

Combiner données mobiles et autres sources pour étudier la population présente

Les travaux de l'Insee et leur insertion dans les projets européens

Marie-Pierre Joubert (de Bellefon)

Insee - Direction de la Méthodologie et de la Coordonation Statistique et Internationale - SSP Lab

CNIS - Commission Territoires

6 juin 2024

Pourquoi la Statistique Publique s'intéresse-t-elle aux données de téléphonie mobile?

Mesurer la population présente

- ▶ Avis de **moyen terme du CNIS 2019 - 2023** : "le Conseil souhaite que les travaux d'expertise des **nouvelles sources de données** se poursuivent dans l'objectif de mesurer la **population présente sur un territoire à différents moments de la journée, différents jours de la semaine ou mois de l'année.**"
- ▶ Avis renouvelé par le CNIS pour le **moyen terme 2024 - 2028** avec le constat que "la seule mesure de la population résidente ne suffit pas pour capter la **dynamique** et l'**attractivité** d'un territoire, la **fréquentation de ses équipements** et l'utilisation de ses ressources" et un souhait que "**l'ensemble des données publiques et privées** soient mobilisées pour mesurer la population présente".
- ▶ Besoin exprimé par les acteurs locaux pour **calibrer** les offres de **soin**, traitement des **déchets**, services de **secours**,...

Mesurer la mobilité quotidienne et de loisir

- ▶ Mieux comprendre les **interactions entre territoires** et les **comportements quotidiens** des individus, vers leur lieu de travail et vers les centres de loisir.
- ▶ Analyser la fréquentation quotidienne des **Aires d'Attraction des Villes** et des **Bassins de Vie** : zonages d'étude construits avec les données usuelles de la statistique publique.

Depuis 2016 : des travaux expérimentaux et un éclairage des mouvements de population durant le COVID

Groupe de travail des instituts de statistique européens

- ▶ ESSNet Big Data 1 et 2. Mené par **Eurostat**.
- ▶ Partage d'expérience (partenariats existant en Belgique, aux Pays Bas, en Italie,...)
- ▶ **Harmonisation des méthodologies** : rapports et outils en accès libre



EUROPEAN
STATISTICAL
SYSTEM

2016 - 2022 : partenariat de recherche avec Orange et Eurostat

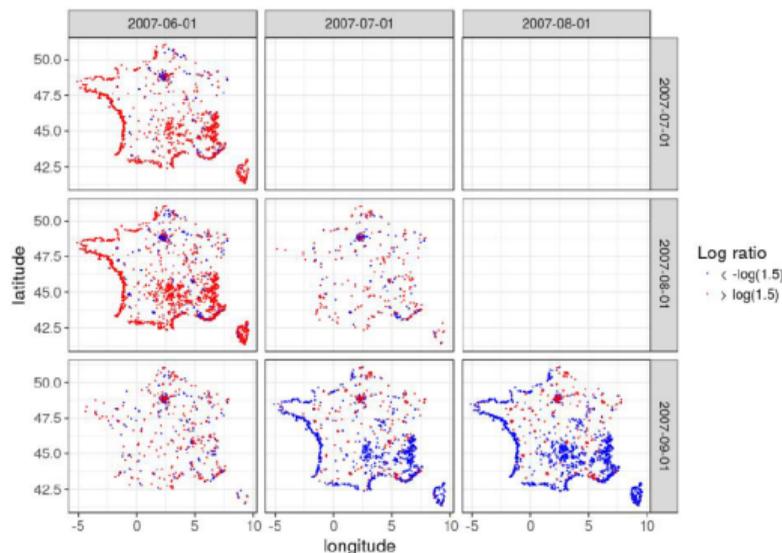
Les acteurs

- ▶ Laboratoire de **recherche en sociologie** d'Orange (Orange Lab - SENSE)
- ▶ **Eurostat** - Task Force Big Data
- ▶ **Insee** - Méthodologie et Coordination Statistique et Internationale

Résultats

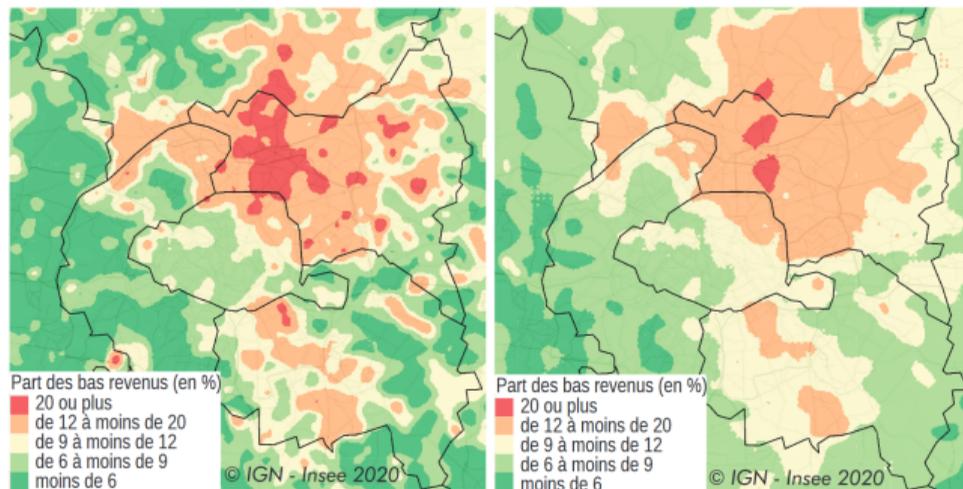
- ▶ plusieurs **articles de recherche** qui éclairent des sujets d'**intérêt pour la statistique publique**
- ▶ une meilleure **compréhension des données** ⇒ bonne base pour de futures **collaborations** avec Orange

Variation de la population présente par mois



Lecture : Entre juin et août la population détectée comme habitant autour des antennes rouges a plus que doublé (essentiellement sur le littoral et en montagne). Source : Orange, CDR 2007; calculs Insee - Bellefon, Givord, Sakarovitch, Vanhoof 2019

La ségrégation résidentielle en combinant données mobiles et données fiscales



Part des personnes à bas
revenus à **6h**

Part des personnes à bas
revenus à **16h**

La mixité sociale est plus forte en journée sur les lieux d'activité que pendant la nuit dans les quartiers de résidence Source : Orange, CDR 2007, Filosofi 2014 ; Calculs Insee - *Galiana, Sakarovitch, Sémécurbe, Smoreda - 2020*

L'expérience de mesure de population lors du Covid

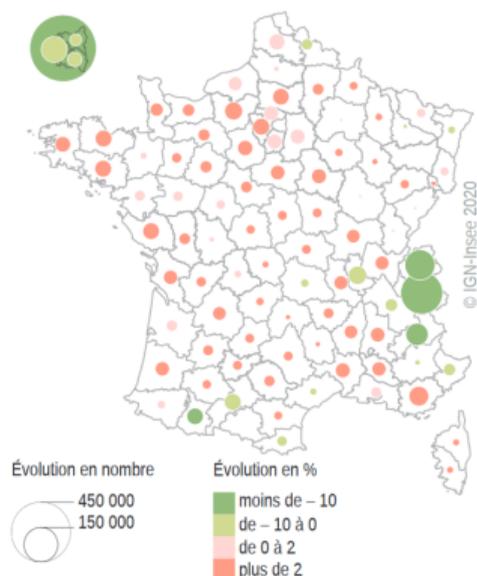
Besoin de statistiques objectives et fiables :

- ▶ Combien de personnes se sont déplacées?
- ▶ Quel profil des personnes s'étant déplacées?

Données

- ▶ Données commerciales mises à disposition gratuitement par Orange, Bouygues, SFR
- ▶ Données journalières de présence au niveau départemental, conditionnées par le département de résidence

Evolution du nombre de métropolitains présents dans le département



During le confinement, Paris a perdu 20% de ses habitants Source : Bouygues Telecom, Orange, SFR (activations du réseau de téléphonie mobile); calculs Insee - *Galiana, Suarez-Castillo, Sémécurbe, Coudin, Bellefon* - Insee Analyse Juillet 2020

Valorisation

Communiqués de presse et Insee Analyse

- ▶ Premiers résultats le 8 avril 2020
- ▶ Résultats consolidés le 18 mai 2020
- ▶ **Retour partiel des mouvements de population avec le déconfinement** *Galiana, Suarez-Castillo, Sémécurbe, Coudin, Bellefon - Insee Analyse Juillet 2020*

Retour d'expérience

- ▶ **Que peut faire l'Insee à partir des données de téléphonie mobile? Mesure de population présente en temps de confinement et statistiques expérimentales** - *Semecurbe, Suarez Castillo, Galiana Coudin, Poulhes - blog de l'Insee, 2020*
- ▶ **The French official statistics strategy : Combining signaling data from various mobile network operators for documenting COVID-19 crisis effects on population movements and economic outlook** - *Coudin, Poulhes, Suarez Castillo - Cambridge University Press 2021*

Actuellement : deux collaborations en cours avec des acteurs académiques et privés

Le projet MOBITIC - subventionné par l'Agence Nationale de la Recherche 2022-2025

Estimation des indicateurs de mobilité et de présence à l'aide des TIC

- ▶ Collaboration entre l'**Insee**, l'opérateur de téléphonie mobile **Orange** (Orange Business Service et Orange Lab) et l'**Université Gustave Eiffel** (laboratoire GRETTIA et LICIT)
- ▶ Combinaison de différents jeux de données, incluant la **téléphonie mobile** agrégée à une échelle spatiale fine, les comptages de flux de **trafic routier**, les données de **transport en commun** et les données **socio-démographiques détaillées** collectées par la statistique publique.



La chaire finance digitale

Utiliser les données de transaction par carte bancaire pour améliorer l'analyse et la prédiction de l'activité économique

- ▶ Fournisseur de données : **Groupement des Cartes Bancaires CB** (GIE CB) : groupement d'intérêt privé qui regroupe la majorité des institutions financières françaises, dans le but de **faciliter l'inter-bancarité** des paiements par carte bancaire.
- ▶ données : données de transaction par carte CB anonymisées.
- ▶ Supervision académique : **Telecom Paris** et l'**Université Paris 2 Panthéon Assas**.
- ▶ Principaux partenaires : **Caisse des Dépôts, Insee**.



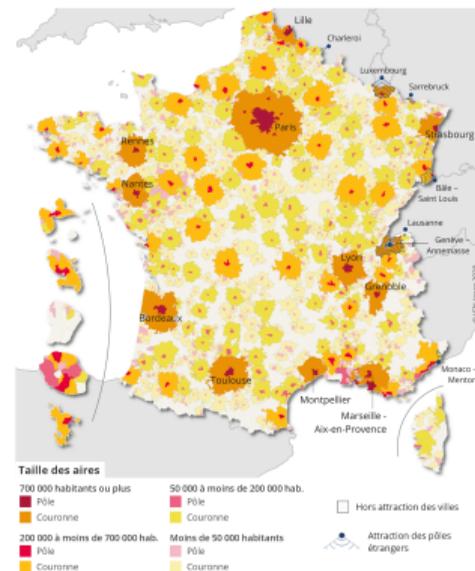
Aires d'attraction des villes : aller au delà des déplacements domicile-travail

Méthodologie

Les centres **densément peuplés** et leurs aires d'attraction, dont au moins **15 % des actifs occupés** travaillent dans le centre d'après le recensement de la population.

Contribution des nouvelles sources

- ▶ comment l'augmentation du **télétravail** fait évoluer la forme des aires en fonction du jour de la semaine ?
- ▶ les personnes **hors du champ des actifs occupés** (retraités, étudiants,...) se déplacent-ils régulièrement vers le centre ?
- ▶ comment les mobilités impactent-elles les **comportements de consommation** ?



Bassins de vie : prendre en compte la fréquentation des équipements

Méthodologie

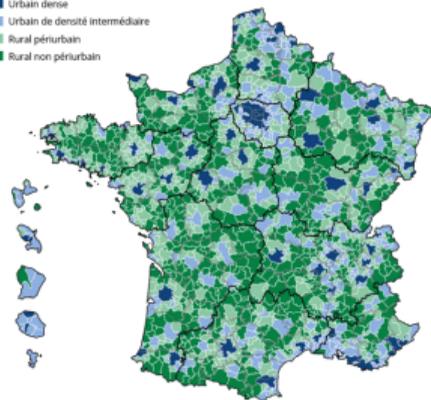
Les bassins de vie sont les plus petits territoires au sein desquels les habitants ont **accès aux équipements et services les plus courants** (commerces, santé, services aux particuliers, sports, loisirs, enseignement...)

Apport des nouvelles sources

- ▶ Les résidents utilisent-ils toujours l'équipement **le plus proche**?
- ▶ Quelle est la **fréquence d'utilisation** de chaque équipement?
- ▶ En quoi la présence d'un grand **centre commercial en périphérie** affecte la **fréquentation des commerces de centre ville**? Cela dépend-il de l'importance du **télétravail**?

type de bassins de vie

- Urbain dense
- Urbain de densité intermédiaire
- Rural périurbain
- Rural non périurbain

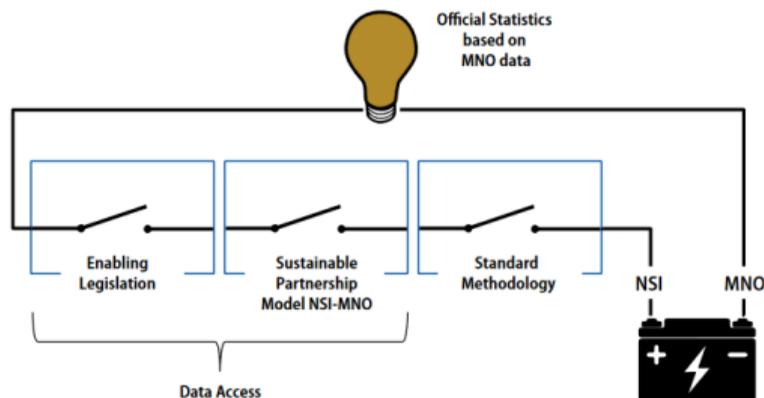


Avec Eurostat : valoriser, échanger et préparer la suite

Deux projets méthodologiques

1 - Task Force pour une méthodologie commune au Système Statistique Européen

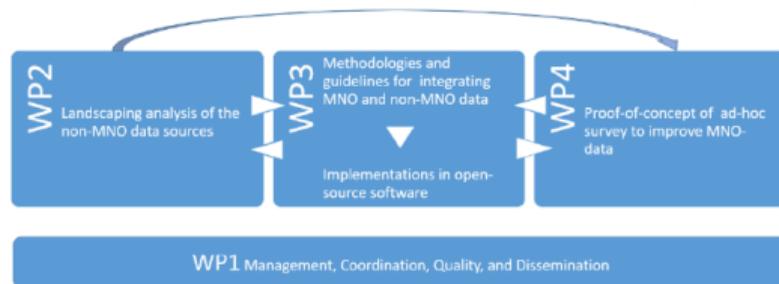
- ▶ **Statistical Reports series** publié en juillet 2023, co-écrit Eurostat - Insee et 17 autres instituts de statistique publique.
- ▶ Standards à respecter : ouverture, transparence, modularité, multi-opérateurs, fusion mobile et non-mobile.



Deux projets méthodologiques

2 - Trusted Smart Statistics : methodological developments based on new data sources

- ▶ **ESS-Net sur 2 ans** Début en novembre 2023. 10 Instituts de Statistique (INS) participants.
- ▶ **Combiner données mobiles et autres sources de données** : panorama des sources possibles, enjeux méthodologiques et réflexion critique. **Insee pilote des travaux sur le panorama des sources à combiner aux données mobiles (WP2).**
- ▶ **Sprint en juin 2024** : deux jours à Vienne pour travailler ensemble au niveau européen.



Un projet technique

Développement d'un pipeline pour intégrer les données de multiples opérateurs

- ▶ Subventionné par Eurostat. Mené par GOPA (cabinet de conseil). Inclut :
 - ▶ 2 entreprises spécialistes de l'analyse des données mobiles
 - ▶ 2 Instituts de Statistique (Italie et Pays-Bas)
 - ▶ 5 opérateurs téléphoniques de 4 pays (Vodafone Italie, Vodafone Espagne, Orange Espagne, A1 Slovénie, POST Luxembourg)
- ▶ **Insee et Orange France** participent au **comité de suivi** (*Advisory Board*)

Conclusion

- ▶ Des données qui peuvent apporter un éclairage des **problématiques locales** très **complémentaire** des données produites par la statistique publique.
- ▶ Un **travail méthodologique** et une bonne **collaboration avec les opérateurs** sont nécessaires car ces données ne sont **pas conçues initialement pour la statistique publique**.
- ▶ Perspective à suivre grâce aux nombreux **projets français et européens** en cours actuellement.

Merci de votre attention

Marie-Pierre Joubert (de Bellefon)
Insee - Responsable adjointe du SSP Lab
marie-pierre.joubert@insee.fr