

La base permanente des équipements (BPE)

Une source statistique singulière et constamment en mouvement



Xavier Helfenstein *

Singularité française, la statistique publique produit chaque année une base permanente des équipements – la BPE – qui recense et géolocalise les équipements, services et infrastructures accessibles à la population sur l'ensemble de la France.

Son intérêt principal est de mettre à disposition de tous un outil simple qui fournit des informations sur le niveau d'équipement, services et infrastructures d'un territoire (présence d'équipements scolaires, médicaux ou encore sportifs, etc.).

Ingrédients clefs de nombreux diagnostics ou analyses de territoire, ses produits de diffusion sont multiples et largement utilisés à l'intérieur comme en dehors de la statistique publique. Héritière de l'inventaire communal, la constitution de cette base s'appuie aujourd'hui sur la collecte de sources, principalement administratives. Bien moins onéreuse à produire et à mettre à jour, elle permet à l'Insee d'entretenir des relations fructueuses tant avec ses utilisateurs que ses fournisseurs de données. L'objectif désormais est de mieux prendre en compte encore les demandes des utilisateurs et de faciliter les échanges à la fois avec les producteurs de données et les utilisateurs, en gérant un système complet de métadonnées.

 *A unique feature of French official statistics is the annual production of a permanent equipment database – the BPE – which lists and geo-locates the equipments, services and infrastructures accessible to the population throughout France.*

Its main interest is to make available to everyone a simple tool that provides information on the level of equipment, services and infrastructure of a territory (presence of school, medical or sports facilities, etc.).

As a key ingredient in many territorial analyses, its dissemination products are numerous and widely used both inside and outside official statistics. Heir to the "communal inventory" survey, the constitution of this database is now based on a collection of sources, mainly administrative. Cheaper to produce and update, it enables INSEE to maintain fruitful relations with both its users and its data suppliers.

The main objective now is to rationalise the work, to take better account of users' requests and to facilitate exchanges with both data producers and users by managing a complete metadata system.

* Chef du pôle Base permanente des équipements, Insee Nouvelle Aquitaine, xavier.helfenstein@insee.fr

Je viens d’emménager, quels sont les équipements disponibles dans mon village ? Y a-t-il un lycée à proximité de la commune où je suis muté ? Aurais-je accès facilement à un centre de dialyse autour de chez moi ? Y a-t-il un terrain de tennis pas trop éloigné ? Tout un chacun peut être amené à se poser ces questions et bien d’autres sur les équipements accessibles et leur localisation.

Les autorités publiques, qu’elles soient nationales ou locales, s’intéressent elles aussi tout naturellement à l’accès de la population aux équipements : quelles sont les zones rurales dans lesquelles les équipements de la vie courante sont difficilement accessibles ? Le vieillissement de la population accentue-t-il l’inégalité d’accès aux équipements ? L’offre de places en crèche est-elle suffisante ? L’équilibre entre habitat et équipement est-il assuré ? (Hautdidier et Kuentz, 2011)

L’intérêt de **la base permanente des équipements (BPE)** est de mettre à disposition un outil simple permettant de répondre à ces différentes questions. Plus précisément, la BPE fournit des informations quantitatives sur le niveau d’équipement d’un territoire. Ce faisant, elle permet de produire les données souhaitées, comme la présence ou l’absence d’un équipement, la densité d’un type d’équipement, etc. (Barbier, Toutin, Levy, 2016), et plus globalement d’analyser la structuration du territoire et notamment d’identifier des zones sous-équipées.

L’Insee utilise d’ailleurs la BPE pour déterminer le zonage des bassins de vie. Pour connaître les zones d’attraction des équipements en milieu rural, des distances entre les communes et les équipements peuvent être calculées. Les données géolocalisées de la BPE conduisent à des analyses territoriales à un niveau très fin. Elles sont souvent associées à d’autres sources, comme le recensement de la population et le dispositif sur les revenus localisés sociaux et fiscaux (Filosofi) pour s’enrichir d’informations relatives à l’offre d’équipements, à l’attractivité des territoires, etc.

► La BPE, une exception française

Au niveau européen, la BPE fait figure d’exception.

Au niveau européen, la BPE fait figure d’exception. Si certains instituts nationaux de statistique mettent à disposition des données géolocalisées sur les équipements, il s’agit le plus souvent de fichiers organisés par thématique¹. Quand la Commission européenne² cherche à mesurer l’accessibilité des services, elle ne dispose pas de données fines comparables sur l’ensemble des États membres³, car la notion d’équipement n’est pas tout à fait cohérente avec le concept qui supporte les implantations des unités légales des répertoires d’entreprises. Elle n’est pas non plus totalement cohérente avec celui des fichiers de bâtis. Dans cet entre-deux, la France a choisi d’insérer un système d’information spécifique, la BPE : il faut sûrement y voir la spécificité française de s’intéresser de longue date aux disparités territoriales.

¹ Voir par exemple les fichiers de l’ONS suisse sur la localisation des commerces, sur le site www.data.europa.eu.

² En l’occurrence la direction générale de la Politique régionale et urbaine appelée DG Regio.

³ Voir l’introduction de l’article de (Kompil et alii, 2018) dont les auteurs utilisent par ailleurs des données privées pour valider leurs hypothèses et leur modèle.

► De quels équipements parle-t-on ?



Pour la BPE, un équipement est un service accessible à la population, qu'il soit marchand ou non.



Définir un équipement n'est pas chose facile : pour la BPE, un équipement est un service accessible à la population, qu'il soit marchand ou non. Le terme de « service » est à prendre dans une acception très large, car il recouvre une grande variété de cas. Il peut s'agir d'un aménagement, tel qu'un jardin remarquable ouvert au public, un lieu de baignade aménagé, une boucle de randonnée. L'équipement peut également être une infrastructure : un gymnase, une piscine, une gare, un établissement de transfusion sanguine ou encore un centre d'accueil

de demandeurs d'asile. Les établissements ouverts au public et offrant des services, comme les commerces, les banques, les tribunaux, les écoles, etc., font aussi partie du champ. Plus exactement, la notion d'équipement se réfère au service rendu au sein d'un établissement et non à l'établissement lui-même. Ce sont ainsi la maternité et le service des urgences qui sont répertoriés et non le centre hospitalier auquel ces services appartiennent. Toutefois, dans la majorité des cas, le service et l'établissement se confondent.

La base permanente des équipements (BPE) répertorie ces équipements au 1^{er} janvier de chaque année. Ils sont classés par type, au nombre de 188 en 2021 (**encadré 1**) et regroupés en sept domaines : les services aux particuliers, les commerces, l'enseignement, la santé et le social, les transports et déplacements, les sports-loisirs et la culture, le tourisme.

Comme souvent en production statistique, le dispositif actuel est le fruit d'une lente évolution sur une longue période.

► Aux origines de la BPE : l'inventaire communal

Historiquement, la BPE succède à l'inventaire communal qui avait lieu tous les 8 à 10 ans. Le dernier inventaire a été réalisé en 1998. Il s'agissait d'une enquête auprès des communes de moins de 30 000 habitants pour la France entière, y compris les DOM, à l'exception de certains départements d'Île-de-France. Une commission communale réalisait l'inventaire des équipements, commerces et services présents et, pour ceux qui étaient absents, les services de remplacement existants (tournée, permanence, rayon dans un commerce multi-services, service identique dans une commune proche, etc.).

Ce dispositif permettait ainsi de disposer de résultats communaux sur la densité d'implantation, la fréquentation des équipements et l'attractivité des communes. Cet outil avait l'avantage d'être une source homogène et facilement mobilisable. Les utilisations locales ou régionales de l'inventaire communal consistaient à décrire le degré d'équipement d'une commune, à le comparer à des communes proches ou de même importance ou encore à dresser la carte d'implantation d'un type d'équipement sur une zone. Les inventaires communaux étaient ainsi mobilisés lors de la mise au point des schémas de services publics en milieu rural, à la demande des collectivités locales ou des régions, ou pour définir des zonages tels les bassins de vie et les pays⁴. Les données disponibles permettaient

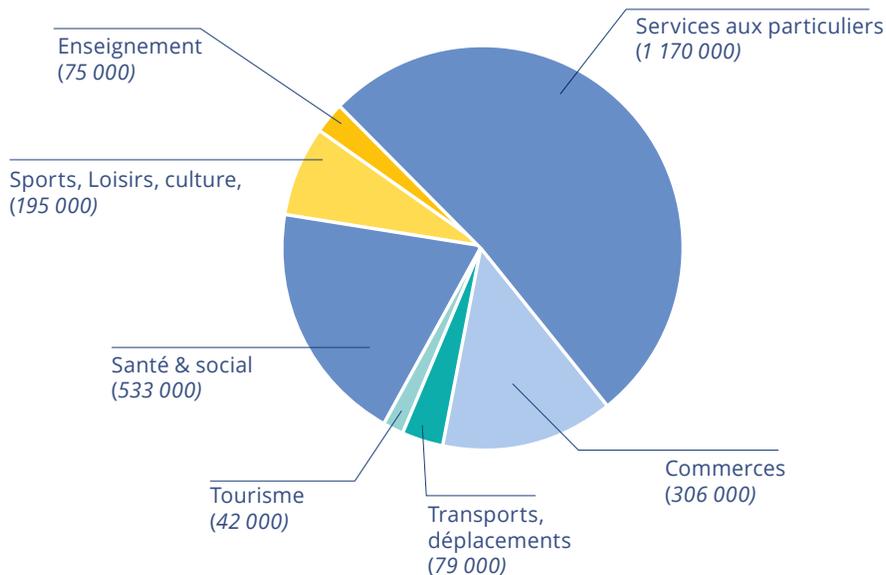
⁴ Le « pays » est une catégorie administrative datant de 1995 ; ce territoire présente une « cohésion géographique, économique, culturelle ou sociale, à l'échelle d'un bassin de vie ou d'emploi » afin d'exprimer « la communauté d'intérêts économiques, culturels et sociaux de ses membres » et de permettre l'étude et la réalisation de projets de développement.

également de repérer des zones rurales en difficulté et de déterminer des pôles à conforter. Certains services des collectivités territoriales et de l'État ont largement utilisé ces référentiels à l'occasion. La même source répondait à de nombreuses demandes d'entreprises pour éclairer leurs études de marché.

L'inventaire communal était cependant contesté par certains acteurs, notamment en raison d'une qualité des réponses à l'enquête jugée insuffisante et de l'absence de procédures homogènes de validation. Les résultats, produits tous les dix ans avec des coûts de collecte élevés, étaient en outre rapidement obsolètes. Enfin, l'enquête ne permettait pas de mesurer les conditions d'accessibilité de la population aux équipements, en termes d'horaires d'ouverture ou de desserte par un transport en commun.

En 1994, une enquête « équipements urbains » a été réalisée sur le champ des communes de plus de 30 000 habitants. Elle avait pour but de suppléer aux lacunes de l'inventaire communal en milieu urbain. L'absence de localisation à l'adresse, la constitution de zones fixes et un processus de diffusion insuffisant ont nui à la publicité de l'enquête et à son utilisation.

► Encadré 1. Une base de 2,4 millions d'équipements*



Les équipements répertoriés relèvent de 7 domaines, qui ont été subdivisés en 27 sous-domaines et 188 catégories. 85 % des équipements sont géolocalisés avec une précision de moins de 100 m, seuls 0,4 % d'entre eux ne sont pas géolocalisés.

* millésime 2021

► L'introduction des sources administratives : des gains multiples

L'Insee décide donc au début des années deux-mille de lancer le projet de production de la base permanente des équipements, reposant sur des sources administratives en lieu et place d'enquêtes. Ces données proviennent soit du service statistique public, soit d'organismes effectuant des missions de service public (La Poste, ANPE puis Pôle emploi, etc.) qui les transmettent dans le cadre de conventions établies avec l'Insee.

Les gains apportés par ce choix sont multiples et visibles dès la première édition. En utilisant au maximum les sources administratives, les coûts sont bien inférieurs



En utilisant au maximum les sources administratives, les coûts sont bien inférieurs à ceux que nécessitaient les inventaires communaux.



à ceux que nécessitaient les inventaires communaux. L'éventail des catégories d'équipements recensés s'accroît nettement : l'inventaire communal de 1998 en comptait une trentaine, elles sont au nombre de 137 dès la BPE 2008 et de 188 dans la BPE 2021. La périodicité des résultats est annuelle, elle était quasi décennale auparavant. Les données millésimées au 1^{er} janvier de chaque année sont mises à disposition au bout de 18 mois. Les données géolocalisées depuis 2013 ouvrent

des perspectives d'utilisation bien supérieures, surtout en les combinant avec d'autres sources, elles aussi géolocalisées (Filosofi, Flores, etc.). Sous l'impulsion donnée par la loi pour une République numérique de 2016, le nombre de nouvelles sources mises à disposition s'accroît, laissant en germe la possibilité pour la BPE d'accroître encore le champ des équipements couverts.

Un autre gain, non des moindres, est celui de la capitalisation de l'expérience acquise au fil des ans. La pérennité du savoir-faire et l'amélioration de la qualité ne peuvent qu'être favorisées par la régularité des travaux et des relations avec les différents partenaires. Des conventions sont ainsi établies formalisant un cadre, définissant la nature de la collaboration, le contenu des données échangées, le calendrier, etc.

Mais comment procède-t-on au juste pour produire annuellement une base d'équipements aussi vaste ?

► **Figure 1 - Des fournisseurs de données variés et nombreux**



**Le dispositif mobilise 6 répertoires,
17 fichiers administratifs, 3 sources en open data,
et s'appuie sur des conventions avec 11 fournisseurs.**

► Un processus simple... en apparence

Chaque année, le processus de production commence par la collecte des données administratives auprès des partenaires (*figure 1*). Trois catégories peuvent être distinguées :

- en premier lieu, on utilise des sources internes à l'Insee. C'est le cas de Sirius qui est une base statistique découlant de Sirene⁵. La base de sondage de l'enquête de fréquentation dans les hébergements touristiques permet par ailleurs d'obtenir des données à l'adresse sur l'ensemble des hôtels et des campings de France ;
- la BPE se fournit également auprès de partenaires externes. Ce sont majoritairement des services statistiques ministériels. Ils couvrent des domaines variés : la santé, le sport, la culture, l'agriculture, la justice, l'éducation nationale, l'intérieur, le développement durable et les collectivités locales. D'autres acteurs institutionnels, comme Pôle emploi, La Poste et l'Agence nationale pour la cohésion des territoires (ANCT), transmettent annuellement des données permettant d'élargir la couverture de la BPE ;
- enfin, certaines données sont collectées en *open data*, telles celles sur les stations-service ou sur les établissements d'accueil auprès de jeunes enfants. Avant d'être utilisées, ces sources sont soumises à expertises amenant souvent à des contacts auprès des producteurs.

Au total, la BPE est alimentée par 6 répertoires et 17 fichiers administratifs.

► Mais le diable se cache dans les détails

La qualité et la nature des informations des différentes sources transmises sont cependant très hétérogènes. Par exemple, certaines sources ne contiennent pas d'informations permettant d'identifier sans ambiguïté la commune de localisation du service⁶. Les adresses des équipements sont par ailleurs présentées de manière très différente d'une source à une autre. Si certaines décomposent chaque élément d'adressage (numéro de voie, indice de répétition, type de voie, libellé de voie, code de la commune, libellé de commune), d'autres les regroupent dans une même variable, en y ajoutant même les noms ou raisons sociales des équipements ainsi que les compléments d'adresse, les boîtes postales et les cedex. Or, ce sont précisément ces éléments d'adressage qui prédéterminent la qualité de la géolocalisation.

L'architecture informatique comporte un ensemble de traitements mutualisés à toutes les sources.

Il est donc nécessaire d'harmoniser les données en entrée, par des traitements automatiques effectués au fur et à mesure de leur réception. L'architecture informatique comporte un ensemble de traitements mutualisés à toutes les sources : la création de variables d'adressage, la génération de codes communes

ou l'attribution du millésime souhaité du code officiel géographique, la suppression des doublons, etc. Des traitements spécifiques les complètent. Ainsi, certaines sources

⁵ Pour illustrer la différence entre Sirene, répertoire inter-administratif des entreprises et Sirius, répertoire statistique, voir l'article d'Ali Hachid et Marie Leclair dans ce même numéro.

⁶ En particulier, il arrive que la source ne contienne pas le code de la commune issu du code officiel géographique, mais seulement le code postal : or le code postal peut désigner plusieurs communes différentes.

contiennent des coordonnées GPS ; elles sont alors projetées dans les systèmes adéquats⁷ (Lapaine, Miljenko et Usery, 2014). De même, des caractéristiques complémentaires sont fournies pour les types d'équipements des domaines de l'enseignement et du sport, loisir et culture : présence de cantine, d'internat, équipement sportif couvert, éclairé, nombre d'aires de pratique, etc.

Une fois les sources harmonisées, elles sont appariées avec les sources de l'année précédente selon différentes clés de fusion qui permettent de repérer les équipements déjà existants, les nouveaux, ceux ayant disparu et ceux pour lesquels l'appariement est présumé. En effet, parmi les non appariés, certains présentent des éléments d'adressage et d'activité proches d'équipements présents l'année précédente. Des gestionnaires peuvent ensuite confirmer ou infirmer manuellement ces identifications présumées. Ces appariements permettent d'attribuer à chaque équipement des identifiants spécifiques à la BPE, ou de récupérer ceux de l'année précédente, et préparent ainsi la phase suivante des contrôles.

► Un contrôle des données nécessaire

Qu'une boulangerie cesse son activité, cela n'est pas étonnant en soi. La fermeture d'un service des urgences ou la disparition d'un aéroport interroge davantage. Ces cas sont plus fréquents qu'on ne pourrait le croire *a priori*. Dans la gestion interne des sources mobilisées par la BPE, des règles particulières peuvent induire des enregistrements tardifs. Par exemple, s'agissant des aéroports, le fournisseur de données ne peut transmettre que la liste de ceux ayant eu plus de 1 000 passagers au départ ou à l'arrivée (hors transit) au cours de l'année précédente : quelques petits aéroports ne franchissent pas cette barre certaines années. Il devient alors nécessaire d'identifier les **équipements structurants**, c'est-à-dire ceux qui sont assimilés à des équipements *a priori* pérennes, comme les hypermarchés, les lycées, les établissements de santé, les ports de plaisance, les zones de mouillage, les aéroports, etc. Quelle qu'en soit la cause, l'apparition ou la disparition d'une année à l'autre d'un équipement structurant dans une commune interpelle au point que cela justifie d'être contrôlé par un gestionnaire.

En complément et jusqu'en 2019, une opération de mesure de la qualité sur le terrain était réalisée chaque année. Elle bénéficiait de la collaboration des superviseurs du recensement de la population. Dans les communes de moins de 10 000 habitants, ceux-ci demandaient aux coordonnateurs communaux du recensement de valider une liste d'équipements, de la modifier et de la compléter le cas échéant. Des taux d'excédent et de déficit entre la base et le terrain étaient ainsi calculés, donnant une idée de la qualité de la couverture de la BPE sur les types d'équipements contrôlés. Cette opération est actuellement en suspens dans l'attente de sa refonte qui permettra d'accroître son efficacité et d'améliorer la représentativité des résultats.

⁷ Le RGF93 pour la métropole, l'UTM40 Sud pour La Réunion, l'UTM20 Nord pour la Martinique et la Guadeloupe, et l'UTM22 Nord pour la Guyane.



Chaque année, une base en évolution est diffusée.



Afin de répondre aux besoins des utilisateurs de la BPE, chaque année, une base en évolution est diffusée. Elle porte sur deux années espacées d'un pas quinquennal et s'effectue en géographie de l'année la plus récente⁸. Deux millésimes de la BPE ne sont pas immédiatement

comparables : estimer dans les comparaisons temporelles ce qui relève d'un changement de nomenclature, d'un changement de source, d'un changement de qualité dans la source ou enfin d'une évolution réelle des services accessibles sur le terrain s'avère particulièrement complexe. Des expertises sont menées pour déterminer quels types d'équipement peuvent être inclus dans le champ de diffusion de la BPE en évolution. La BPE en évolution 2016-2021 en géographie 2021 ; elle porte sur 119 types d'équipements (sur les 188 que contient la BPE 2021).

► Des coordonnées géographiques complètent les informations collectées

Après la phase des contrôles en bureau, vient celle de la géolocalisation des équipements. Elle est réalisée en deux parties, par une géolocalisation automatique d'abord, puis par des reprises manuelles. Celles-ci sont effectuées pour des équipements dont la qualité de géolocalisation est jugée insuffisante⁹. Au préalable, on réutilise le résultat des recherches manuelles des années passées pour réduire le volume des recherches. Des zonages sont aussi ajoutés lors de cette phase : les arrondissements, les zones d'emploi, les bassins de vie, etc.

Une fois ces opérations terminées, démarre l'aval du processus de production : la préparation des produits de diffusion. Les tâches sont nombreuses et sont coordonnées avec les différents acteurs impliqués au sein de l'Insee, qui parfois les mettent à disposition en l'état, parfois les valorisent en les transformant en d'autres produits (des indicateurs cartographiques, des tableaux de dénombrement, etc.). La mise à jour des métadonnées associées à la source BPE fait partie intégrante de cette phase. Sa volumétrie et les multiples formes qu'elle revêt nécessitent une attention particulière (voir *infra*).



Au total, le processus de production d'un millésime de la BPE dure 18 mois.

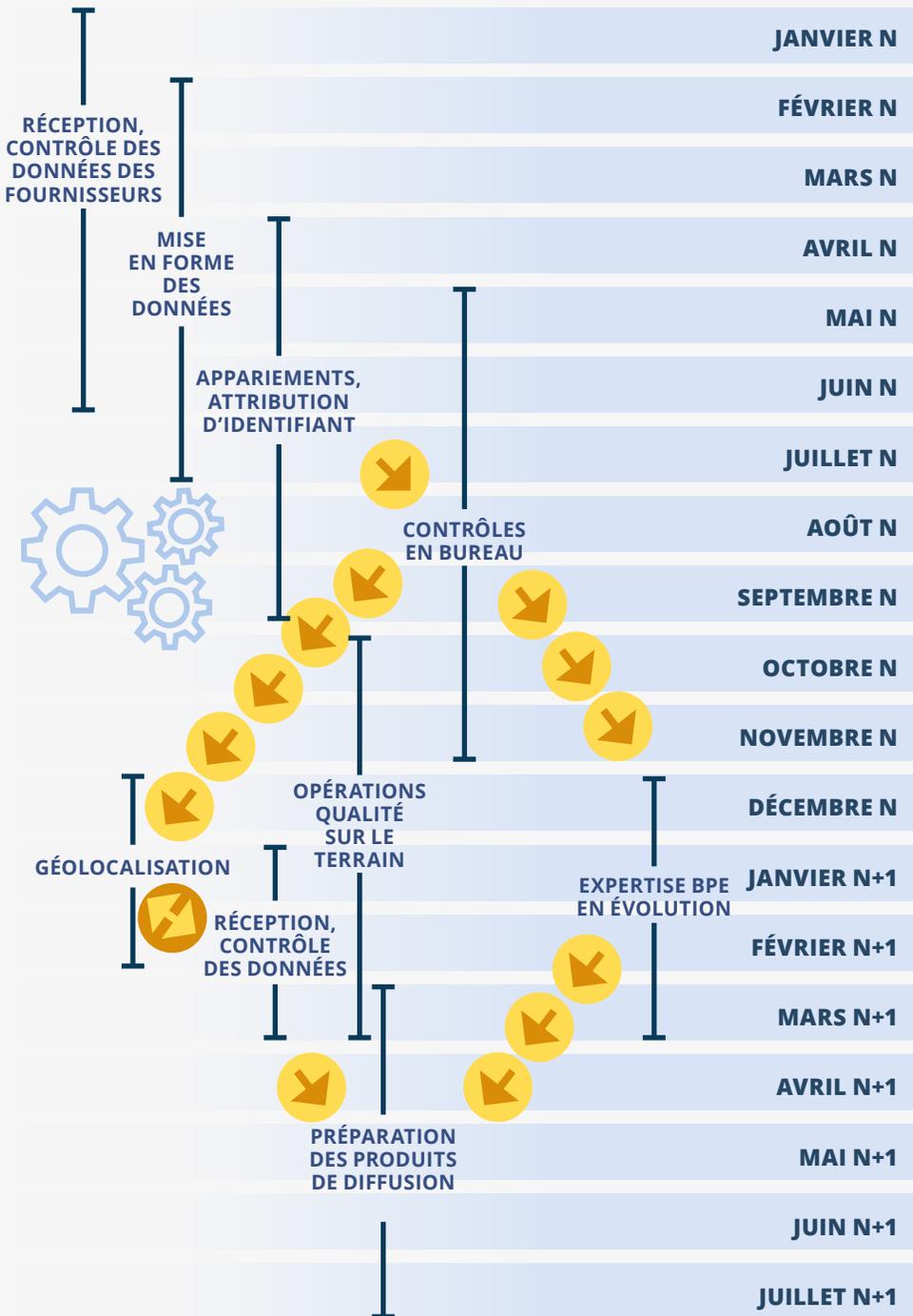


Au total, le processus de production d'un millésime de la BPE dure 18 mois (*figure 2*). Les différents produits d'un millésime dont les données sont référencées au 1^{er} janvier de l'année n sont diffusés au plus tard au début du mois de juillet de l'année n+1. Cette production permet de délivrer une offre de services diversifiée à différents utilisateurs (*figure 3*).

⁸ En cinq ans, le périmètre des communes peut évoluer, par suite de fusions de communes, de scission de portions de territoires, etc. Il importe donc d'avoir des « géographies » parfaitement comparables.

⁹ Les choix des types d'équipements faisant l'objet de reprises manuelles sont décidés en concertation entre le pôle BPE, la division Statistiques et analyses urbaines et le groupe Production localisée géographique (PLG).

► **Figure 2 - Une production cadencée sur 18 mois**



► Une offre qui s'adapte aux utilisateurs

Trois types de publics sont identifiés pour trois niveaux de diffusion¹⁰ :

- le grand public (incluant notamment les particuliers, les journalistes et les entreprises) ;
- les acteurs publics nationaux et régionaux¹¹ ;
- et les fournisseurs des données qui alimentent la BPE.

Les fichiers tous publics sont mis en ligne sur le site de l'Insee avec des fichiers d'ensemble et des bases de données thématiques. Il s'agit de dénombrements par type d'équipement, allant jusqu'au niveau de l'Iris¹², sur l'ensemble du territoire y compris les cinq départements d'outre-mer. Des fichiers détail sont également mis en ligne contenant les équipements avec leurs coordonnées géographiques et, le cas échéant, des caractéristiques complémentaires sur les domaines de l'enseignement (présence d'une cantine, d'un internat, secteur public ou privé, etc.) et des sports-loisirs (équipement couvert, éclairé, nombre de salles par structure culturelle, etc.). La mise en ligne d'un millésime annule et remplace le millésime précédent pour éviter des comparaisons temporelles inappropriées. Par ailleurs, à l'instar d'autres données, les données géolocalisées de la BPE sont utilisables sur l'espace « statistiques locales » pour concevoir des cartes à façon.

Certaines demandes sont adressées directement à l'Insee, émanant de journalistes, d'administrations, d'entreprises ou encore de particuliers et portant sur des sujets très divers : des précisions de concepts, des mises à disposition de données complémentaires, des données en évolution, etc. Les sites Insee Contact prennent en charge les réponses et font appel au pôle BPE le cas échéant. Toute structure chargée d'une mission de service public peut demander un extrait de la base sur sa zone de compétence et comportant des informations supplémentaires. Sous réserve de la signature d'une licence d'usage, cet extrait est diffusé gratuitement en application de l'arrêté du 20 juillet 2012 relatif aux modalités de diffusion de la BPE. Outre la dénomination et la localisation de l'équipement à l'adresse, sont fournies des variables de gestion (par exemple la date de création de l'équipement ou sa date de dernière mise à jour) et les variables spécifiques sur les champs de l'enseignement et des sports-loisirs (voir *supra*).

Enfin, la base est fournie chaque année gratuitement aux contributeurs, c'est-à-dire les fournisseurs alimentant la BPE, dans le cadre des conventions. Ceux-ci disposent ainsi de l'intégralité des données de la BPE, y compris les noms/raisons sociales des équipements et leur adresse.

Les produits mis à disposition sont majoritairement des produits intermédiaires adressés aux services de l'Insee chargés de les transformer en produits finaux. Ils sont formatés pour intégrer un entrepôt de données locales, dans lequel les services et les pôles d'études puisent pour réaliser leurs analyses ou actualiser leurs produits, le plus souvent en complémentarité avec d'autres sources. Ils sont transformés en indicateurs cartographiques permettant de réaliser des cartes à façon.

¹⁰ Voir l'arrêté du 20 juillet 2012 modifiant l'arrêté du 11 janvier 2008 portant création d'une BPE.

¹¹ Les personnes de droit public et les personnes de droit privé chargées d'une mission de service public (agence d'urbanisme, agence régionale de la santé, conseil départemental, Communauté d'agglomération, etc.).

¹² Îlots regroupés pour l'information statistique : quartier d'environ 2 000 habitants pour les communes de 5 000 habitants et plus, l'Iris constitue la brique de base en matière de diffusion de données infra-communales.

► Figure 3 - Que peut-on faire avec les produits de la BPE ?

LE PROCESSUS BPE :

1

UNE GAMME DE PRODUITS VARIÉS...

INDICATEURS
CARTOGRAPHIQUES

FICHIERS
ANONYMISÉS DES
ÉQUIPEMENTS
GÉOLOCALISÉS

EXTRAITS
DE LA BASE

BASES DE
DÉNOMBREMENT
PAR THÈME ET
À DIFFÉRENTS
ZONAGES

GAMMES
D'ÉQUIPEMENTS



2

DE MULTIPLES MOYENS D'ACCÈS AUX DONNÉES ...

 FranceArchives
PORTAIL NATIONAL DES ARCHIVES



UTILISATEURS
INTERNES



3

... POUR UN LARGE SPECTRE D'UTILISATEURS

ACTEURS
PUBLICS



CHERCHEURS



FOURNISSEURS
DES DONNÉES



CHARGÉS
D'ÉTUDE



GRAND PUBLIC



Ils deviennent des tableaux de comptage ou d'indicateurs de présence/absence à différents niveaux géographiques. Les produits finaux élaborés sont les bases fournies aux acteurs publics régionaux, aux fournisseurs et aux chercheurs¹³ et celles transmises pour archivage. Ces bases sont aussi accessibles en *open data* sur le site www.data.gouv.fr.

► La BPE utilisée pour enrichir les analyses territoriales

Favoriser l'accès des populations aux équipements et services publics est une problématique récurrente de nombreuses politiques publiques territoriales. Les plans d'action associés s'appuient sur un diagnostic territorial partagé à l'échelle du territoire. Un bilan de l'offre existante sur le territoire est alors réalisé, pour mettre en exergue d'éventuels déficits territoriaux et les rapporter aux populations concernées. La BPE est la source incontournable pour ce type de diagnostic. Elle permet ainsi de présenter les disparités locales en termes de taux d'équipement mais surtout en termes de temps d'accès de la population à ces équipements. Elle est pour cela couplée avec l'utilisation du distancier Metric¹⁴ développé par l'Insee. Les territoires éloignés des équipements sont ainsi repérés et la population concernée est estimée.

Par exemple, la BPE a été utilisée à la suite de la mise en place fin 2020 des conférences régionales du sport chargées « *d'instaurer une stratégie de développement du sport à l'échelle de la région à travers un projet sportif territorial* ». Ont été mis à disposition des acteurs locaux concernés des éléments tels que l'offre d'équipements sportifs, avec des comparaisons appropriées avec d'autres équipements, des cartes par établissement public de coopération intercommunale (taux d'équipement et accès), des éléments sur les disparités d'accès en milieu urbain et en zone rurale ou encore sur les temps d'accès des établissements scolaires aux équipements sportifs, et notamment à la piscine dans le cadre du plan « *aisance aquatique* ». Des dossiers thématiques et des analyses ont été livrés à l'image de celui de Normandie (*Chandavoine, Guérente, Hurard, Merel, Mura et Nadaud, 2021*).

La BPE est également utile pour mener des travaux sur le logement étudiant, avec l'approche fine souhaitée par les observatoires territoriaux du logement étudiant (OTLE). Lors de leur création en 2021, ces structures cherchaient à identifier les communes ayant sur leur territoire un établissement d'enseignement supérieur, pour affiner le champ des étudiants du supérieur tel que connu à partir du recensement de la population. Ces études ont conduit à diverses publications régionales, notamment en Auvergne-Rhône-Alpes (*Aude et Bianco, 2021*). La base a été mobilisée également pour les travaux sur l'école en milieu rural, en lien avec les conventions ruralité, pour aider à l'identification des territoires « *fragilisés* » où le temps d'accès des enfants devant quitter leur commune pour aller à l'école est potentiellement plus élevé qu'ailleurs. Il en est de même pour l'accès des jeunes non insérés à Pôle emploi ou aux établissements d'enseignement et de formation dans le cadre des pactes régionaux d'investissement dans les compétences.

¹³ Via Quetelet-Progedo-Diffusion.

¹⁴ Le distancier Metric (Mesure des trajets inter-communes / Carreaux) calcule des distances et des temps de parcours d'une commune à une autre ou d'un point à un autre lorsque les données sont géolocalisées. Il est mis à disposition via Quetelet-Progedo Diffusion.

► Des paniers d'équipements pour des catégories de population



Le panier généraliste « Vie courante », regroupe des besoins de tous les jours, les paniers « Seniors », « Jeunes » et « Familles » ciblent les besoins des populations plus âgées, plus jeunes et des familles avec enfants.



Pour des études plus fines, quatre **paniers thématiques** de commerces et services ont été définis par l'Insee, avec l'Agence nationale pour la cohésion des territoires et l'Institut d'aménagement et d'urbanisme d'Île-de-France. Ces paniers regroupent des équipements structurants pour certains types de population : le panier généraliste « Vie courante », regroupe des besoins de tous les jours, les paniers « Seniors », « Jeunes » et « Familles » ciblent les besoins des populations plus âgées, plus jeunes et des familles avec enfants.

Ainsi, les cartes d'accès au panier « Vie courante » et la part des populations éloignées constituent une base de travail utile pour les travaux d'élaboration de schémas d'aménagement du territoire, tels que les SDAASP¹⁵ et les SRADDET¹⁶. Il en est de même de l'accès au panier « Jeunes » pour les politiques territoriales dédiées à la jeunesse : une étude menée à La Réunion s'est ainsi intéressée à l'accès aux équipements favorisant l'insertion sociale et professionnelle des jeunes sur l'île (*Grangé et Merceron, 2020*). Le panier « Seniors » a été utilisé dans le cadre des schémas départementaux d'autonomie. À l'instar des travaux menés dans les Pays de la Loire (*Barre et Chesnel, 2019*), ces derniers permettent aux décideurs publics d'anticiper l'impact du vieillissement de la population dans les années à venir avec l'arrivée aux grands âges des générations issues du baby-boom, accentué par l'attractivité des territoires.

► Des métadonnées volumineuses et rationalisées

Dans toute production statistique, il importe que le producteur fournisse aux utilisateurs une documentation adaptée à chaque produit et à chaque usage. Cette étape importante du processus statistique s'appuie sur ce qu'on appelle les métadonnées. À l'Insee, ces informations sont gérées dans le référentiel de métadonnées statistiques RMÉS (*Bonnans, 2019*).

Dans le cas de la BPE, la diversité des produits livrés, intermédiaires et finaux et la multiplicité des destinataires directs en aval du processus de production induisent de produire – et de mettre à jour – une documentation particulièrement abondante, sous des formats variés.

De nombreuses métadonnées sont ainsi associées à chaque produit : un dictionnaire des variables, un dessin de fichier, une note de mise à disposition, une note explicative complémentaire le cas échéant, la définition des types d'équipements, etc. Pour un même produit, ces métadonnées diffèrent selon les destinataires et sont à actualiser chaque année.

¹⁵ Schéma départemental d'amélioration de l'accessibilité des services au public.

¹⁶ Schéma régional d'aménagement, de développement durable et d'égalité des territoires.

Lors de leur refonte, ces métadonnées ont été structurées selon les normes adoptées par Eurostat : le SIMS (*Single integrated metadata structure*¹⁷) et le standard DDI (*Data Documentation Initiative*). Pour ce faire, les producteurs de la BPE ont bénéficié des services du référentiel RMÉS et de l'appui des experts « qualité » de l'Insee. Au final, dans RMÉS, la BPE est dotée d'un système de métadonnées original, qui lui est propre, conçu pour optimiser les mises à jour et en alléger la charge induite.

Les métadonnées en sortie de RMÉS ont été construites pour répondre aux besoins en documentation des utilisateurs de la BPE, mais également pour être utilisées en *input* de leur propre processus. Par exemple, les notes de mise à disposition ou les notes explicatives complémentaires sont constituées de « briques de métadonnées » venant de RMÉS. Certains outputs de la BPE, destinés à préparer des produits finaux de diffusion, sont accompagnés de fichiers *XML* qui peuvent être injectés en entrée d'un processus de conception d'indicateurs cartographiques ou de tableaux thématiques, et mettre à jour automatiquement la documentation qui les accompagne.

Avec la mise au point de ce système de métadonnées, la BPE a également gagné en cohérence et en homogénéité des documents et des fichiers produits en aval.

► Une démarche qualité comme fil conducteur des évolutions

La rationalisation entreprise avec succès sur les métadonnées, a succédé aux gains déjà constatés sur la qualité des données elles-mêmes, lors de la bascule de SAS® en R des programmes qui harmonisent les différentes sources (voir *supra*).

► Encadré 2. Un nécessaire comité des utilisateurs de la BPE

Pour pouvoir dresser le bilan des utilisations passées de la BPE et recueillir les nouveaux besoins des utilisateurs, qu'ils proviennent de l'Insee ou de l'externe, un comité des utilisateurs est une instance précieuse. Cette instance décrit et priorise les évolutions attendues et qui seront soumises au comité de pilotage de la BPE.

Le comité se réunit une fois par an. Il est constitué de représentants de l'Insee, de représentants de la Fédération nationale des agences d'urbanisme et de l'Agence nationale de la cohésion des territoires. Les demandes récurrentes des internautes sont également discutées, celles-ci parvenant régulièrement au pôle BPE *via* le service Insee contact. Le comité se réunit une fois par an.

Cette instance très utile est à l'origine de la diffusion de la BPE en évolution sur un pas quinquennal glissant (2015-2020, 2016-2021, etc.) en complément de la base millésimée, à la suite de nombreux besoins exprimés. Elle permet aussi de cibler les investigations quant à l'extension du champ couvert par la BPE. De nouveaux types d'équipement ont ainsi été intégrés : les mairies, les théâtres, les salles de concert, les implantations France services, etc. Des remarques sur la présence de faux actifs dans la BPE formulées lors de ce comité ont aussi eu pour conséquence le remplacement de la source la plus volumineuse (Sirene) par la source Sirius, répertoire statistique des entreprises (voir l'article 6 « Sirius » dans ce numéro du *Courrier des statistiques*).

¹⁷ Voir <https://ec.europa.eu/eurostat/fr/data/metadata/metadata-structure> : Eurostat distingue les métadonnées de référence (concepts, méthodologies utilisées, qualité des données) et les métadonnées structurelles (pour identifier les données statistiques).

La modernisation des autres étapes du processus nécessitait une analyse préalable et des outils adéquats. Une démarche qualité a donc été menée en 2021 par le pôle de production de la BPE, avec la collaboration de l'Unité Qualité de l'Insee : il a fallu remettre à plat l'existant, identifier les principaux risques du dispositif, ainsi que ses forces et faiblesses. Classiquement, l'enjeu pour la BPE était double : il s'agissait d'une part de rationaliser les procédures et d'en alléger la charge, et d'autre part de gagner en souplesse pour faciliter la prise en compte des demandes des utilisateurs (**encadré 2**).

En premier lieu, il s'agissait de situer le processus de production dans son environnement immédiat, en décrivant les échanges qui se déroulent avant et après le processus, selon l'approche du *Sipoc*¹⁸. Une analyse détaillée du processus a ensuite permis de le segmenter en briques appuyées sur le phasage du GSBPM¹⁹ et d'identifier plus précisément, à chaque phase, les *inputs*, les traitements opérés, les *outputs*, les acteurs et leur rôle. S'en est suivie une analyse des forces et faiblesses selon les critères du Code des bonnes pratiques de la statistique européenne. La mise en exergue des risques de tout ordre (techniques, organisationnels, managériaux, etc.) pesant sur le bon déroulement des opérations a permis d'établir un plan de maîtrise des risques, identifiant trente actions pour les réduire. À titre expérimental, une analyse de la validation des données avant diffusion a aussi été testée.

Au final, un premier plan d'action qualité a été établi et validé par la maîtrise d'ouvrage comme fil conducteur de la modernisation de la BPE. Il comporte 18 actions s'étalant de juin 2021 à juin 2023. Certaines portent sur la rénovation de la chaîne de traitement, d'autres sur la description de l'opération conformément aux standards de RMÉS. Le plan aborde aussi la révision du processus de contrôle et le partage des connaissances pour assurer la continuité de service.

► Un processus en constante évolution

L'analyse menée a notamment pointé le caractère extrêmement mouvant du processus de production de la BPE. Par nature, celui-ci est basé sur une évolution permanente des sources en entrée ; chaque année, à la différence des répertoires comme Sirene ou le RNIPP²⁰, de nouvelles sources sont intégrées à la base. Par ailleurs, compte tenu du concept même d'équipement, l'exhaustivité est impossible à atteindre. Si l'exhaustivité est visée, ce n'est pas sur l'ensemble des équipements offerts à la population, mais plutôt sur le champ couvert par les sources. Ainsi, certains services offerts en activités secondaires de commerces (par exemple des points de retraits de commerces en ligne) ou des infrastructures mises à disposition de tous (comme des aires de covoiturage ou des bornes de recharge de véhicules électriques) sont encore mal couverts par la statistique publique. Il est de ce fait difficile de parler de *répertoire* des équipements, encore moins de référentiel²¹.

18 *Supplier input process output customer* : « fournisseur /intrant /processus /extrant /client ».

19 *General statistical business process model* ou modèle générique de description des processus de production statistique. Voir (Unece, 2019) et (Erikson, 2020). Les phases de ce modèle offrent une grille d'analyse utile.

20 Voir l'article de Lionel Espinasse et Valérie Roux dans ce même numéro.

21 Voir l'article de Pascal Rivière dans ce même numéro, pour retrouver la définition d'un répertoire ayant statut de référentiel.

Cependant, le pôle BPE recherche constamment à intégrer de nouveaux types d'équipement. Régulièrement, de nouveaux fournisseurs sont associés au processus, par exemple en 2019 le Commissariat général à l'égalité des territoires, devenu Agence nationale pour la cohésion des territoires. Par ailleurs, les sources évoluent au fil des ans : les fournisseurs adoptent de nouveaux systèmes d'information, sont soumis à de nouvelles règles de gestion et de transmission d'information, etc. Ainsi, la classification *ad hoc* des types d'équipement évolue-t-elle chaque année pour intégrer les nouveautés.

L'adaptabilité des traitements et la réactivité des personnes en charge de la production de la BPE sont importantes. Cette nécessité s'illustre notamment dans le développement des traitements informatiques. Certaines opérations sont réalisées en interne, au moyen de programmes en R. Ces programmes viennent en complément d'une application développée et maintenue par l'Insee, qui permet de bénéficier d'une interface utile pour structurer les phases et faciliter la visualisation des opérations. La collaboration entre les agents du pôle BPE et les informaticiens permet de sécuriser la production de la base et de répondre aux normes de sécurité informatique de l'Insee.

Cependant, le caractère fortement évolutif des sources, du champ couvert, des informations disponibles imposent une réactivité forte aussi bien de l'informatique que du pôle BPE. Or, les modifications de la structure des sources sont parfois constatées une fois la source transmise par le fournisseur. De même, les informations nouvelles disponibles qui permettraient d'enrichir la BPE ne sont pas signalées dans des délais permettant de les prendre en compte. Elles peuvent en effet nécessiter de réaliser des développements informatiques pour faire évoluer l'application.

Depuis 2020, des travaux sont en cours pour passer d'un processus centré autour d'une application monolithe à une chaîne multi-outils. En segmentant le processus et en ciblant les *inputs*, les *outputs* et les traitements à effectuer, il est possible d'envisager des outils simples, dédiés à un nombre limité de tâches. Ce changement de paradigme permettra de simplifier la maintenabilité de la chaîne de traitement, de faciliter la prise en compte de nouvelles données disponibles et donc d'enrichir plus facilement la BPE. Il permettra d'être plus réactif aux demandes des utilisateurs. L'occasion se présente aussi de développer un nouveau support : une API²² dont l'ambition serait de permettre aux fournisseurs de données d'insérer automatiquement dans leur système d'information les améliorations de la BPE (sur la qualité de l'adressage, sur l'ajout de zonages, etc.). L'année suivante, ce serait la BPE qui en serait bénéficiaire : un cercle vertueux se mettrait en place à chaque fourniture de données.

²² *Application Programming Interface* ou interface de programmation applicative en français. Il s'agit d'un ensemble normalisé de classes, de méthodes, de fonctions qui sert de façade par laquelle un logiciel offre des services à d'autres logiciels.

► Bibliographie

- BARBIER, Max, TOUTIN, Gilles, LEVY, David, 2016. *L'accès aux services, une question de densité des territoires*. [en ligne]. 6 janvier 2016. Insee Première, n°1579. [Consulté le 4 septembre 2022]. Disponible à l'adresse : <https://www.insee.fr/fr/statistiques/1908098>.
- BARRÉ, Martine et CHESNEL, Hélène, 2019. *La hausse du nombre de seniors dépendants accélérerait à partir de 2023*. [en ligne]. 13 juin 2019. Insee Analyses Pays-de-la-Loire, n°75. [Consulté le 4 septembre 2022]. Disponible à l'adresse : <https://www.insee.fr/fr/statistiques/4169347>.
- BONNANS, Dominique, 2019. RMÉS, le référentiel de métadonnées statistiques de l'Insee. In : *Courrier des statistiques*. [en ligne]. 27 juin 2019. Insee. N° N2, pp. 46-57. [Consulté le 4 septembre 2022]. Disponible à l'adresse : <https://www.insee.fr/fr/statistiques/fichier/4168396/courstat-2-6.pdf>.
- CHANDAVOINE, Bruno, GUÉRENTE, Sylvie, HURARD, Camille, MEREL, Aubin, MURA, Bruno et NADAUD, Laurence, 2021. *Le sport en Normandie : pratiques, équipements et emplois*. [en ligne]. 7 octobre 2021. Insee Dossier Normandie, n°19. [Consulté le 4 septembre 2022]. Disponible à l'adresse : <https://www.insee.fr/fr/statistiques/5435270>.
- ERIKSON, Johan, 2020. Le modèle de processus statistique en Suède – Mise en œuvre, expériences et enseignements. In : *Courrier des statistiques*. [en ligne]. 29 juin 2020. Insee. N° N4, pp. 122-141. [Consulté le 4 septembre 2022]. Disponible à l'adresse : <https://www.insee.fr/fr/statistiques/fichier/4497085/courstat-4-8.pdf>.
- GRANGÉ, Claire et MERCERON, Sébastien, 2020. *Une proximité qui ne suffit pas à réduire les difficultés de formation et d'emploi. Équipements pour l'insertion sociale et professionnelle des jeunes Réunionnais*. [en ligne]. 19 novembre 2020. Insee Analyses Réunion, n°50. [Consulté le 4 septembre 2022]. Disponible à l'adresse : <https://www.insee.fr/fr/statistiques/4964208>.
- HAUTDIDIER, Baptiste et KUENTZ, V., 2011. Quelles invariances d'échelle entre densité de la population et des équipements ? Une approche empirique sur données communales françaises. In : *Rencontres de Théo Quant 2011*. Février 2011. Besançon, France. pp.32.
- KOMPIL, Mert, JACOBS-CRISIONIA, Chris, DIJKSTRA, Lewis et LAVALLE, Carlo, 2018. Mapping accessibility to generic services in Europe: A market-potential based approach. In : *Sustainable Cities and Society*. [en ligne]. 5 décembre 2018. n°47 (May, 2019) 101372. [Consulté le 4 septembre 2022]. Disponible à l'adresse : <https://doi.org/10.1016/j.scs.2018.11.047>.
- LAPAINE, Miljenko et USERY, E. Lynn, 2014. Projections cartographiques et systèmes de références. In : *Comité français de cartographie (CFC)*. [en ligne]. Septembre 2014. N°221, Chapitre 9, pp.87-101. Traduction par Didier Halter et Jean-François Hangouët. [Consulté le 4 septembre 2022]. Disponible à l'adresse : <https://www.lecfc.fr/new/articles/221-article-11.pdf>.
- UNECE, 2019. *Generic Statistical Business Process Model*. [en ligne]. Janvier 2019. Version 5.1. [Consulté le 4 septembre 2022]. Disponible à l'adresse : <https://statswiki.unece.org/display/GSBPM/GSBPM+v5.1>.