

## La statistique dans la cité n° 14 - avril 2019

*Lettre bimestrielle du groupe « Statistique et enjeux publics »*

### Sommaire du n°14 – avril 2019

Éditorial :	- <i>L'Europe et la statistique</i>
Droit et statistique :	- <i>Du nouveau sur les données de santé</i>
Méthodes :	- <i>Indice des prix : de nouveau la polémique ?</i> - <i>La nouvelle nomenclature des professions</i>
Vie des institutions :	- <i>Japon : un « bug statistique » ?</i> - <i>Suite du RGPD : la Cnil est active !</i> - <i>L'affaire Andreas Georgiou (suite)</i>
Lu pour vous :	- <i>La question du PIB irlandais</i> - <i>Le secret statistique</i>
Humour :	- <i>L'inflation vénézuélienne</i>
Les Cafés de la statistique :	- <i>Séances passées et à venir</i>

### Éditorial

#### ***L'Europe et la statistique***

En ces temps de Brexit et d'élections au Parlement européen, il est intéressant de se poser la question de la statistique européenne. À quoi sert-elle et comment est-elle fabriquée ?

La statistique publique est considérée comme une activité régaliennne. S'il n'existe pas d'État européen, l'Union européenne a néanmoins acquis de nombreuses prérogatives jusque-là réservées aux États-Nations. Elle établit et adopte des projets de politiques publiques, accorde des aides, définit des normes... Toutes activités qui nécessitent une bonne connaissance des situations économique, démographique, sociale et environnementale dans l'ensemble de l'Union et dans chacun des États-membres, voire dans les régions qui les composent. Une des directions de la Commission, officiellement encore appelée « Office statistique des communautés européennes » (OSCE), mais maintenant connue sous le nom d'Eurostat, est chargée d'établir les statistiques nécessaires au bon gouvernement de l'Union européenne.

Eurostat produit peu de statistiques elle-même. Pour l'essentiel, elle coordonne les instituts nationaux de statistique et les autres autorités nationales de statistique, pour produire des informations cohérentes, harmonisées et comparables au niveau européen. L'établissement de ces statistiques est régi par des règlements et directives auxquels doivent se conformer les États-membres. Mais le domaine d'intervention de l'Union européenne est si vaste qu'aujourd'hui, la plupart des statistiques nationales sont établies dans le cadre de ces normes européennes.

Eurostat ne se contente pas de définir les méthodes de production des statistiques. Elle fixe aussi un cadre plus général sur ce que l'on peut appeler l'environnement de la statistique : organisation des services producteurs, respect du secret statistique, déontologie. Eurostat a en particulier établi un « Code de bonnes pratiques de la statistique européenne » qui fait maintenant référence dans tous les États-membres et même au-delà des frontières de l'Union. Par ailleurs, les instituts nationaux de statistique et les autres autorités nationales responsables de l'élaboration, de la production et de la diffusion des statistiques européennes, s'engagent à respecter le code de bonnes pratiques.

**Pour nous écrire :** [sep@sfds.asso.fr](mailto:sep@sfds.asso.fr)

## **Droit et statistique**

### **Du nouveau sur les données de santé**

**Épisode précédent** : une loi santé de janvier 2016 a créé le Système national des données de santé (SNDS), une immense base de données qui a réuni trois bases existantes : les feuilles de soins de l'assurance maladie, les résumés de sortie des hôpitaux et les causes de décès - avec un « chaînage » permettant un suivi dans le temps de toutes les personnes résidant en France. La responsabilité en a été confiée à la Caisse nationale de l'assurance maladie (Cnamts) ; un Institut national des données de santé a été chargé d'instruire les demandes d'accès des chercheurs et bureaux d'études publics ou privés, la décision finale appartenant cependant à la Cnil, au vu des finalités et du sérieux des projets. On en attendait un fort développement de la recherche, de l'expertise et des innovations en France, en prenant toutes les précautions pour que seules les personnes habilitées dans le cadre de chaque projet autorisé puissent accéder aux données : même si elles ne sont pas directement nominatives, ces bases, très détaillées, comportent des risques de réidentification, avec un fort enjeu de protection de la vie privée.

**Nouvel épisode** : l'article 11 du projet de loi sur l'organisation et la transformation du système de santé, adopté en première lecture à l'Assemblée nationale et qui sera prochainement débattu au Sénat, crée une *Plateforme des données de santé* (se substituant à l'Institut national des données de santé). Le SNDS sera autant que possible étendu à d'autres données de santé recueillies à l'occasion de soins financés par la collectivité (dossiers hospitaliers, dossier médical partagé par exemple). La Plateforme (et non plus la Cnamts) gèrera elle-même la mise à disposition des données du SNDS notamment lorsque ces données seront appariées à d'autres données de santé ; elle pourra facturer ses services aux utilisateurs, aider les gestionnaires de données publics ou privés à les réorganiser pour en faciliter l'utilisation et pour les apparier avec les données actuelles du SNDS. Cette base encore plus riche et techniquement plus facile d'accès permettra, espère-t-on, que la France devienne effectivement une championne des études et recherches utilisant les données de santé.

Cette ambition était déjà celle de la loi précédente mais la Cnil s'était opposée à la constitution de bases de données appariées et pérennes, qui puissent servir à plusieurs projets de recherche ou d'études : la nouvelle formulation de la loi est supposée le permettre.

Pour en savoir plus, voir : <http://www.assemblee-nationale.fr/15/projets/pl1681.asp>

## **Méthodes**

### **Indice des prix : de nouveau la polémique ?**

Connaît-on bien l'évolution du pouvoir d'achat ? Les chiffres que l'Insee publie à ce sujet sont-ils fiables ? Inutile de souligner l'importance de telles questions en ces temps de « gilets jaunes ». Fin 2018, un économiste, Philippe Herlin, a publié un livre intitulé « Pouvoir d'achat : le grand mensonge » dans lequel il attaque l'indice des prix à la consommation de l'Insee (IPC). Le propos est polémique : l'indice est un « mensonge », un « subterfuge », une « tricherie », etc. Ce livre a eu un large écho dans la presse et dans les grands médias. La dénonciation de l'auteur porte essentiellement sur deux points : le poids trop réduit accordé selon lui aux dépenses de logement dans l'indice, et la prise en compte des changements de qualité des produits dans les variations de prix. Les choix de la statistique publique sur ces deux points aboutiraient, selon lui, à minorer la hausse des prix, et donc à surestimer les progrès du pouvoir d'achat des Français.

De telles attaques ne sont pas nouvelles : dans les années 1970, certains syndicats remettaient en cause l'IPC (« Indice Insee, indice truqué ») ; en 2007, Michel-Édouard Leclerc, patron d'une grande surface commerciale, faisait de même. Dans tous les cas, le message est le même : « on ment aux Français en leur faisant croire que leur pouvoir d'achat, en moyenne, progresse ». Ce genre de dénonciation remporte toujours du succès, en flattant une partie du public « dans le sens du poil ».

Les questions soulevées – et notamment l'effet qualité – sont importantes : elles méritent un débat sérieux, qui fasse apparaître les incertitudes qu'elles induisent <sup>(1)</sup>. Malheureusement, l'étude de Philippe Herlin est superficielle, et la violence de ses propos tranche avec la faiblesse de son argumentation technique. Pour ne donner qu'un exemple : changer le poids des dépenses courantes de logement ne modifie que très peu l'évolution de l'indice global. L'auteur n'en dit rien, de même qu'il passe sous silence tout ce qui pourrait contredire sa thèse ! Reste le message négatif, qui a pu impressionner une partie de l'opinion publique. Pourquoi faut-il que les grands médias cèdent toujours

à la tentation de relayer les polémiques ?

À consulter : la [déclaration du directeur général de l'Insee au bureau du Conseil national de l'information statistique du 5 décembre 2018](#) (p. 24)

*(1) Un tel débat existe dans les milieux académiques : c'est d'ailleurs plutôt l'hypothèse inverse, celle d'une surestimation de la hausse des prix par les indices officiels, qui a la faveur de certains économistes. Voir « La croissance est-elle sous estimée ? » dossier paru dans « L'économie française en 2018 », Insee Références édition 2018.*

### **La nouvelle nomenclature des professions**

La nomenclature des professions et catégories socio-professionnelles (PCS) actuellement en vigueur date de 2003. Fin 2017, le Conseil national de l'information statistique (Cnis) a créé un groupe de travail pour rénover la PCS à l'horizon 2020 en lui donnant quatre objectifs :

- 1) actualiser la liste des professions pour prendre en compte l'évolution des métiers, tout en améliorant son articulation avec les nomenclatures internationales (mais sans modifier les catégories et groupes socio-professionnels pour assurer la continuité des comparaisons temporelles),
- 2) améliorer l'utilisation de la PCS, en simplifiant le processus de recueil des professions, afin de les collecter dans la plupart des enquêtes auprès des ménages,
- 3) proposer de nouveaux agrégats et regroupements, pour les individus comme pour les ménages, en complément des agrégats historiques (groupes sociaux, ...),
- 4) améliorer la documentation et la diffusion de la nomenclature, afin de faciliter son appropriation par un large public (chercheurs, journalistes, grand public...).

S'appuyant sur une soixantaine d'auditions et les contributions de quatre sous-groupes, le groupe de travail a formulé les propositions suivantes à la commission Emploi du Cnis du 12 avril dernier.

La P2020 détaillerait 317 professions, au lieu des 486 professions de la PCS 2003 : elles seraient donc moins nombreuses (et plus faciles à coder), de taille plus homogène, avec une distinction plus nette entre secteurs public et privé.

Trois nouveaux agrégats ont été présentés :

- une agrégation des professions prendrait en compte à la fois le statut (indépendant-salarié), une actualisation du niveau de qualification, le secteur (public-privé) et la nature du contrat (durée limitée ou non),
- la PCS Ménage s'appuierait sur un principe de croisement des groupes socio-professionnels des adultes d'un même ménage,
- un autre type d'agrégat permettrait l'analyse de domaines professionnels transversaux aux catégories et groupes existants, par des regroupements ad-hoc de libellés (exemples non exhaustifs : les enseignants, les professions numériques, les métiers « verts », les professions de l'élite).

Parallèlement, une application informatique de saisie des professions en auto-complétion a été développée et testée avec l'aide de l'Insee : dès que l'enquêteur (ou l'enquêté lors d'une collecte par Internet) a saisi trois caractères du libellé d'une profession, une liste apparaît, qui le guide vers le code approprié de la P2020, la liste variant selon la réponse aux questions concernant le statut, le secteur, et le nombre de salariés pour les indépendants.

Enfin, un site Internet dédié à la nomenclature socio-professionnelle a été ébauché : organisé en quatre grandes rubriques ("découvrir" ; "explorer, naviguer" ; "décrire" ; "collecter, coder"), il répondrait (d'un rappel de l'historique à l'aide pratique au codage des professions) à différents besoins et à un large spectre d'utilisateurs.

Suite à la commission du Cnis, la nouvelle nomenclature pourrait être mise en production dès 2020 ou 2021 pour les enquêtes Emploi et Logement de l'Insee et les enquêtes Génération du Céreq...

Pour en savoir plus, consulter le [site du Cnis](#).

### **Vie des institutions**

#### **Japon : un « bug statistique » ?**

Le *Monde de l'Économie* fait état dans son édition du 19 février dernier d'un « scandale ayant ébranlé la confiance dans le système statistique japonais ». La presse anglo-saxonne (*Financial Times*, *New York Times*, ...) s'en était fait l'écho quelques semaines auparavant. Les statistiques tirées de l'enquête mensuelle sur les revenus produites par le ministère du travail de 2004 à 2017 auraient été, d'après ces informations, considérablement surévaluées, ce qui aurait eu des conséquences

notamment pour les services sociaux qui s'appuient sur ces statistiques pour le calcul de diverses allocations : selon une première estimation, près de vingt millions de personnes auraient ainsi été privées d'un total de près de 80 milliards de yens (plus de 630 millions d'euros).

L'opposition au Premier Ministre Shinzo Abe s'est évidemment saisie de cette affaire en se demandant s'il n'y avait pas eu manipulation, ou volonté de dissimulation. Le très sérieux Nikkei Asian Review (hebdomadaire économique) a parlé lui aussi de « scandale ». Ce qui semble malgré tout être à l'origine de cette affaire, c'est l'utilisation de méthodes non conformes aux standards internationaux due à la faiblesse de la coopération entre le service statistique du ministère du travail et la direction générale des normes statistiques chargée de coordonner le système statistique japonais dans un environnement d'un système statistique extrêmement décentralisé mais faiblement coordonné.

Dans cette affaire, *la Statistique dans la Cité* ne peut que rappeler deux des Principes fondamentaux de la statistique officielle adoptés par les Nations-Unies, le principe 8 qui stipule que, « à l'intérieur de chaque pays, il est essentiel que les activités des différents organismes responsables de la statistique soient coordonnées pour assurer la cohérence et l'efficacité du système statistique », et le principe 9 qui rappelle que « l'utilisation, par les organismes responsables de la statistique de chaque pays, des concepts, classifications et méthodes définis à l'échelon international favorise la cohérence et l'efficacité des systèmes statistiques à tous les niveaux officiels ». Il y a probablement eu dans cette affaire un non-respect de certains des principes fondamentaux qui sont depuis 25 ans le pivot des règles éthiques à la base du fonctionnement de nos systèmes statistiques.

### **Suite du RGPD : la Cnil est active !**

Comme nous l'avons déjà indiqué dans « *La statistique dans la cité* » (n°9 – avril 2018), le RGPD européen (Règlement Général sur la Protection des Données) est entré en vigueur le 25 mai 2018. Selon *Le Monde* (21/02/19), une trentaine d'entreprises se sont déjà vues menacées de sanction ou même sanctionnées par la Cnil (Commission nationale de l'informatique et des libertés). Les entreprises concernées n'avaient pas encore bien identifié leurs nouvelles obligations. Mais une prise de conscience est en train de se faire sur la collecte, le traitement, l'exploitation commerciale des données, sur le consentement de l'individu et le maintien de la protection de ses données durant leur conservation. Elle entraîne la création d'une nouvelle fonction, celle d'un délégué à la protection des données (DPO : Data Protection Officer), désormais bien identifié au sein des entreprises et consulté par les salariés avant de lancer un projet.

### **L'affaire Andreas Georgiou (suite)**

« *La Statistique dans la Cité* » a déjà évoqué, dans ses n° 6 (octobre 2017) et 10 (juin 2018), le harcèlement juridique que subit Andreas Georgiou, ancien président d'Elstat, l'institut national de statistique de Grèce, de 2009 à 2014. Le numéro de novembre 2018 de la revue « *Statistique et Société* », mais aussi de nombreux journaux ou revues, ont également consacré des articles ou des dossiers à cette affaire. Parmi les multiples procédures auxquelles Andreas Georgiou a dû faire face suite à la révision en hausse des statistiques de la dette et du déficit publics fin 2009, l'une d'elles concernait la « complicité de fausses déclarations ayant causé de graves dommages à l'État ». À deux reprises, en 2015 et en 2017, un juge d'instruction avait prononcé un non-lieu, décision confirmée les deux fois par la chambre d'instruction de la Cour d'Appel. Après une première annulation de ce verdict par la Cour Suprême (qui joue plus ou moins en Grèce le rôle de la Cour de cassation en France), la Cour Suprême avait à nouveau annulé la décision prise en mai 2017 par la Cour d'Appel. Andreas Georgiou a donc été jugé une troisième fois pour les mêmes charges, nonobstant le principe *non bis in idem* généralement en vigueur dans la plupart des pays, principe qui veut qu'on ne puisse jamais poursuivre pour des faits déjà jugés avec les mêmes motifs. Le 28 février dernier, la chambre d'instruction de la Cour d'Appel a de nouveau prononcé un non-lieu mais cette fois le procureur auprès de la Cour Suprême n'a pas demandé l'annulation de ce troisième verdict dans le délai légal d'un mois dans lequel il aurait pu le faire. La justice grecque reconnaît donc enfin qu'Andreas Georgiou n'est pas complice de fausses déclarations.

Mais ce n'est malheureusement pas encore la fin de l'histoire car de nombreuses autres procédures restent encore pendantes ou sont même encore en phase d'investigations préliminaires. Toutefois, le jugement du 28 février entraîne que les charges de complicité de fausses déclarations qui jouaient un rôle central dans les accusations auxquelles Andreas Georgiou était confronté sont donc définitivement abandonnées ; c'est bien une étape importante dans le déroutement de l'affaire. « *La Statistique dans la Cité* » continuera à informer ses lecteurs des suites de cette affaire qui nous concerne tous.

## **Lu pour vous**

### **La question du PIB irlandais**

Nous avons signalé, dans un précédent numéro de « *La statistique dans la cité* » (n°7 décembre 2017), le problème posé par l'évolution surprenante du PIB irlandais en 2015. Ce sujet est analysé et décortiqué dans deux articles de la revue *Variances*, qui se présente comme « le Webzine des Ensaë alumni » :

#### **- Les multinationales rendent-elles le PIB obsolète ?**

Publié par François Lequiller et François Ecalte ( <http://variances.eu/?p=3923> )

#### **- Économie numérique, globalisation : de nouveaux problèmes pour la mesure de la croissance ?**

Publié par Didier Blanchet ( <http://variances.eu/?p=3877> )

### **Le secret statistique**

Jean-Pierre Le Gléau vient de rédiger un livre consacré au secret statistique : son origine, son évolution, les modalités pratiques de sa mise en œuvre. Il élargit également la réflexion au secret fiscal et à celui qui concerne les données de santé, et à la façon dont ce secret s'applique dans d'autres pays. Cet ouvrage est publié dans la collection « Le monde des données » créée par la SFdS. Il pourra être commandé sur le site de l'éditeur à partir du 2 mai 2019.

<https://aboutique.edpsciences.fr/produit/1074/9782759823420/Le%20secret%20statistique>

## **Humour (noir)**

### **L'inflation vénézuélienne**

Depuis plus de trois ans, la Banque Centrale et l'INS du Venezuela ne publient pratiquement plus de statistiques. L'Assemblée Nationale (présidée par Juan Guaidó) essaie d'estimer l'inflation et a proposé le chiffre de 2 300 000 % en février 2019 comparé à février 2018, après avoir atteint 2 700 000 % en janvier. Malgré ce beau score, le Venezuela est loin de pouvoir postuler au Livre Guinness des records. En Allemagne, en 1923, l'inflation annuelle a été de 16 000 000 %. Plus près de nous, l'inflation au Zimbabwe a atteint un sommet en juillet 2008 avec un rythme annuel de 231 000 000 % (peut-être est-ce là le record mondial). On peut toutefois se demander comment sont établis de tels résultats. À ce degré de magnitude, la méthodologie est sûrement très différente de celle utilisée dans l'Union européenne ; à commencer par les problèmes de collecte : au restaurant à Berlin en 1923, les clients demandaient à payer leur commande en début de repas pour éviter que l'addition n'ait doublé s'ils attendaient la fin du repas. Et à supposer que ce soit commensurable en termes métrologiques, on peut aussi se demander ce que ça signifie conceptuellement.

## **Les Cafés de la statistique**

Les derniers « cafés » se sont tenus :

- le mardi 12 mars 2019 sur le thème « *La transition énergétique. Préparation de la PPE* », avec Richard Lavergne (CGE)
- le mardi 9 avril 2019, sur le thème « *Le décompte des sans domicile* », avec Stéphane Legleye (Insee)

Les suivants se tiendront :

- le 21 mai 2019 sur le thème « *Les data pour le développement durable dans une perspective d'économie numérique* », avec comme invité Lionel Janin (CGDD)
- le 11 Juin 2019 sur le thème « *Le revenu global des familles* », avec comme invité Michel Villac

Les « Cafés de la statistique » ont lieu de 19h00 à 21h30 au Café du Pont Neuf, 14 quai du Louvre, Paris 1er (Métro Pont-Neuf, Châtelet ou Louvre-Rivoli)

**Tous nos lecteurs sont invités à proposer des thèmes qui pourraient être retenus pour de futurs « Cafés de la statistique ».**

*Responsable de l'infolettre : Marion Selz, présidente du groupe SEP*

*Rédacteur en chef : Jean-Pierre Le Gléau*

*Secrétaire de rédaction : Jean-Louis Bodin*

*Webmestre : Érik Zolotoukhine*