

Le 22 septembre 2021

Délibéré sur les appariements de données individuelles au sein du service statistique public (SSP)

Après avoir pris connaissance du rapport consacré aux appariements de données individuelles au sein du service statistique public (SSP) par l'Inspection générale de l'Insee et des projets développés par l'Institut en vue de faciliter cette démarche, l'Autorité de la statistique publique souligne la place essentielle qu'est amené à prendre ce « troisième type » de collecte de l'information, à côté des enquêtes et de l'exploitation directe des données administratives, dans la production des statistiques publiques et des études qui en découlent.

Ces opérations permettent en effet une granularité très fine, une rapidité accrue et un enrichissement substantiel des informations obtenues, en vue de répondre à de nouveaux besoins, de se substituer à certaines enquêtes ou de leur servir plus utilement de point d'appui. Elles donnent néanmoins lieu à des contraintes juridiques (anonymisation, règles d'utilisation du numéro d'inscription au répertoire (NIR) et/ou d'identifiants non significatifs...), administratives et informatiques qui rendent souhaitables une mutualisation au sein du SSP.

Face à ces besoins et à la diversité des pratiques des services statistiques ministériels (SSM), l'ASP salue les travaux menés par l'Insee ces dernières années sur l'utilisation de données administratives à des fins statistiques, et en particulier le projet de développement par l'Insee d'un « code statistique non significatif » (CSNS) destiné à faciliter et harmoniser les appariements de données individuelles au sein du SSP, en s'appuyant sur une organisation permettant à la fois d'en étendre et d'en sécuriser l'usage. Elle insiste sur l'intérêt et la nécessité d'accompagner ce projet d'une « offre de services » adaptée, laquelle est indispensable pour permettre au SSP de se saisir de l'ensemble des potentialités de ces techniques de traitement de l'information, en plein essor mais complexes. Elle encourage également l'Insee à poursuivre son examen des expériences internationales et à en tirer les enseignements adaptés.

L'ASP estime en outre que les opérations d'appariement réalisées par le SSP devraient faire l'objet de davantage de transparence et d'une information systématique sur leurs objectifs, leur contenu et leurs conditions de réalisation : mention explicite dans tous les programmes statistiques transmis au Conseil national de l'information statistique (Cnis) et discussion, pour les opérations d'ampleur significative, au sein de ses commissions thématiques ; examen par le Comité du label lorsque les opérations d'appariement réalisées permettent d'enrichir les bases de sondage ou des exploitations afférentes aux enquêtes de la statistique publique.

L'ASP souhaite par ailleurs que les réflexions soient poursuivies sur deux points, en vue de simplifier le contexte actuel des opérations d'appariement, qui demeure complexe et fragmenté :

- la possible convergence, dans le respect de leur cadre juridique respectif, des processus d'appariement mis en œuvre au sein du SSP et de ceux proposés aux chercheurs dans le cadre du Centre d'accès sécurisé aux données (CASD) ;
- le possible rapprochement des techniques d'appariement proposées par l'Insee aux SSM (utilisation du code statistique non signifiant), notamment en vue de l'utilisation des fichiers socio-fiscaux, et de celles mises en place dans le cadre du système national des données de santé (SNDS) par l'entremise de la Caisse nationale d'assurance maladie (CNAM) ou du *Health data Hub*, sachant que des appariements de données individuelles mobilisant simultanément ces deux types de sources sont nécessaires aux SSM comme aux chercheurs pour l'appréhension des inégalités sociales de santé.