

La statistique dans la cité n° 28 - février 2022

Lettre bimestrielle du groupe « Statistique et enjeux publics »

Sommaire du n°28 – février 2022

Éditorial

Outils :

- *Faut-il inclure l'indice des prix des logements occupés par leur propriétaire (LOP) dans l'indice trimestriel des prix à la consommation harmonisé (IPCH) calculé par Eurostat ?*
- *La Nuit nationale de la solidarité*

Vie des institutions :

- *Les services statistiques de deux pays membres de l'OCDE dans la tourmente*

Feuilleton :

- *Coup de théâtre : le HDH retire sa demande d'autorisation à la Cnil*

Courrier des lecteurs

Nous avons lu :

- *Le Courrier des Statistiques n° 7 – 2022 est paru le 20 janvier 2022*
- *L'Ined annonce un bilan démographique de l'année 2020 bouleversé par la pandémie de la Covid-19*

Vie du groupe

Tous les numéros de l'infolettre sont téléchargeables sur le site du groupe SEP ([cliquez ICI](#))

Éditorial

Avec ce premier numéro de l'année 2022, La statistique dans la cité entre dans sa septième année calendaire.

Ce début d'année est aussi à marquer d'une pierre blanche pour le service statistique public français car l'Insee est amené depuis le 1er janvier 2022 et jusqu'au 30 juin 2022, dans le cadre de la présidence française (1) de l'Union européenne, à présider les réunions du groupe statistique du Conseil, chargé de préparer les positions du Conseil en vue de l'adoption, en codécision avec le Parlement européen, des textes législatifs (règlements et directives) proposés par Eurostat pour le compte de la Commission (2). L'Insee assure cette mission avec l'appui des services statistiques ministériels en charge des domaines concernés par ces textes.

Dans [un des articles de son blog](#), l'Insee rappelle l'enjeu majeur de cette législation statistique qui est de pouvoir disposer de statistiques comparables d'un pays à l'autre au sein de l'UE afin d'assurer la meilleure gouvernance économique possible à ses actions, notamment en matière de cohésions sociale, économique et territoriale, de développement durable, de politique agricole ou encore de mondialisation (3). Cette comparabilité suppose que dans un domaine donné, l'ensemble des pays de l'UE travaillent avec les mêmes concepts et les mêmes nomenclatures, et même parfois, dans certains domaines, avec les mêmes outils comme par exemple dans le cas de l'enquête Emploi (*Labour Force Survey*). Cela ne peut se faire sans l'adoption de textes législatifs s'imposant à tous les pays membres. C'est dire l'importance du rôle du pays qui assure la présidence de l'Union pour six mois.

On trouvera dans ce numéro 28 de *La statistique dans la cité* les rubriques habituelles mais il faut signaler un nouveau chapitre du feuilleton HDH (*Health Data Hub*) qui vient de connaître un énième

coup de théâtre. Nous nous réjouissons également qu'un lecteur nous ait écrit et nous souhaitons que vous soyez de plus en plus nombreux à le faire en adressant commentaires, critiques ou observations à sep@sfds.asso.fr.

(1) [Le site « Toute l'Europe » propose une description](#) du rôle, des missions et des responsabilités du pays qui assure la présidence de l'Union européenne.

(2) La procédure d'adoption des textes législatifs européens peut sembler complexe à un observateur extérieur. Elle suppose la collaboration de trois acteurs, la Commission, qui a seule l'initiative de proposer ces textes, le Parlement européen et le Conseil. [Le site « Toute l'Europe » propose une infographie](#) qui décrit la procédure législative en vigueur.

(3) En 2021, plus des deux-tiers des statistiques produites en France par le service statistique public sont couvertes par des textes législatifs européens.

Pour nous écrire : sep@sfds.asso.fr

Outils

Faut-il inclure l'indice des prix des logements occupés par leur propriétaire (LOP) dans l'indice trimestriel des prix à la consommation harmonisé (IPCH) calculé par Eurostat ?

À l'issue de la revue de la stratégie monétaire début juillet 2021, le Conseil des gouverneurs de la Banque centrale européenne a proposé une feuille de route en plusieurs étapes pour l'inclusion, dans l'indice des prix à la consommation, du coût des logements occupés par leurs propriétaires. Cette feuille de route sera discutée lors d'une prochaine réunion du Comité du Système Statistique Européen (CSSE). Si elle était approuvée, Eurostat pourrait publier d'ici 2023 un indice expérimental trimestriel qui intégrerait cet indice dans l'Indice européen des Prix à la Consommation Harmonisé (IPCH). Des changements réglementaires seraient nécessaires par la suite.

En effet, l'IPCH ne prend pas en compte la totalité du coût du logement pour les ménages. Il inclut bien sûr les loyers payés par les locataires (environ 7,5 % du panier de consommation des ménages en zone euro et 7,8 % en France), ainsi que les charges, les frais d'entretien et de réparation, etc. (4,4 % en zone euro et 3,8 % en France), mais pas le coût des logements occupés par leurs propriétaires.

De ce fait, lorsque ce coût évolue différemment des autres prix, sa prise en compte pourrait affecter la mesure de l'inflation. Mais comment mesurer ce coût ? Dans [le n° 253 de son bloc-notes](#), la Banque de France propose plusieurs méthodes.

Mais les économistes sont partagés sur la signification d'une telle prise en compte. Peut-on intégrer des dépenses d'investissement dans un indice des prix à la consommation ? En outre, pour reprendre une expression du monde du journalisme, cette proposition, déjà faite à plusieurs reprises dans le passé, est un véritable « marronnier » !

La Nuit nationale de la solidarité

Plusieurs municipalités, et notamment Paris qui avait lancé la première édition de la *Nuit de la solidarité* en février 2018, ont mis en place des opérations visant à dénombrer les personnes sans abri sur leur territoire, et à recueillir des informations sur leur profil. Le n° 18 (février 2020) de [La statistique dans la cité](#) en avait fait état.

Par ailleurs, l'Insee dans le cadre des opérations de collecte du recensement de la population mène tous les cinq ans le recensement des sans-abri dans les communes comptant 10 000 habitants ou plus. Ce comptage des personnes sans abri a été réalisé les 20 et 21 janvier 2022. Aussi a-t-il paru pertinent de mutualiser les efforts entre ces deux opérations qui poursuivent des objectifs similaires, tout en respectant l'identité de chacune et en les soumettant à des conditions de faisabilité et de réussite assez strictes.

Dans la nuit du 20 au 21 janvier 2022, c'est donc une démarche de la *Nuit nationale de la solidarité*, qui a été initiée par le gouvernement afin de coordonner la réalisation simultanée de l'opération par plusieurs collectivités (4). Celles-ci peuvent s'appuyer sur le [guide méthodologique](#) produit par l'Insee en collaboration avec la Dihal (Délégation Interministérielle à l'Hébergement et à l'Accès au Logement).

(4) De nombreuses grandes villes (dont Lyon, Marseille, Bordeaux et dix villes du Grand Paris) participent à ce dispositif de collecte. Le Monde indique dans son édition datée du 3 février que 2 600 personnes sans abri ont été décomptées à Paris, soit 220 de moins qu'en 2021 et 1061 de moins qu'en 2020.

Vie des institutions

Les services statistiques de deux pays membres de l'OCDE dans la tourmente

Deux pays membres de l'OCDE se sont retrouvés en janvier sur le devant de la scène pour des manquements supposés à l'éthique et aux bonnes pratiques statistiques.

Mardi 25 janvier, le premier ministre du Japon, Fumio Kishida, a regretté les « erreurs » commises dans les demandes budgétaires formulées par plusieurs ministères. Quelques jours auparavant, dix hauts fonctionnaires avaient été sanctionnés pour avoir couvert une surestimation des statistiques mensuelles des mises en chantier dans l'immobilier, l'un des 53 indicateurs-clés du gouvernement pour mesurer l'activité économique. Le quotidien Asahi n'a pas hésité à comparer cette manipulation des chiffres à celle du déficit public grec, qui avait précipité la zone euro dans la crise des dettes souveraines en 2010, d'autant plus que cette révélation n'est pas une première au Japon : fin 2018, le ministère du travail avait reconnu la manipulation de données sur les salaires ; or ces chiffres servaient à calculer les allocations-chômage ou les primes d'assurance versées aux accidentés du travail et près de vingt millions de personnes avaient été lésées, avait admis le ministère. Il faut dire que le système statistique japonais est extrêmement décentralisé avec une très faible coordination, ce qui ne rend pas aisés les contrôles de qualité.

Début janvier, l'Institut de statistique de Turquie (TÜİK - *Türkiye İstatistik Kurumu*) annonçait une augmentation annuelle de l'indice des prix à la consommation en décembre de 36,08 %, la plus haute depuis près de vingt ans. Or, sans doute conscient des dommages causés non seulement à l'économie suite à l'effondrement de la livre turque mais aussi à sa cote de confiance, le président turc, Recep Tayyip Erdoğan avait promis début janvier de « ramener l'inflation à un chiffre le plus vite possible ». Et donc, suivant la logique de cette promesse, il a brutalement limogé le président de TÜİK, Sait Erdal Dinçer, par un décret publié le 29 janvier, sans respecter les procédures en vigueur (5).

Le Japon et la Turquie sont deux pays membres de l'OCDE dont le Conseil a adopté en 2015 une [recommandation \(6\) concernant les bonnes pratiques statistiques](#). Contacté par la rédaction de *La Statistique dans la cité*, le secrétariat de l'OCDE nous a indiqué qu'il examinait quelle suite appropriée il convenait de donner à ces deux situations.

(5) Il faut toutefois signaler que TÜİK n'a pas été contraint de réviser les chiffres publiés pour le mois de décembre et vient en outre d'annoncer une augmentation annuelle de 48,69 % de l'indice des prix à la consommation pour janvier.

(6) Texte amendé en 2019. En 2020, le secrétariat de l'OCDE a préparé un [rapport sur la mise en œuvre de cette recommandation](#), rapport qui a été déclassifié par le Conseil le 24 septembre 2020.

Feuilleton

Coup de théâtre : le HDH retire sa demande d'autorisation à la Cnil

Nouveau rebondissement dans les mésaventures du *Health Data Hub* (HDH (7)) : le ministère des solidarités et de la santé vient de retirer « temporairement » la demande d'approbation par la Cnil du projet d'arrêté que nous évoquions dans le numéro 27 de *La statistique dans la cité* (décembre 2021) et qui aurait autorisé le HDH :

- à récupérer auprès de la Caisse nationale de l'assurance maladie (Cnam) la base du système national des données de santé (SNDS) « historique » (8) ;
- à y ajouter d'autres bases de données de santé (dossiers hospitaliers notamment et bases destinées à la recherche).

La loi exige l'approbation de la Cnil parce que ces données sont pseudonymisées (dans chaque base les données individuelles sont rattachées à un numéro unique obtenu par une « opération cryptographique non réversible » à partir du NIR, le numéro d'inscription au répertoire national d'identification des personnes physiques géré par l'Insee, de la personne concernée) mais qu'il est possible de les ré-identifier si on connaît la personne et son parcours de soins...

Le retrait de la demande du HDH traduit sans doute la crainte du gouvernement d'un refus de la Cnil qui aurait pu se traduire par une crise grave, coalisant les oppositions d'un grand nombre d'acteurs au contrat confiant l'hébergement des données du HDH au « cloud » de Microsoft. Les opposants invoquent les lois autorisant les autorités états-uniennes à accéder à des données personnelles gérées par des entreprises américaines même hors du territoire des États-Unis (9).

Une autre raison des oppositions au HDH est la réticence des gestionnaires de données de santé (assurance maladie et hôpitaux notamment) à une centralisation des bases de données de santé par le HDH, même si la loi de juillet 2019 la prévoit pour les données dont le recueil est financé sur fonds publics. Le Conseil d'administration de la Cnam qui gère aujourd'hui le SNDS s'interroge sur le bien-fondé du financement du HDH à hauteur de 11,5 millions d'euros qui lui est demandé pour cette année (10).

De son côté, le gouvernement français espère que ce conflit peut être l'occasion de faire émerger des « clouds » européens, français notamment, capables de concurrencer les GAFAs tout en respectant la réglementation européenne sur la protection des données personnelles, mais il considère que ce remplacement prendra du temps.

Cela étant, il existe des solutions. En effet, le HDH est supposé exercer deux métiers différents :

- il est d'abord, comme son prédécesseur, l'Institut national des données de santé, un guichet unique pour instruire les demandes d'accès des chercheurs et assimilés aux bases de données de santé, ce qu'il continuera de faire ;
- il est aussi censé être une plateforme hébergeant et appariant ces bases de données ; mais, cette plateforme a à peine démarré (seuls quatre traitements sur les cinquante-cinq annoncés y seraient aujourd'hui effectués). Or rien n'oblige à passer par elle. Le Centre d'Accès Sécurisé aux Données (CASD) héberge déjà [quatre-vingts traitements de données de santé](#) et les hôpitaux peuvent créer leurs propres entrepôts de données de santé conformément au référentiel que la Cnil a justement publié en octobre dernier (11).

(7) La statistique dans la cité *fait régulièrement part des aventures du HDH (Health Data Hub) qui constituent un véritable feuilleton à rebondissements ! Nous en avons parlé pour la première fois dans le n° 18 (février 2020) et le feuilleton s'est poursuivi dans les numéros 21, 22, 23, 26 et 27 !*

(8) Le SNDS historique défini par la loi de janvier 2016 comprend notamment les données exhaustives de l'assurance maladie (actes et médicaments présentés au remboursement...), ainsi que les données d'activité hospitalière (hospitalisations, diagnostics, actes chirurgicaux et thérapeutiques...) et les causes de décès.

(9) C'est pour la même raison que l'équivalent autrichien de la Cnil a interdit le 13 janvier aux sites web autrichiens de vendre les données de navigation des internautes à Google Analytics (qui les traite à des fins publicitaires). Un [communiqué de la Cnil daté du 10 février 2022](#) va dans le même sens.

(10) Voir notamment <https://interhop.org/2022/01/17/cnam-doit-verser-millions>.

(11) Ce référentiel a déjà été évoqué dans le numéro 27 (décembre 2021) de *La statistique dans la cité*.

Courrier des lecteurs

Nous avons reçu la réaction d'un lecteur à propos de la note de lecture publiée dans le numéro 27 (décembre 2021) de *La statistique dans la cité* sur le « trois pages » du Monde du 5 novembre 2021 intitulé « *Dans la fabrique opaque des sondages* ».

Nous encourageons nos lecteurs à réagir aux articles publiés dans *La statistique dans la cité*. N'hésitez pas à nous faire part de vos commentaires et de vos attentes en écrivant à sep@sfds.asso.fr.

Jean-Marc nous écrit :

J'ai été un peu surpris par la note de lecture du dossier du Monde intitulé « Dans la fabrique opaque des sondages » publiée dans le n° 27 de *La Statistique dans la Cité*.

Dans une démocratie, le peuple est censé décider. Avant une décision collective, celui qui peut affirmer « les Français pensent que... » dispose donc d'un énorme pouvoir pour éclairer ou orienter la décision, voire la manipuler. Ce pouvoir est exercé majoritairement par des couples médias / sociétés de sondages ; le fonctionnement même de la démocratie impose de vérifier qu'ils ne déforment pas l'expression de la volonté populaire. J'ai ainsi évidemment apprécié le commentaire d'humeur, dans ce même numéro 27, sur une publication semble-t-il très hasardeuse du *Figaro* reprenant un sondage non identifié.

Je trouve par conséquent totalement légitime de la part d'un journaliste d'investiguer sur la façon dont ce pouvoir est exercé par les sociétés de sondages, et en l'occurrence d'enquêter sur leurs pratiques,

d'autant que celles-ci ont fortement évolué dans le temps. Le même travail resterait à faire sur les pratiques des médias, au moins autant responsables du bon usage des sondages.

Loin de considérer que cette investigation est « contestable sur le plan déontologique », il me semble qu'elle était la seule permettant de connaître et valider ou pas ces pratiques (essentiellement de gestion de panels). Même si vous admettez que « les critiques ne sont pas totalement infondées » (ce qui implique qu'elles le sont partiellement), au lieu de répondre techniquement aux différents problèmes soulevés, vous affirmez que la « quasi-totalité des sociétés de sondage sont sérieuses », ce qui m'amène à trois questions :

- Sur quoi se fonde cette assertion ? S'agit-il d'une simple opinion en forme de plaidoyer pro domo (« il est excessif de condamner les informations fournies par les sociétés de sondages »), ou cela résulte-t-il d'une analyse de la façon dont ces sociétés remédient aux problèmes soulevés par *Le Monde* ? Indépendamment de ceux-ci, le rapport de la Commission des sondages de 2017, faisant suite à la dernière élection présidentielle, mentionnait plusieurs manquements graves à la déontologie (comme le fait de ne pas indiquer les marges d'erreurs correctes ou d'avoir effectué et laissé publier des résultats sortis des urnes avant 20 h) qui interrogeaient en tout cas sur le sérieux de certaines d'entre elles.
- Quelle est la fraction de la population électorale qui n'accède pas à Internet (quelle qu'en soit la raison : idéologie, âge etc.) ? Fonder des sondages exclusivement sur des panels qu'on interroge par Internet ne permet pas de représenter la population qui n'a pas Internet et induit probablement un biais non mesurable et non redressable.
- Des études universitaires comparatives, menées pour la Commission nationale consultative des droits de l'homme, ont montré que, sur des sujets sensibles, il pouvait y avoir vingt points d'écart entre une interrogation en face à face et la même en ligne. Or on est passé de méthodes d'interrogation en face à face d'abord sur un échantillon tiré au hasard, puis constitué par quotas mais représentatifs de l'ensemble de la population, à des échantillons extraits de panels interrogés exclusivement en ligne. Non seulement une partie de la population échappe donc aux sondages actuels, mais le mode d'interrogation influe notablement sur le résultat. Pour affirmer le sérieux des sociétés de sondage, y a-t-il véritablement eu des études montrant qu'il n'y avait pas de biais ?

Nous avons lu ...

Le Courrier des Statistiques n° 7 – 2022 est paru le 20 janvier 2022

Ce [n°7 du Courrier des statistiques](#), paru le 20 janvier 2022, comprend huit articles, consacrés aux problématiques actuelles de la statistique publique.

La revue traite, dans les deux premiers articles de ce numéro, d'une orientation majeure : l'intégration du multimode dans les enquêtes, l'un présentant l'évolution de ce dispositif qui répond notamment à la difficulté de contacter les ménages, l'autre détaillant la construction progressive de l'outillage conceptuel et technique des enquêtes « réellement » multimode, dans toute leur complexité. Un troisième article présente les innovations importantes du recensement agricole de 2020, lesquelles concernent les enquêtés, les enquêteurs et les statisticiens.

Les deux articles suivants portent sur les besoins d'outillage en data science : le SSPCloud, nouvelle infrastructure mise en place par l'Insee, devrait offrir, sur le versant informatique, un environnement « lab » propice à l'engagement d'expérimentations, et le statisticien « selfeur » doit désormais « savoir coder » !

Le sixième article décrit l'élaboration de la nouvelle base statistique créée à partir de multiples sources administratives pour étudier le patrimoine immobilier des ménages et le profil redistributif de la taxe foncière. Le septième article fournit un tour d'horizon des méthodes utilisées pour mesurer les compétences des élèves. Enfin le dernier article relate la mise au point par l'ONU d'une nomenclature internationale des infractions, laquelle a été déclinée fin 2021 au niveau français.

L'Ined annonce un bilan démographique de l'année 2020 bouleversé par la pandémie de la Covid-19

Comme chaque année, l'Ined a publié dans sa revue *Population* (n° 2021/4, volume 76) son analyse des évolutions démographiques en 2020 (12) et recherché l'influence qu'a pu avoir la pandémie de la Covid-19 ; [cette étude peut être librement consultée sur son site](#). En contrepoint, l'Insee, dans son [Insee Focus n° 257](#) (29 décembre 2021) analyse l'évolution du solde naturel de 2013 à 2019.

Les auteurs de l'étude menée par l'Ined observent que la pandémie a bouleversé toutes les composantes de la dynamique du solde démographique naturel : diminution du nombre de naissances (– 17 000), des interruptions volontaires de grossesse (– 10 000) et du nombre de mariages (– 70 000) ; le nombre de décès a été historiquement élevé (+ 56 000). Il faut remonter à la seconde guerre mondiale pour retrouver un nombre de naissances aussi faible. Du fait du prolongement de la crise sanitaire, l'année 2021 sera encore très –voire davantage– affectée. Au 1er janvier 2021, la France comptait 67,4 millions d'habitants soit 120 000 de plus qu'au 1er janvier 2020. Si la crise de la Covid n'a pas entraîné de diminution de la population, elle a très fortement ralenti sa progression, mais aussi les composantes de celle-ci : en 2020, le solde naturel (+ 67 000) a été à peine plus important que le solde migratoire. Le solde naturel est même négatif le dernier trimestre de l'année 2020 et le premier trimestre 2021. La France demeure malgré tout le pays avec la plus forte croissance démographique de l'Union européenne, qui a perdu environ 300 000 habitants en 2020. Si la crise sanitaire a perturbé la dynamique démographique de la France, elle n'a modifié qu'à la marge sa structure par âge, n'entraînant qu'un très léger rajeunissement.

(12) L'Insee avait présenté le 19 janvier 2021 une note rapide sur le bilan démographique de l'année 2020 dans son [Insee Première n° 1834](#). L'étude de l'Ined confirme ce bilan rapide en l'enrichissant de façon très importante, notamment par des comparaisons avec la situation dans d'autres pays européens.

Vie du groupe

Deux Cafés ont été organisés en ce début d'année 2022, le jeudi 20 janvier sur le thème des modélisations de l'épidémie de Covid-19 et leurs utilisations (notre invité était Antoine Flahault, docteur en biomathématiques et médecine, directeur de l'Institut de santé globale, Université de Genève) et le mardi 15 février sur le thème du service statistique public face à la pandémie de Covid-19 (avec deux invités, Jean-Luc Tavernier, directeur général de l'Insee, et Fabrice Lenglard, directeur de la recherche, des études, de l'évaluation et des statistiques au ministère des solidarités et de la santé). Ces deux Cafés ont été organisés en visioconférence. Comme à l'accoutumée, les vidéos des exposés introductifs des deux séances seront mises en ligne sur le site du groupe *Statistique et Enjeux Publics*.

Le Café suivant sera organisé le mardi 8 mars avec Jean Bosvieux qui a été directeur des études de l'Anil, l'Agence nationale pour l'information sur le logement ; le sujet de ce café sera « Comment répondre à la demande de logements au niveau local ? ». En avril, un Café sera organisé le mardi 12 avril autour du commerce de l'armement pour lequel nous avons invité Christian Schmidt, professeur émérite à l'Université de Paris IX-Dauphine.

Responsable de l'infolettre : Chantal Cases, présidente du groupe SEP

Rédacteur en chef : Jean-Louis Bodin

Secrétaire de rédaction : Jean-Pierre Le Gléau

Webmestre : Érik Zolotoukhine