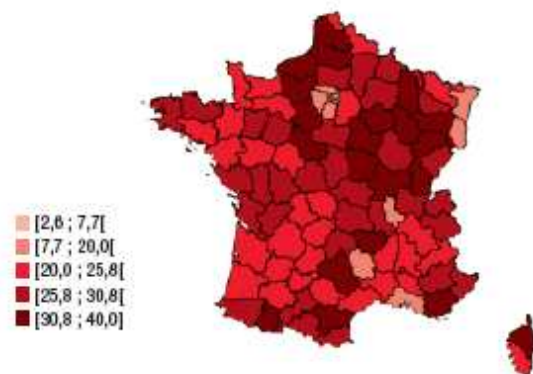
a. Densification (moyenne nationale : 42 %)



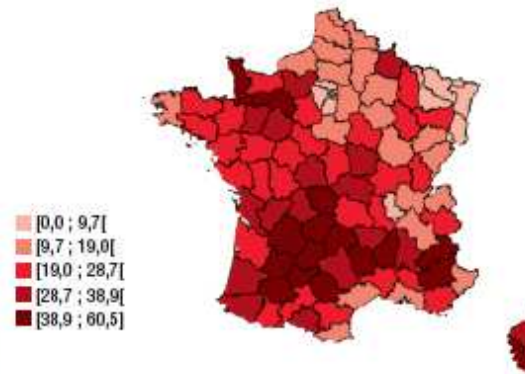
b. Artificialisation de masse (moyenne nationale : 8 %)



c. Continuité du bâti (moyenne nationale : 26 %)



d. Mitage (moyenne nationale : 24 %)



Lecture : entre 2005 et 2013, dans le Finistère, 43 % de la surface consacrée au bâti correspond à de la densification du bâti, 11 % à de l'artificialisation de masse, 28 % à de la construction en continuité de bâti et 18 % à du mitage.

Sources : SDES, Sit@del, CORINE Land Cover ; IGN, BD CARTO® ; DGFIP, Fichiers fonciers.

Les enjeux en termes de gouvernance : principes du code particulièrement sollicités

- 2- mandat pour la collecte des données*
- 4- engagement sur la qualité,*
- 5- secret statistique,*
- 7- méthodologie solide*
- 8- procédures statistiques adaptées,*
- 12- exactitude et fiabilité,*
- 14- cohérence et comparabilité,*
- 15- accessibilité et clarté.*

Conditions d'accès aux données privées (1)

Principe 2: Mandat pour la collecte de données et l'accès aux données. *Les autorités statistiques disposent d'un mandat légal clair les habilitant à collecter et à accéder à des informations issues de sources de données multiples pour les besoins des statistiques européennes. À la demande des autorités statistiques, les administrations, les entreprises et les ménages ainsi que le public en général peuvent être contraints par la loi à permettre l'accès à des données ou à fournir des données pour l'établissement de statistiques européennes.*

Indicateur 2.1: Le mandat des autorités statistiques les habilitant à collecter des informations issues de multiples sources de données et à y accéder pour l'élaboration, la production et la diffusion des statistiques européennes est inscrit dans le droit.

Indicateur 2.2: Les autorités statistiques sont autorisées par la législation à accéder rapidement et gratuitement aux données administratives et à les utiliser à des fins statistiques. **Elles participent, dès le démarrage, à la conception, au développement et à l'abandon des fichiers administratifs, afin, le cas échéant, de rendre leurs données plus adaptées à des finalités statistiques.**

Indicateur 2.3: Les autorités statistiques peuvent rendre obligatoire la réponse aux enquêtes statistiques en se fondant sur un acte juridique.

Indicateur 2.4: L'accès à d'autres données, telles que les données privées, à des fins statistiques est facilité, dans le respect du secret statistique et de la protection des données.

Conditions d'accès aux données privées (2)

Principe 5: Secret statistique et protection des données. *Le respect de la vie privée des fournisseurs de données, la confidentialité des informations qu'ils fournissent, l'utilisation de celles-ci à des fins strictement statistiques et la sécurité des données sont absolument garantis.*

Indicateur 5.1: Le secret statistique est garanti par le droit.

Indicateur 5.2: Le personnel signe un engagement de confidentialité au moment de l'entrée en fonction.

Indicateur 5.3: Des sanctions sont prévues pour toute violation délibérée du secret statistique.

Indicateur 5.4: Des principes directeurs et des instructions sont fournis au personnel concernant la protection du secret statistique tout au long des processus statistiques. Les règles de confidentialité sont portées à la connaissance du public.

Indicateur 5.5: Les mesures réglementaires, administratives, techniques et organisationnelles nécessaires sont en place afin d'assurer la sécurité et l'intégrité des données statistiques et de leur transmission, conformément aux bonnes pratiques, aux normes internationales, ainsi qu'aux législations européenne et nationale.

Indicateur 5.6: Des protocoles stricts s'appliquent aux utilisateurs extérieurs ayant accès aux microdonnées statistiques à des fins de recherche.

Fiabilité, cohérence et comparabilité...

Indicateur 12.3: Les révisions sont régulièrement analysées afin d'améliorer les données collectées, les processus statistiques et les résultats.

Indicateur 14.1: Les statistiques présentent une cohérence interne

Indicateur 14.2: Les statistiques permettent la comparaison sur une période raisonnable.

Indicateur 14.3: Les statistiques sont élaborées sur la base de normes communes en ce qui concerne l'étendue, les définitions, les unités et les nomenclatures dans les différentes enquêtes et sources de données.

Indicateur 14.4: Les statistiques provenant de différentes sources de données et ayant une périodicité différente sont comparées et conciliées.

Peut-on se passer des enquêtes?

Exemple: cadre de vie et sécurité

Graphique 4

Nombres estimés de personnes de 14 ans et plus s'étant déclarées victimes de « Vols ou tentatives de vol personnels avec violences ou menaces » de 2006 à 2014 (En milliers de victimes et intervalles de confiance à 95 %)



Champ : Personnes de 14 ans et plus vivant en ménages ordinaires en France métropolitaine.

Source : Enquêtes « Cadre de vie et sécurité » 2007 à 2015, INSEE-ONDRP.

Note de lecture : Selon les résultats de l'enquête « Cadre de vie et sécurité » 2015, on estime à 178 000 le nombre de personnes de 14 ans et plus ayant déclaré avoir été victimes de vols ou tentatives de vol avec violences ou menaces en 2014.

Tableau synthétique du mois de juillet

	Faits constatés sur les trois derniers mois (CVS-CJO)	
	Nombre	Variation en % ²
Homicides (y compris coups et blessures volontaires suivis de mort) ¹	239	+16
Vols avec armes (armes à feu, armes blanches ou par destination)	1 858	-3
Vols violents sans arme	20 943	+6
Vols sans violence contre des personnes	175 134	+3
Coups et blessures volontaires (sur personnes de 15 ans ou plus)	62 091	+9
Cambriolages de logements	59 615	+2
Vols de véhicules (automobiles ou deux roues motorisés)	35 867	+2
Vols dans les véhicules	64 842	+1
Vols d'accessoires sur véhicules	23 238	+1

¹ Les statistiques et commentaires pour l'indicateur « homicides » portent sur les données brutes. Ces données ont été corrigées sur les années 2015 et 2016 (cf. infra « Méthodes de traitement des données »).

² Cumul des trois derniers mois (mai 2018 à juillet 2018) rapporté au cumul des trois mois précédents (février 2018 à avril 2018), CVS-CJO.

Champ : France métropolitaine.

Source : SSMSI - Base des crimes et délits enregistrés par la police et la gendarmerie.

Complémentarités entre enquêtes et autres sources: chômage (1)

1. Évolution du nombre de demandeurs d'emploi inscrits à Pôle emploi en catégorie A et du nombre de chômeurs au sens du BIT



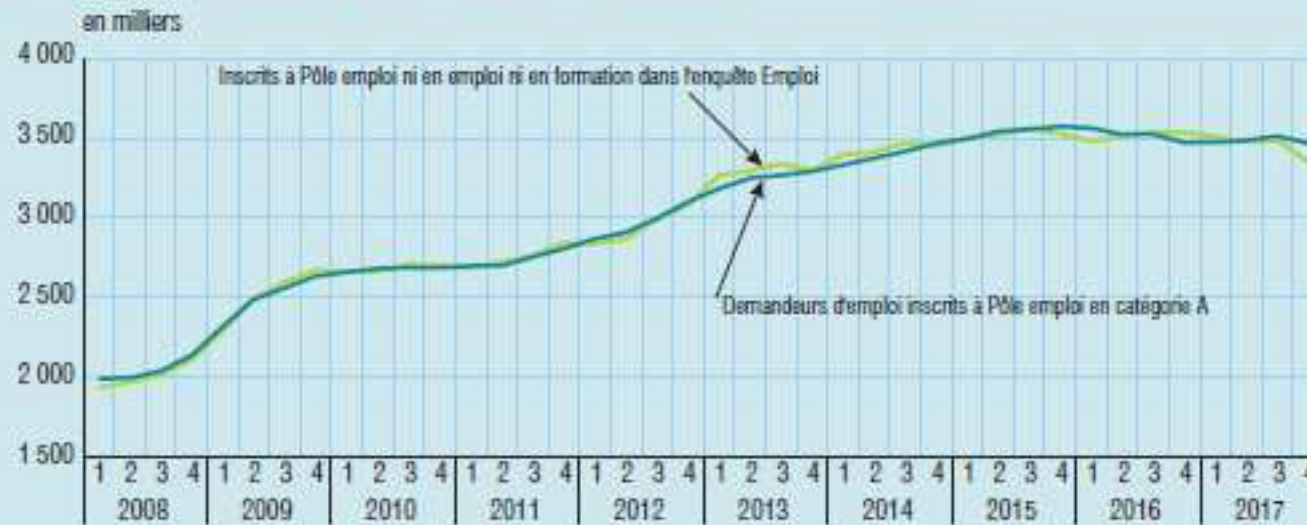
Champ : France métropolitaine.

Note : données trimestrielles CVS (pour le chômage au sens du BIT), CVS-CJO (pour les demandeurs d'emploi).

Sources : Dares-Pôle emploi, fichiers STMT (demandeurs d'emploi) ; insee, enquête Emploi (chômeurs au sens du BIT).

Complémentarités entre enquêtes et autres sources: chômage (2)

2. Évolution de la catégorie A selon l'enquête Emploi et des DEFMA



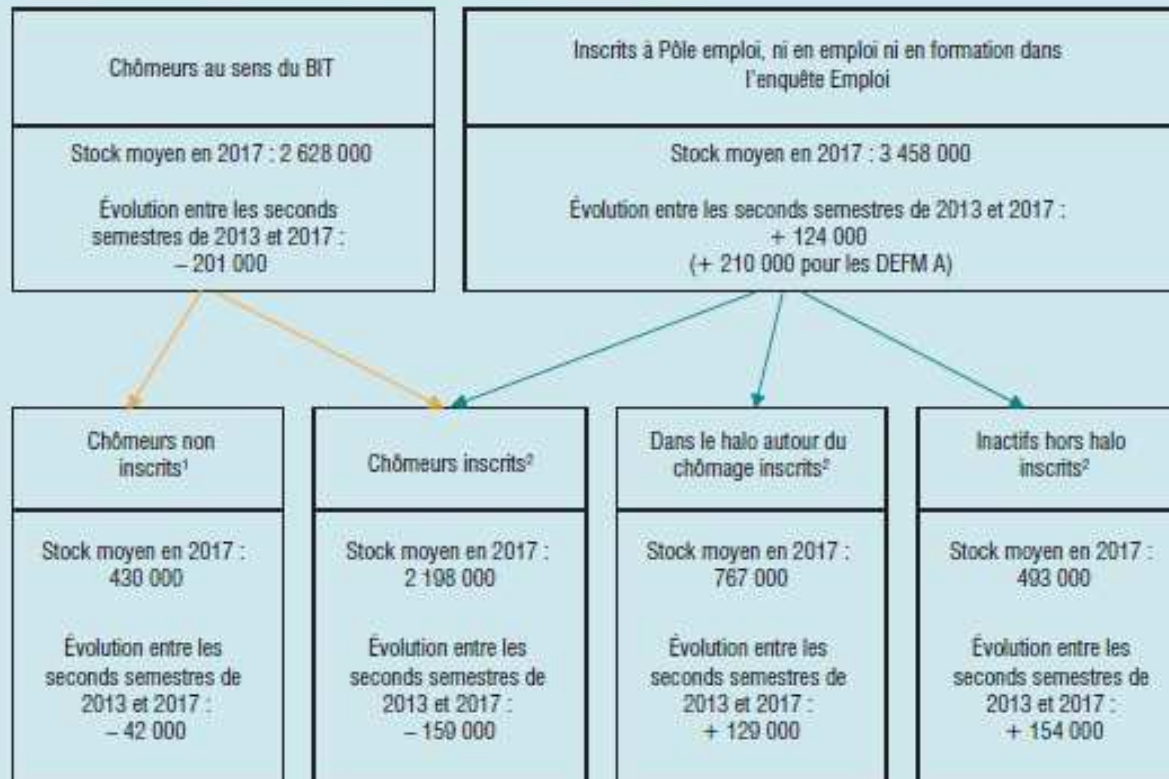
Champ : France métropolitaine.

Note : données trimestrielles CVS pour les inscrits à Pôle emploi sans emploi et hors formation, et CVS-CJO pour les demandeurs d'emploi inscrits à Pôle emploi en catégorie A.

Sources : Dares-Pôle emploi ; Insee, enquête Emploi.

Complémentarités entre enquêtes et autres sources: chômage (3)

3. Répartition des personnes inscrites à Pôle emploi ni en emploi ni en formation dans l'enquête Emploi ou au chômage suivant leur statut au sens du BIT et leur statut déclaratif d'inscription à Pôle emploi



1. Non inscrit au sens «non inscrit à Pôle emploi, ni en emploi ni en formation dans l'enquête Emploi».

2. Inscrit au sens «inscrit à Pôle emploi, ni en emploi ni en formation dans l'enquête Emploi».

Champ : France métropolitaine.

Source : Insee, enquête Emploi.