

PRÉSENTATION DU NUMÉRO

L'histoire de la statistique, les outils liés aux enquêtes, les règlements européens, les sources administratives... Depuis sa renaissance en 2018, le Courrier des statistiques vous invite à découvrir le système statistique national au travers d'articles explorant des sujets très divers, faisant ainsi intervenir des auteurs issus des différentes structures de la statistique publique. Ce numéro ne fait pas exception à cette règle, avec quatre articles issus de services statistiques ministériels (SSM) et d'une autre autorité nationale (en anglais, *other national authority* : ONA) qui participent pour la première fois au Courrier des statistiques. Trois articles complètent ce numéro : un article socle sur la statistique structurelle d'entreprises, un article plus institutionnel sur le processus de la revue par les pairs et un dernier article informant sur l'utilisation de données privées, prélude au prochain numéro qui y consacrera un dossier.

La production de toute statistique provient d'un besoin, besoin datant parfois de plusieurs siècles. Il en est ainsi des statistiques du commerce extérieur. L'article de Ketty Attal-Toubert et Boris Guannel revient sur leur genèse et témoigne de leur importance pour la France et à l'échelon européen, hier comme aujourd'hui. En effet, mesurer les importations, les exportations et le solde commercial reste au cœur des débats, et l'intérêt qu'on leur prête a été amplifié par la mondialisation. Récemment, la coopération entre les Autorités nationales statistiques de l'Union européenne en charge du commerce extérieur s'est développée afin de faire face aux enjeux émergeant de la numérisation de l'économie. Des règlements européens se sont imposés pour permettre l'élaboration et la diffusion de statistiques cohérentes entre les États membres.

Dans un autre registre, celui des statistiques de la culture, Amandine Schreiber et Laure Turner nous font découvrir comment depuis plus de 60 ans, les enquêtes comme celle sur les pratiques culturelles des Français permettent d'établir des constats objectifs utiles aux politiques, dans un environnement en perpétuel mouvement du fait des innovations technologiques successives. Par ailleurs, l'enjeu économique est fort et les acteurs sont très divers. Pour suivre ces professionnels et mesurer leur activité, un cadre européen harmonisé a été établi. Mais les particularités du secteur culturel imposent d'avoir recours à des nomenclatures spécifiques.

Le CépiDc (Centre d'épidémiologie sur les causes médicales de décès) de l'Inserm est, avec le CASD (Centre d'accès sécurisé aux données), un des deux ONA à fournir des statistiques officielles, et nous entraîne dans le domaine médical. Pouvoir mesurer précisément les causes de décès est un impératif pour comprendre la situation sanitaire de la France. Cela requiert une codification pertinente de ces causes à partir du certificat de décès renseigné par les médecins. En préambule, Élise Coudin et Aude Robert racontent l'histoire des statistiques sur les causes de décès depuis le XIX^e siècle puis elles expliquent les méthodes de codage utilisées, de la codification manuelle au système de règles puis à l'utilisation de l'Intelligence artificielle avec le codage par prédictions de réseaux de neurones profonds.

Comment mesurer l'insertion professionnelle des étudiants quittant l'enseignement supérieur ? Karl Even nous présente le dispositif InserSup déployé par la sous-direction des systèmes d'information et des études statistiques (Sies) du ministère de l'Enseignement supérieur et de la Recherche fin 2023. Ce dispositif rassemble de nombreuses données administratives exhaustives, émanant de différents ministères. Ces données sont appariées afin de produire et mettre à disposition, à un niveau très fin, une multitude d'indicateurs

pour qualifier l'insertion des jeunes, selon leur formation par exemple. Ce système permet, entre autres, aux étudiants de choisir une formation en pleine connaissance des possibilités d'emploi à l'issue de ce cursus, en se connectant aux différentes plateformes de poursuite d'études.

Après la découverte de quatre milieux très différents les uns des autres, un retour à l'Insee, avec un article sur la statistique structurelle d'entreprises. Dans les années 70, l'Insee a conçu un système statistique très élaboré pour l'observation sur les entreprises. Sylvain Moreau relate cette construction, bâtie sur des données administratives complétées par des enquêtes permettant de construire des bases de données micro-économiques sur les entreprises. Tout en conservant les définitions du début, ce système a cependant considérablement évolué et doit continuer de s'adapter à la complexité des organisations. C'est tout l'enjeu des prochaines années.

Dans ce numéro également, une fois n'est pas coutume, un article institutionnel. Mylène Chaleix et Jean-Michel Durr décrivent le processus de la revue par les pairs. Afin de s'assurer que le Code des bonnes pratiques en matière de statistiques est respecté par les instituts nationaux de statistiques (INS), une équipe constituée de quatre experts va auditionner un système statistique national (INS et ONA). Dans cet article, les trois cycles des revues sont présentés avec leurs recommandations et leurs conséquences pour la France mais aussi dans d'autres pays. La confiance dans les statistiques officielles est un enjeu important pour le système européen. Si les revues par les pairs existent depuis presque 20 ans, leur méthodologie est susceptible d'évoluer lors des prochains cycles.

Enfin, un article sur l'utilisation de données privées à des fins statistiques nous décrit les données bancaires exploitées par l'Insee pour documenter l'évolution de la situation financière des Français lors de l'épisode de la Covid-19 ou encore les impacts d'une remise à la pompe sur la consommation de carburant des ménages. Odran Bonnet et Tristan Loisel expliquent comment l'Insee a accès depuis 2020 à des échantillons anonymisés de comptes bancaires de deux grandes banques : La Banque Postale et Crédit Mutuel Alliance Fédérale. De nombreux défis ont été relevés pour exploiter ces données, qui apportent un regard complémentaire sur l'analyse conjoncturelle économique. Comment résoudre les difficultés inhérentes à l'utilisation de données conçues à des fins de gestion bancaire ? Des premières réponses sont apportées.

L'Insee s'intéresse à d'autres données privées à des fins d'usage statistique : par exemple, les données de carte bancaire ou les données de téléphonie mobile. Mais cela est une autre histoire à découvrir dans un dossier du prochain numéro du Courrier des statistiques. À suivre...

Catherine Fresson-Martinez
Rédactrice en chef, Insee