
LE PROJET NAUTILE (NOUVELLE APPLICATION UTILISÉE POUR LE TIRAGE DES INDIVIDUS ET DES LOGEMENTS DES ENQUÊTES)

Ludovic VINCENT, Sébastien FAIVRE

Insee, Département des Méthodes Statistiques, Division Sondages

ludovic.vincent@insee.fr

Mots-clés : Échantillonnage, enquêtes ménages, échantillon-maître, sondage en grappes, équilibrage spatialement réparti

Résumé

Actuellement, les enquêtes auprès des ménages de l'INSEE sont tirées, pour la plupart, dans le recensement de la population (RP). La particularité de ces enquêtes réside dans le fait que bon nombre d'entre elles se font en face-à-face et demandent la présence d'un enquêteur. Pour pouvoir réaliser ces enquêtes, l'Insee a mis en place une méthodologie consistant :

- à sélectionner un échantillon de zones de collecte (les « unités primaires ») d'étendue acceptable représentant l'ensemble du territoire français,
- puis, au sein de chacune des zones sélectionnées, et pour chaque enquête ménages, à échantillonner les logements qui seront interrogés.

Cet échantillon de zones (les « zones d'action enquêteur » - ZAE) est appelé l'Échantillon Maître (EM). L'EM actuel a été mis en place en 2009 pour une durée de vie de 10 ans.

Le tirage de l'Enquête Emploi en Continu (EEC) a été effectué dans la taxe d'habitation 2006, donc avant le renouvellement de l'EM de 2009 et sans lien avec ce dernier. Il a consisté à tirer 3 200 zones de 120 logements – les secteurs –, divisés en 6 grappes de 20 logements. Chaque trimestre, une grappe de chaque secteur est enquêtée ; chaque grappe est enquêtée 6 trimestres de suite. La durée de vie de l'échantillon est donc, par construction, de 36 trimestres, soit 9 ans. Entré progressivement en service entre 2009 et 2010, il est impératif de renouveler l'échantillon de l'EEC pour mi 2019.

En 2019, la question du renouvellement de l'EM s'impose donc à l'Insee ainsi que celui de l'échantillon de l'EEC.

La division Sondages a mené des analyses préparatoires montrant un déséquilibre croissant des groupes de rotations du RP au cours du temps, impactant directement la qualité de l'EM actuel et de tout EM issu de la même méthode de tirage.

Par ailleurs, l'apparition d'une nouvelle source, Fidéli (Fichier démographique des logements et des individus) a ouvert de nouvelles opportunités. Fidéli est un fichier d'individus et de logements issu des fichiers fiscaux, apurés, complétés par le répertoire des communautés et celui des résidences hôtelières, et enrichis d'informations de géolocalisation (coordonnées, zonage) et d'informations sur les revenus (issues de Filosofi).

Ce fichier, mis à jour chaque année, présente ainsi les propriétés d'une bonne base de sondage : l'exhaustivité, l'unicité (absence de doublons), et la fraîcheur des données.

L'utilisation de Fidéli permet de s'affranchir des groupes de rotation du RP, et ainsi de diminuer l'étendue des zones d'enquêtes et d'améliorer la qualité des zones sélectionnées (et donc des futures enquêtes). Elle rend également possible le tirage d'individus.

De plus, certaines enquêtes étant déjà sélectionnées au sein des fichiers fiscaux, Fidéli offre la possibilité de tirer toutes les enquêtes dans une unique base, facilitant ainsi la disjonction et évitant la sélection d'un même logement pour plusieurs enquêtes, y compris l'EEC.

Enfin, le système Nautile a également vocation à prendre en charge à terme l'échantillonnage des enquêtes ménages dans les Départements d'Outre-Mer (qui disposent pour l'instant d'un échantillonnage distinct de celui de la Métropole), dans le cadre d'une réflexion méthodologique adaptée permettant de tenir compte des spécificités de la source FIDELI dans ces départements.

Cette communication donne une présentation d'ensemble de la construction des nouveaux échantillons des enquêtes ménages de l'INSEE à partir de la source FIDELI, les aspects méthodologiques les plus innovants du projet (processus de constitution des unités primaires de l'Echantillon-Maître et des grappes Emploi, tirage équilibré réparti spatialement des nouvelles zones de collecte, réflexions sur les possibilités de coordination entre Echantillon-Maître et échantillon Emploi) étant développés dans des communications dédiées.