

D'Octopusse à Nautille ou changement de source pour la base de sondage des enquêtes ménages de l'Insee

**10^e Colloque francophone sur les Sondages
Lyon**

Ludovic VINCENT

**Direction de la Méthodologie et de la
Coordination Statistique Internationale**

Insee - France



Plan de la présentation

L'échantillonnage des enquêtes-ménages à l'Insee

- Le principe de l'échantillon-maître
- Le recensement comme base de sondage

Pourquoi changer ? Pour quoi changer ?

- Changer l'échantillon-maître et revoir la méthode
- Une nouvelle source, Fidéli

Les conséquences du changement

- Un nouvel échantillon-maître
- De nouveaux concepts, de nouvelles variables

L'échantillonnage des enquêtes-ménages à l'Insee

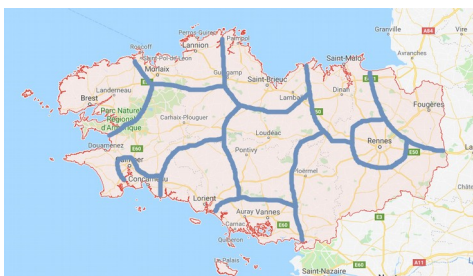
Le principe de l'échantillon-maître

- Les enquêtes ménages de l'Insee :
 - Enquêtes sur les revenus, la santé, les conditions de vie, la mobilité...
 - Souvent en face-à-face
 - Devant représenter tout le territoire
- L'Insee possède un réseau d'enquêteurs sous contrat
 - Mais un nombre limité d'enquêteurs...
 - ...ne pouvant pas arpenter l'ensemble du territoire.

→ Recours à la méthode de l'échantillon-maître (EM)

Le principe de l'échantillon-maître

- D'un point de vue pratique



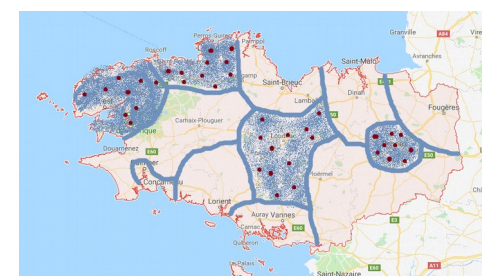
Créer une partition
du territoire



Sélectionner
intelligemment les
zones



Affecter un enquêteur
à chaque zone



Tirer des logements dans
chaque zone, pour chaque
enquête, à partir d'une base de
sondage.

- D'un point de vue théorique : tirage à deux degrés
 - 1^{er} degré : Tirage d'unités primaires (UP), les zones constituant l'échantillon-maître
 - Un échantillon de zones représentant le territoire
 - Des zones qu'un enquêteur peut couvrir
 - 2^e degré : Tirage d'unités secondaires au sein de chaque unité primaire, les logements

Le recensement comme base de sondage

- Depuis les années 60, utilisation des recensements de la population pour la définition de l'échantillon-maître
- Avec le recensement rénové de 2004
 - Le nouveau recensement en deux mots,
 - Répartition initiale en 5 groupes
 - Des communes de moins de 10 000 habitants
 - Des adresses dans les communes de plus de 10 000 habitants
 - Chaque année, une enquête (EAR – enquête annuelle de recensement) sur un groupe telle que :
 - les communes de moins de 10 000 habitants du groupe sont recensées
 - 40 % des adresses du groupe dans les communes de plus de 10 000 habitants sont recensées

Le recensement comme base de sondage

- Avec le recensement rénové de 2004
 - Constitution de l'échantillon-maître
 - Partition du territoire en unités primaires
 - Comme ensemble de communes
 - Chaque UP possède suffisamment de logements pour que l'enquêteur affecté ait du travail chaque année.
 - Chaque UP possède suffisamment de logements pour éviter des réinterrogations sur 10 ans (au moins 300 résidences principales de chaque groupe de rotation dans chaque UP).
 - Tirage équilibré de l'échantillon-maître en 2009
 - équilibré sur les 5 groupes de rotation
 - Au deuxième degré, sélection des logements
 - au sein des unités primaires de l'échantillon-maître
 - parmi les logements recensés dans l'année.

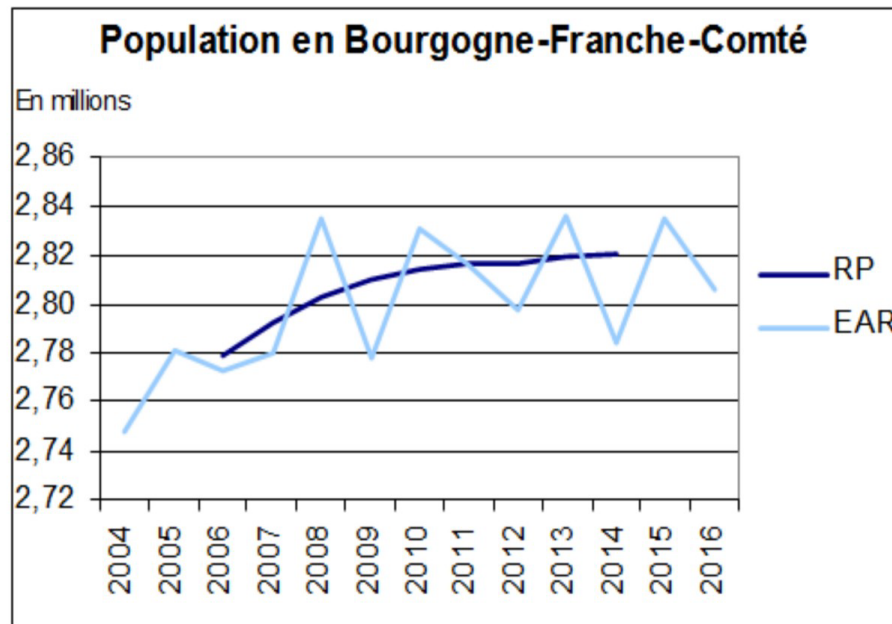
Pourquoi changer ? Pour quoi changer ?

Changer l'échantillon-maître et revoir la méthode

- Un besoin régulier de renouveler l'échantillon-maître
 - Ne pas toujours réinterroger les mêmes logements
 - Épuisement des zones de collecte
 - Des zones qui ne représentent plus le territoire
 - Perte de qualité de l'échantillon-maître

Changer l'échantillon-maître et revoir la méthode

- Une méthode à revoir
 - Déséquilibre des groupes de rotation et donc de tout échantillon-maître issu de la même méthode



Source : Insee

Figure 6 : Évolution des estimations de la population de la Bourgogne-Franche-Comté entre 2004 et 2016, à partir du RP d'une part et de l'EAR d'autre part.

Changer l'EM et revoir la méthode

- Un système trop complexe
 - Répartition en groupes de rotation, sélection d'adresses, partition liée au groupe, tirage équilibré par groupe...
- Deux sources de base de sondage rendant difficile la disjonction des enquêtes
 - Certains tirages sont effectués dans les sources fiscales
 - Difficile de repérer dans le recensement de la population les logements interrogés après tirage dans les sources fiscales

Une nouvelle source, Fidéli

- Constitution d'un nouveau fichier : Fidéli
 - Fichier DÉmographique des Logements et des Individus
- Issu des sources fiscales
 - Taxe foncière : impôt pour le propriétaire d'un logement
 - Taxe d'habitation : impôt pour l'occupant d'un logement
 - Impôts sur le revenu
- Apuré
 - Suppression de doublons
 - Redressement de variables
- Complété
 - Géocodage, ajout des prestations sociales...

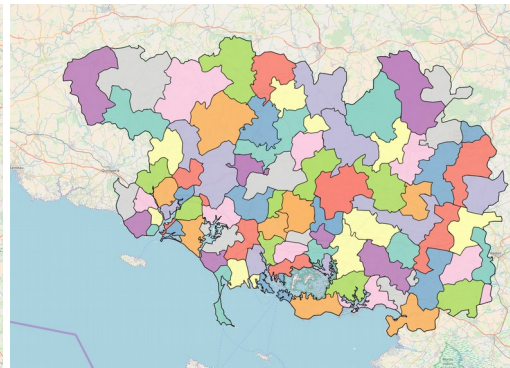
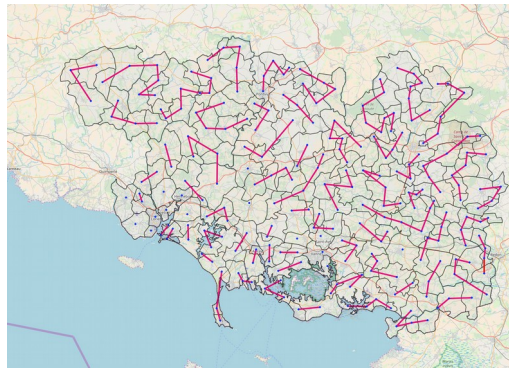
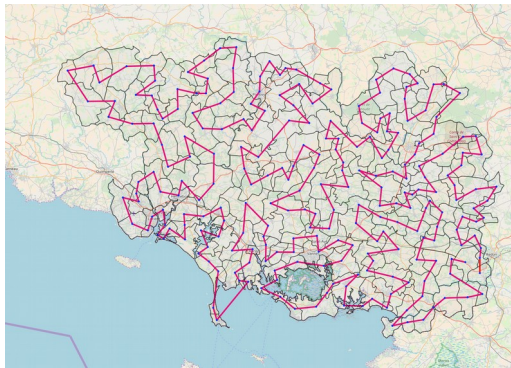
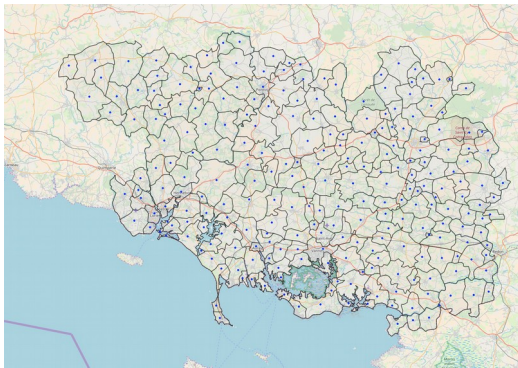
Une nouvelle source, Fidéli

- Les avantages d'une bonne base de sondage
 - Unicité – absence de doublon
 - Fraîcheur
 - Exhaustivité
 - Informations auxiliaires intéressantes
 - Une seule source pour toutes les enquêtes
- Mais règles administratives de gestion et non statistiques
 - Domiciliation, rattachement des enfants aux deux parents...

Les conséquences du changement

Un nouvel échantillon-maître

- Partition du territoire en 5000 zones
 - Ensemble de communes
 - De 2500 résidences principales minimum
 - Les moins étendues possible
 - Grâce à une réponse au problème du voyageur de commerce



Un nouvel échantillon-maître

- L'abandon du recensement de la population permet de s'affranchir de la contrainte sur les cinq groupes de rotation
 - Des unités primaires avec plus de résidences principales
 - donc pouvant durer plus longtemps, sans réinterrogation
 - Des unités primaires moins étendues
 - inutile d'aller chercher une commune de chaque groupe dans chaque unité primaire

Un nouvel échantillon-maître

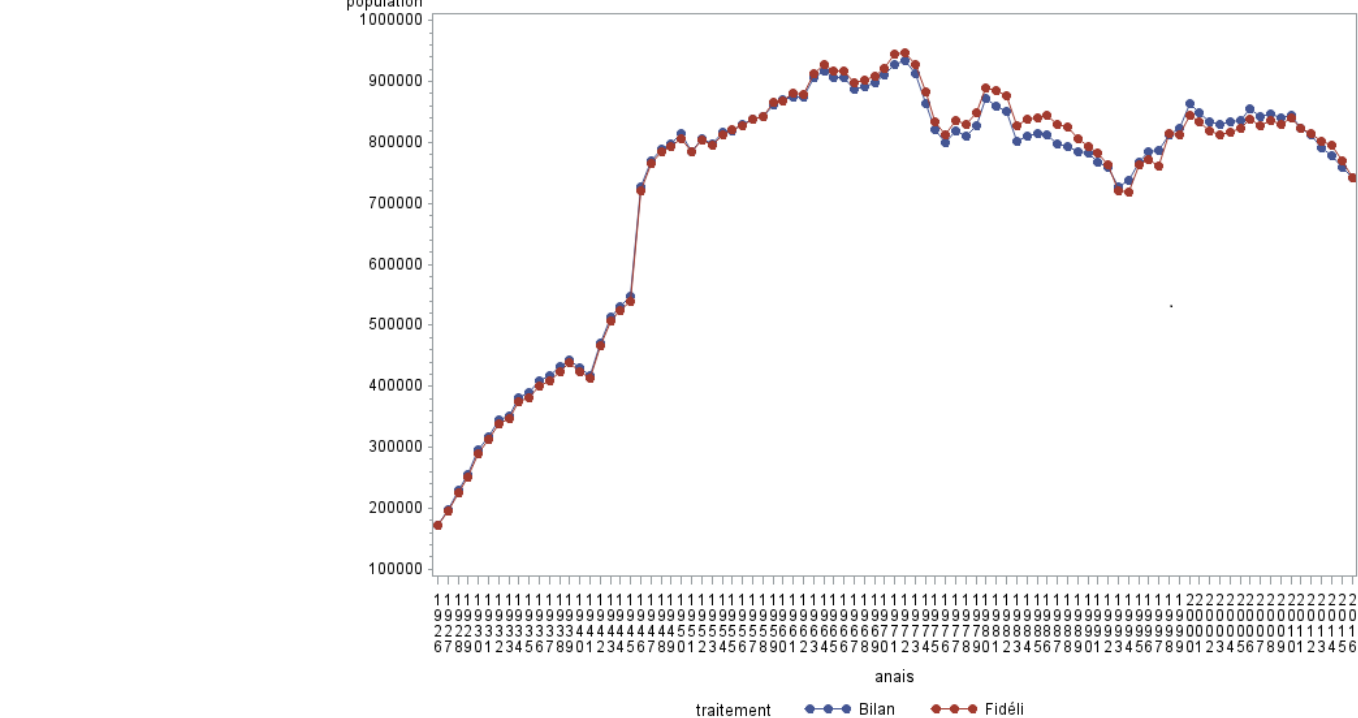
- Sélection de l'échantillon-maître
 - Les objectifs :
 - Obtenir un échantillon-maître représentatif du territoire
 - pour un maximum d'enquête, sur des thèmes différents
 - Pendant 10 ans
 - Respecter les contraintes de précision imposées par Eurostat
 - Sélectionner un nombre de zones que le réseau des enquêteurs peut couvrir

Un nouvel échantillon-maître

- Sélection de l'échantillon-maître
 - La méthode
 - Tirage spatialement équilibré (Grafström et Tillé, 2008)
 - Équilibrage par région
 - Sur des variables *majeures*
 - Sur les premiers axes d'une ACP sur les UP portant sur un grand nombre de variables socio-démographiques
 - Conséquences :
 - L'échantillon-maître n'est pas directement équilibré sur toutes les variables socio-démographiques *souhaitées*
 - Mais est plus représentatif pour un plus grand nombre de variables et d'enquêtes
 - Coordination avec l'échantillon de l'enquête emploi en continu
 - À travers des unités de coordination de 10 000 résidences principales minimum
 - Pour faciliter le travail de collecte

Un nouvel échantillon Maître

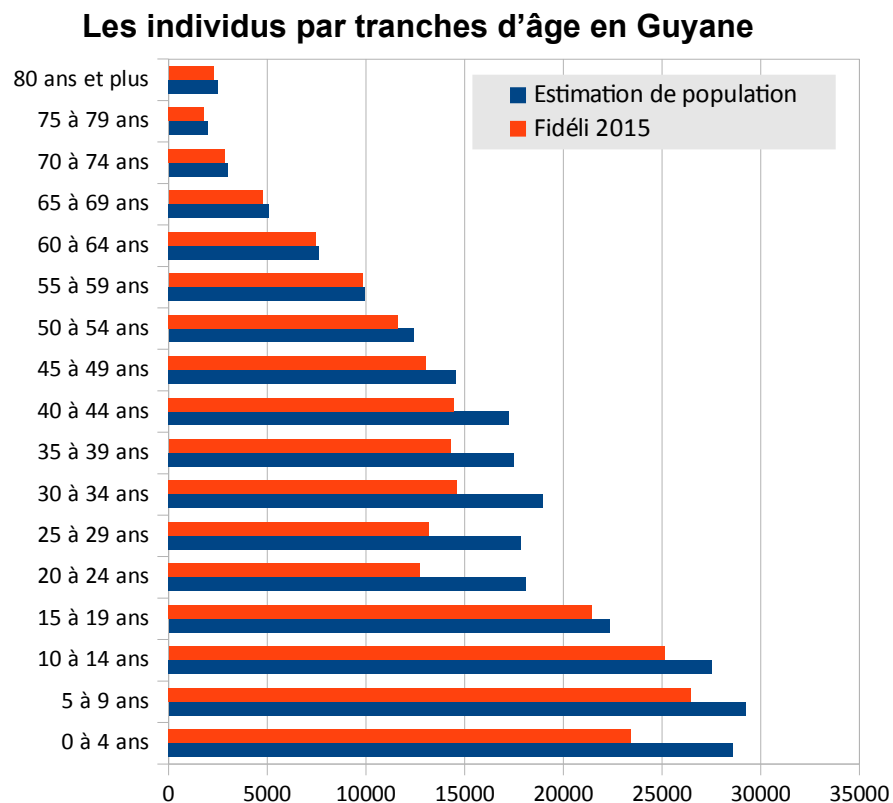
- L'abandon du recensement de la population permet
 - d'utiliser des variables de revenus
 - Très explicatives pour beaucoup d'enquêtes
 - de s'affranchir de contraintes de représentativité sur les cinq groupes de rotation
 - Moins de contraintes, donc plus de précision
 - Ou moins de zones, et donc moins de territoire à couvrir



Champ : Tous les individus 2015

De nouveaux concepts, de nouvelles variables

- Même si localement, on peut trouver des différences
 - Notamment dans les DOM



Champ : Tous les individus 2015

De nouveaux concepts, de nouvelles variables

- Stratification et post-traitements
 - Quelques variables ne sont plus disponibles
 - Diplôme
 - Catégorie socio-professionnelle
 - Mais de nouvelles variables importantes
 - Montant des revenus individuels (et par logements)
 - Types de revenus (salaires, pensions...)

De nouveaux concepts, de nouvelles variables

- Un repérage qui évolue
 - Disparition des informations « recensement »
 - Repérage « normalisé » pour le recensement
 - Mais déjà, des défauts avec le recensement pas internet
 - De nouvelles opportunités d'interrogation et de repérage à étudier
 - Une information géolocalisée et géocodée
 - Présence du mail et de numéros de téléphones des déclarants
 - Informations complémentaires sur l'adresse
 - Dépendance, présence de commerces, communautés...
 - Plusieurs adresses pour un individu

De nouveaux concepts, de nouvelles variables

- Des changements intervenant dans la définition du concept
 - Le logement ordinaire
 - Des logements connus de la taxe d'habitation
 - D'autres logements aux contours plus flous
 - Présence de communautés
 - Foyers fiscaux non rattachés à des structures identifiées
 - Le concept de résidence principale
 - Des intérêts divergents entre bases administratives et enquêtes de recensement

De nouveaux concepts, de nouvelles variables

- Les risques
 - Mauvaise couverture de la base
 - Hors champ et champ non couvert
 - Doubles comptes
 - Un concept différent
 - Impact sur les estimations produites
 - Réflexion sur la définition du logement ordinaire, vis-à-vis du recensement
- Les solutions envisagées
 - Travaux pour rapprocher le concept de résidences principales de celui du recensement
 - Études approfondies des populations mal appréciées dans les bases
 - Enquêtes et tests sur les informations Fidéli

Conclusion

- Une simplification de la méthode par l'abandon du recensement de la population
 - Gain en précision en s'affranchissant des groupes de rotation
 - Utilisation d'une base exhaustive
- Permettant une méthodologie innovante pour la sélection du nouvel échantillon-maître
 - Tirage spatialement équilibré sur variables et axes d'ACP
 - Coordonné avec l'échantillon de l'enquête emploi en continu
- Mais un changement de source à anticiper
 - Modification des variables disponibles pour la stratification et les post-traitements
 - Changement d'informations pour le repérage
 - Modification de l'appréhension du champ des enquêtes ménages du fait de sources divergentes

D'Octopusse à Nautille ***ou changement de source pour la base de sondage des*** ***enquêtes ménages de l'Insee***

Merci de votre attention

Avez-vous des questions ?

Insee
www.insee.fr

Ludovic VINCENT
ludovic.vincent@insee.fr