

DREES MÉTHODES

N° 9 • mars 2023

Identification des seniors en hébergements dédiés aux personnes âgées dans les données Fidéli

Litti Esteban (DREES)

Identification des seniors en hébergements dédiés aux personnes âgées dans les données Fidéli

Litti Esteban (DREES)

Remerciements : Patrick Aubert, Angélique Balavoine, Stéphanie Boneschi, Amélie Carrère, Camille Freppel, Julie Labarthe, Pierre Lamarche, Albane Miron de L'Espinay, Layla Ricoch

Retrouvez toutes nos publications sur : drees.solidarites-sante.gouv.fr

Retrouvez toutes nos données sur : data.drees.solidarites-sante.gouv.fr

DREES MÉTHODES

N° 9 • mars 2023

Synthèse Identification des seniors en hébergements dédiés aux personnes âgées dans les données Fidéli

Litti Esteban (DREES)

Retrouvez toutes nos publications sur : drees.solidarites-sante.gouv.fr

Retrouvez toutes nos données sur : data.drees.solidarites-sante.gouv.fr

SYNTHÈSE

Peu de sources permettent d'estimer les revenus des seniors en établissements pour personnes âgées ou en résidences-services seniors. L'identification de ces populations dans les données individuelles de Fidéli (fichiers démographiques sur les logements et les individus) de l'Insee permet d'en produire une description fine, mise à jour annuellement, et robuste. Cette identification est possible après un appariement entre les données Fidéli et les tables des lieux de vie des seniors, de deux types :

- Les établissements sanitaires, sociaux ou médicaux sociaux pour personnes âgées recensés dans le répertoire Finess (fichier national des établissements sanitaires et sociaux). En particulier les EHPAD (établissement d'hébergement pour personnes âgées dépendantes), EHPA (établissement d'hébergement pour personnes âgées), USLD (unités de soins de longue durée) et résidences autonomie ;
- Les résidences-services seniors pour lesquelles il n'y a pas de répertoire officiel, ce qui nécessite un recensement sur internet.

Cet appariement se base sur l'adresse postale, qui se présente sous forme de libellé dans Finess et dans la base des résidences-services seniors, et sous forme de code (code Rivoli) dans les données Fidéli. Il est donc nécessaire d'utiliser une table de passage, la base Fantoir (fichier annuaire topographique initialisé réduit, anciennement fichier Rivoli).

Dans un premier temps, les établissements Finess et les résidences-services seniors sont appariées à Fantoir en utilisant les similarités textuelles des libellés d'adresse. Ceci permet d'obtenir l'adresse codée (adresse brute) de tous les établissements et résidences-services. Puis dans un second temps ces deux bases sont appariées aux données Fidéli, en utilisant une stratégie différente pour les résidences-services seniors et pour les établissements Finess. Les résidences-services sont directement appariées aux individus de Fidéli grâce à l'adresse brute issue de l'appariement résidences-services - Fantoir. Pour les établissements Finess en revanche, deux approches sont combinées. D'une part, les adresses déclarées dans la table des individus de Fidéli sont analysées afin de déterminer si elles sont susceptibles d'être des adresses d'établissements pour personnes âgées ou non, i.e. elles ont été repérées comme telles par l'Insee ou repérées comme telles dans en raison d'une accumulation de personnes âgées à l'adresse. D'autre part, les adresses brutes issues de l'appariement Finess – Fantoir sont recherchées dans la liste des adresses de Fidéli. Ces deux approches se combinent comme suit : les adresses brutes de Finess sont recherchées dans la liste des adresses brutes de Fidéli, puis pour chaque établissement, les adresses proches dans Fidéli (dans la même rue notamment) identifiées comme étant susceptibles d'être des établissements pour personnes âgées sont recherchées. Les adresses ainsi retrouvées sont associées à l'établissement. Ceci permet de palier aux erreurs légères des adresses Finess, et surtout aux erreurs de déclaration des résidents. Un établissement Finess est apparié de manière sûre à Fidéli si au moins une adresse de Fidéli susceptible d'être l'adresse d'un établissement pour personnes âgées lui a été associée. Une fois l'appariement avec les adresses brutes de Fidéli terminé, on peut associer les individus de Fidéli vivant à chaque établissement.

Les résultats de ces appariements sont satisfaisants. Selon une étude de *Matières Grises*¹, on dénombrait autour de 750 résidences-services seniors en 2019. Avec la méthode d'appariement, 638 résidences-services seniors sont retrouvées, pour 32 900 résidents, dont 83 % ont plus de 60 ans. Concernant les établissements sanitaires et sociaux pour personnes âgées (Finess), les résultats dépendent de la catégorie de l'établissement. 94 % des établissements d'hébergement pour personnes âgées dépendantes (EHPAD) de départ sont retrouvés, et 3 résidences autonomie sur 4 sont retrouvées. La méthode est en revanche moins adaptée pour les EHPA avec 61 % des établissements retrouvés, et elle n'est en pas adaptée pour les USLD, avec moins d'un établissement sur deux retrouvé. Dans l'ensemble, 88 % des établissements sont retrouvés dans les tables Fidéli (tableau A), et regroupent 437 300 individus, dont 90 % ont plus de 60 ans.

Une expertise qualité de la méthode d'appariement est menée en confrontant les caractéristiques de la population identifiée en établissements sanitaires et sociaux dans Fidéli aux caractéristiques connues de cette population dans d'autres jeux de données (enquête Care Institution, enquête EHPA). Dans l'ensemble la population retrouvée dans Fidéli a des caractéristiques proches de ce qui est attendu en termes de répartition par âge et sexe, et de niveau de pension.

¹ Kuhn Lafont, A., Troutot, R. (2022). Résidences services seniors : appréhender une offre devenue incontournable pour les territoires. Les études de *Matières Grises*, mars 2022.

Tableau A • Scores d'appariement par catégorie d'établissement

Catégorie	Nombre	Répartition (%)	Part à être appariée de manière sûre (%)
EHPAD privés à but lucratif	1817	17.7	95.4
EHPAD privés à but non lucratif	2359	22.9	92.6
EHPAD publics hospitaliers	1210	11.8	90.2
EHPAD publics non hospitaliers	2133	20.7	95.3
Résidences autonomie	2290	22.3	74.2
EHPA	314	3.1	61.1
USLD	166	1.6	45.8
Ensemble	10 289	100	87.6

Champ > Établissements pour personnes âgées dans Finess, France métropolitaine et DROM.

Source > Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Cependant par construction cette méthode ne permet pas de retrouver l'ensemble des résidents d'établissements pour personnes âgées ou de résidences-services seniors. La base Individu de Fidéli étant un assemblage de données administratives, notamment fiscales, l'adresse associée à un individu provient du fichier d'imposition des personnes (FIP) et correspond donc à l'adresse à laquelle l'individu est fiscalement déclaré. Ainsi l'appariement entre les établissements et résidences-services seniors et la base Individu ne permet de retrouver que les individus qui sont fiscalement déclarés à leur établissement d'hébergement.

Dans les établissements sanitaires et sociaux, on estime que c'est le cas pour seulement 62 % des résidents. Parmi les 38 % restant, on trouve par exemple les seniors en couple dont le conjoint est à domicile (situation imputable à 10 % des résidents d'établissements pour personnes âgées²), puisqu'il est probable que l'adresse fiscale du couple soit celle du domicile. Ceci est cohérent avec le fait que la méthode identifie moins de seniors en couple que ce qui est obtenu dans d'autres enquêtes (7 % des seniors en établissement identifiés par la méthode sont mariés ou pacés, contre 11 % des résidents dans l'enquête Care-I 2016). Il peut également arriver qu'un senior soit toujours propriétaire de son ancienne résidence principale, et continue de s'y déclarer fiscalement, même sans avoir de conjoint qui y habite. Enfin parmi les résidents non repérés fiscalement à leur établissement on trouve également les seniors dont la durée de séjour en établissement est trop courte pour qu'ils aient le temps de s'y déclarer fiscalement (25 % des résidents en hébergement permanent dans un établissements pour personnes âgées ont une ancienneté inférieure à 11 mois à la fin de l'année 2019³).

Ce type d'analyse ne peut pas être mené dans les résidences-services seniors. En effet, la capacité totale n'étant pas disponible, il n'est pas possible de calculer la proportion des résidents à être déclarés à la résidence. Le taux de déclaration y est toutefois supposé plus élevé car les logements en résidences-services sont des logements ordinaires, il est donc probable qu'ils constituent le domicile principal des seniors, auquel ils se déclarent fiscalement. Aussi, il est sûrement plus courant d'emménager en couple dans une résidences-services seniors que dans un EHPAD, où l'un des membres du couple peut rester au domicile tandis que l'autre a besoin de soins médicaux dans un établissement spécialisé. On fait l'hypothèse que les situations telles que l'un des deux conjoints est en résidences-services seniors et l'autre à l'ancien domicile sont marginales. Enfin, les résidences-services seniors étant non médicalisées, il est raisonnable de penser que les individus qui y emménagent sont en meilleure santé que dans les EHPAD, et donc avec une espérance de vie plus longue. Il y a une plus grande probabilité qu'ils aient le temps de se déclarer fiscalement à leur nouveau domicile en résidences-services seniors.

L'identification de la population en établissements sanitaires et sociaux pour personnes âgées et en résidences-services seniors dans les données individuelles de Fidéli permet de produire des statistiques annuelles et robustes sur leur composition en termes d'âge, de sexe et de situation familiale et de décrire leurs revenus individuels, en particulier leur pension. Pour les seniors en résidences-services seniors, aucune étude de ce type n'est à l'heure actuelle disponible. De plus, les seniors vivant en résidences-services seniors étant en logement ordinaire, les données Fidéli permettent d'avoir accès aux revenus à l'échelle de leur ménage, notamment leurs prestations sociales, ou leurs revenus fonciers et mobiliers. L'identification des adresses d'établissements pour personnes âgées et de résidences-services seniors dans Fidéli permet également d'étudier les mobilités d'un type d'habitat à l'autre, lorsque le changement a eu lieu au cours de l'année 2019.

² Enquête EHPA 2019, 9.5% des résidents d'établissements ont leur conjoint hors de l'établissement (data.drees).

³ Enquête EHPA 2019, DREES.

SOMMAIRE

■ INTRODUCTION	2
■ BASES DE DONNÉES ET PRÉTRAITEMENTS	4
Finess	4
Résidences-services seniors (RSS)	4
Mise en forme des bases « sanitaire et social » et « cap résidence »	5
Ajout du code géographique	5
Repérage des doublons	5
Fidéli	6
Fantoir	6
■ MÉTHODE D'APPARIEMENT	7
Appariement des établissements et des résidences-services seniors avec Fantoir. 7	
Étapes de l'appariement	7
Résultats RSS	8
Résultats Finess	9
Synthèse	9
Retrait des établissements Finess dans la base RSS	10
Appariement des établissements et résidences enrichis de Fantoir avec Fidéli	11
Établissements RSS	11
Établissements Finess	12
Résultats des appariements	18
Résultats généraux	18
Scores d'appariement Finess par catégorie d'établissement	19
Adresses identifiées par un des critères	21
Apport de la technique d'appariement des adresses proches pour Finess	22
■ EXPERTISE QUALITÉ DE LA MÉTHODE ET PREMIÈRES STATISTIQUES SUR LA POPULATION IDENTIFIÉE	24
Taux de déclaration	24
Comparaison avec EHPA 2019	26
Structure générale	27
Établissement par établissement	28
Analyse de la pension selon la catégorie de l'établissement Finess	29
Comparaison avec Care-Institution	30
Données sur les revenus du ménage	31
Description de la population identifiée	32
■ CONCLUSION	35
■ POUR EN SAVOIR PLUS	36
Annexe 1. Caractéristiques des établissements Finess non appariés	37
Annexe 2. Mesures de similarité textuelle	40
Annexe 3. Détermination du critère d'accumulation	41
Annexe 4. Autres tableaux	46

■ INTRODUCTION

Peu de sources permettent d'avoir une description de la composition des revenus des seniors vivant en établissements pour personnes âgées et en résidences-services seniors (RSS, voir *encadré 1* pour une définition) et elles sont produites à une fréquence supra-annuelle. Ce constat s'explique en partie car les données de la statistique publique couvrent en général le champ des ménages dits « ordinaires », et excluent donc les seniors en institution. La source Fidéli (fichiers démographiques sur les logements et les individus) fournit les revenus individuels déclarés de l'ensemble des individus vivant en France métropolitaine et en départements et régions d'outre-mer (DROM), qu'ils vivent ou non en logement ordinaire. Ainsi, il est possible de rechercher dans cette table les seniors vivant en établissement pour personnes âgées ou en RSS, afin de produire une description annuelle de cette population en termes de revenus, mais aussi d'âge et de situation matrimoniale. Toute la difficulté réside dans cette identification, qui constitue alors un apport non négligeable pour parvenir à une description fine des seniors.

L'ampleur et les enjeux de cet apport ne sont pas les mêmes selon le champ des seniors considérés (établissements pour personnes âgées ou RSS). Concernant les seniors en établissements pour personnes âgées, nous disposons déjà de données administratives et d'enquête permettant de les décrire en partie

Du côté des données administratives, les remontées administratives RI-APA-ASH auprès des conseils départementaux permettent d'avoir des informations individuelles sur les demandeurs et les bénéficiaires de l'APA et de l'ASH. Ces remontées obligatoires permettent de collecter des caractéristiques sociodémographiques des bénéficiaires, leur niveau de dépendance détaillé ainsi que l'historique des évolutions du niveau de dépendance depuis la première demande d'APA, leurs ressources au sens de l'APA, les montants de leurs plans d'aide et le contenu de ces plans. Cependant, ces données ne permettent pas de retrouver l'ensemble des seniors en établissement, mais seulement ceux qui ont demandé l'APA ou l'ASH⁴. Enfin, contrairement à Fidéli qui est mise à jour annuellement, ces remontées sont quadriennales.

Du côté des données d'enquête, nous disposons des enquêtes Care (Capacités, Aides et REssources des seniors) de 2015 et 2016. Les enquêtes Care ménage (15 000 seniors) et Care institution (3 000 seniors) en France métropolitaine permettent d'obtenir des informations sur l'âge, le GIR et également sur les revenus et les prestations sociales des seniors, car elles ont été appariées aux déclarations fiscales de revenus et aux données de certaines caisses d'allocations⁵. Malheureusement cette enquête n'est conduite que rarement, et les données ne sont pas exhaustives. Une autre source importante provient de l'enquête EHPA (enquête auprès des établissements d'hébergement pour personnes âgées). Elle recueille de manière exhaustive des informations sur l'activité des établissements, sur le personnel et sur les personnes âgées qui y résident. Il n'y a cependant aucune variable relative au revenu des résidents. Elle présente également d'autres limites : elle n'a lieu que tous les 4 ans, et le taux de non-réponse concernant les données individuelles des résidents est très élevé (autour de 60 %). Enfin, l'enquête Autonomie (dont les résultats ne sont pas encore disponibles) permet d'obtenir des renseignements relatifs à la perte d'autonomie sur l'ensemble de la population, y compris la population en établissements. Elle sera appariée à l'ERFS et permettra donc d'obtenir une description fine des revenus des individus en établissement. Cependant les enquêtes de ce type n'ont lieu que tous les 10 ans.

Un repérage des seniors vivant en établissements pour personnes âgées dans Fidéli permet plusieurs apports : des données annuelles, exhaustives, et incluant des variables sur les revenus individuels des seniors (en particulier la pension). Enfin, les seniors en résidences-services seniors font partie du champ ordinaire des données d'enquêtes classiques (ERFS par exemple). Cependant, ils ne sont pas directement identifiables et isolables. Peu d'études portent sur les RSS. On note par exemple l'étude du think tank *Matières Grises* (Kuhn Lafont, Troutot, 2022), mais il n'y a que très peu de statistiques sur les caractéristiques socio-démographiques des résidents. Repérer les résidents des RSS dans la table Fidéli permet donc de produire une première description de ces individus, et en particulier de leurs revenus.

L'identification des seniors en résidences-services seniors et en établissements d'hébergement pour personnes âgées est possible grâce à un appariement entre les tables des lieux de vie des seniors et la table des individus de Fidéli, en utilisant l'adresse de résidence. Les lieux de vie des seniors sont recensés d'une part dans le répertoire Finess (fichier national des établissements sanitaires et sociaux) pour les établissements sanitaires, sociaux ou médicaux sociaux pour personnes âgées, et d'autre part dans une table construite grâce à un recensement web pour les résidences-services seniors. Dans ces tables de lieux de vie, l'adresse est codée sous forme de libellé, tandis qu'elle est sous forme de code Rivoli dans la table Fidéli. Une table de passage, la table Fantoir (fichier annuaire topographique initialisé réduit, anciennement fichier Rivoli) doit donc être utilisée.

⁴ En 2017, 90% des résidents d'EHPAD, EHPA ou USLD en France métropolitaine bénéficient de l'APA, EHPA 2019, RI-APA-ASH 2017.

⁵ Caisse nationale des allocations familiales (Cnaf), Caisse centrale de mutualité sociales agricole (CCMSA), Caisse nationale d'assurance vieillesse (Cnav), Caisse centrale de la mutualité sociale agricole (CCMSA).

Après une présentation des bases de données utilisées, ce dossier détaille la méthode d'appariement : l'appariement entre les tables de lieux de vie et la table Fantoir, puis l'appariement avec les données Fidéli. Une expertise qualité de la méthode est ensuite menée, notamment en comparant la structure de la population identifiée avec ce qui est obtenu par d'autres sources. Enfin, une première description des seniors en établissements pour personnes âgées et en résidences-services seniors est disponible à la fin de ce dossier.

■ BASES DE DONNÉES ET PRÉTRAITEMENTS

Finess

La table des établissements pour personnes âgées est extraite du répertoire Finess, qui recense tous les établissements sanitaires, sociaux et médico-sociaux. Elle comporte une ligne par établissement, identifié par son numéro Finess, avec le code commune Insee, l'adresse en toute lettre (libellé d'adresse) au 1^{er} janvier 2020, la catégorie de l'établissement et sa capacité.

Sont sélectionnés les établissements Finess de 2019 qui sont de l'une de ces catégories : EHPAD, EHPA, USLD, résidences autonomie⁶. Ce sont 10 289 établissements pour personnes âgées qui sont ainsi recensés au niveau métropole et DROM en 2019.

L'ensemble de ces établissements est apparié à Fidéli. Cependant, 232 n'ont pas d'adresse renseignée, soit 2,3 % des établissements. Parmi ces établissements sans adresse, la moitié environ a tout de même un « Lieu-Dit » renseigné (120 établissements), qui sera utilisé comme libellé d'adresse. Seuls les 112 établissements restant (1,1 %) ne pourront pas être traités. Ils sont plus petits que l'ensemble des établissements, avec une médiane de capacité installée⁷ de 42 places, contre 71 pour l'ensemble des établissements Finess.

Par ailleurs, la base de passage d'un libellé d'adresse à une adresse brute (la base Fantoir) ne couvre ni Saint Barthelemy ni Saint Martin, ce qui conduit à supprimer 2 établissements supplémentaires de la base. 10 175 établissements sont ainsi à traiter.

Avant de procéder à l'appariement avec Fantoir, les libellés d'adresse Finess sont mis en forme. Le libellé est décomposé en *code nature de la voie + libellé de la voie* (par exemple « avenue du Maine » est décomposé en *code nature de la voie : « avenue » + libellé de la voie : « du Maine »*). Puis le code nature voie est modifié de sorte à le rendre similaire au code utilisé dans Fantoir (« avenue » devient « av »). Les textes sont ensuite nettoyés de tout accent ou ponctuation.

Résidences-services seniors (RSS)

Les RSS n'étant pas des établissements et services sociaux et médico-sociaux, elles ne sont pas répertoriées dans la base Finess. Afin d'obtenir une liste des adresses des RSS sur le territoire français, la DREES a procédé à un recensement web (effectué début 2022) pour construire deux bases de résidences services-seniors :

- La base « sanitaire et social », issue du site <https://www.sanitaire-social.com/>. Elle comporte 825 lignes et des informations sur l'adresse (l'adresse en toute lettre, le code postal, le nom de la commune) ;
- La base « cap résidence », issue du site <https://www.capresidenceseniors.com/>. Elle contient 1 079 lignes et des informations sur l'adresse (l'adresse en toute lettre, le code postal, le nom de la commune, les coordonnées x,y).

Encadré 1 • Les résidences-services seniors

Une résidence-service est « un ensemble d'habitations constitué de logements autonomes permettant aux occupants de bénéficier de services spécifiques non individualisables »¹ (services qui bénéficient par nature à l'ensemble des occupants, comme par exemple une conciergerie ou des espaces de convivialité). Plus spécifiquement, la loi du 28 décembre 2015 relative à l'adaptation de la société au vieillissement précise que « ces résidences-services s'adressent à des âgés autonomes, valides et semi-valides, de plus de 60 ans² qui désirent vivre en appartement ou en maison, tout en profitant de la convivialité et de la sécurité assurées par les équipes en place ».

Les RSS se distinguent des résidences autonomie, car ces dernières appartiennent à la catégorie des établissements et services sociaux et médico-sociaux (leur fonctionnement est régi par le Code de l'action sociale et des familles, le conseil départemental leur délivre l'autorisation de fonctionner et vérifie la qualité des prestations par des évaluations régulières).

⁶ Dans le répertoire Finess, les « catégories d'ET » à sélectionner sont : EHPAD (500), EHPAD perc crédit AM (501), EHPAD sans crédit AM (502), établissements expér. PA (381), CHR (101), CH (106, 355), établissements de soins longue durée (362) et résidences autonomie (202).

⁷ La capacité installée d'un établissement Finess est le nombre total de places offertes par l'établissement.

Les résidences-services seniors sont en pleine expansion en France : de 540 en 2016, leur nombre est estimé à 972 en 2022 (Kuhn Lafont, Troutot, 2022).

1. Article L631-13 du code de la construction et de l'habitation, création avec la loi du 28 décembre 2015.
2. En pratique, aucun âge minimal n'est requis pour intégrer une résidence-services.

Mise en forme des bases « sanitaire et social » et « cap résidence »

Les deux bases sont nettoyées :

- Retrait des accents, ponctuation, mise en majuscule pour l'adresse et le nom et retrait de quelques mots non porteurs de sens pour le nom ;
- Suppression des doublons, lorsque le code postal, le nom nettoyé et l'adresse nettoyée sont identiques (retrait de 4 lignes dans « sanitaire et social », aucune dans « cap résidence »).

Dans la base « cap résidence », 104 lignes avec une adresse nettoyée vide, et 4 sans commune ni code postal sont retirées.

De plus, dans la base « sanitaire et social », à chaque fois que des lignes partagent le même code postal et la même adresse nettoyée, les noms sont proches également mais l'un précise « résidence autonomie » et l'autre « résidence-service ». Nous supprimons la ligne qui précise « résidence autonomie ». Ceci supprime 13 lignes de la base « sanitaire et social ». Dans la base « cap résidence », nous supprimons 5 lignes qui ont été détectées comme doublon.

Nous ajoutons enfin un identifiant à chaque ligne, pour chaque base. « Sanitaire et social » compte 808 lignes, et « cap résidence » compte 966 lignes.

Ajout du code géographique

On concatène les deux bases pour obtenir une base RSS de 1 774 lignes. Tous les départements de cette base sont bien dans la base Fantoir. Seules 5 résidences sont dans les DROM.

Pour permettre l'appariement avec la base Fantoir, il faut introduire le code géographique Insee dans la base. On utilise une base de passage code postal / code géographique (*codgeo*) fournie par l'Insee pour associer les codes géographiques à chaque ligne (à chaque code postal), ainsi que le nom de commune selon l'Insee. À la suite de cet appariement basé sur le code postal, 18 lignes sur les 1 774 (1,0 %) ne sont associées à aucun code commune Insee (leur code postal n'est pas retrouvé dans la base de passage). Inversement, il peut y avoir plusieurs *codgeo* pour un établissement (car un même code postal peut correspondre à différents *codgeo*). Les noms de communes (RSS et Insee) sont nettoyés en retirant les accents, la ponctuation, et en recodant de manière uniforme les arrondissements. Pour 93,3 % des 1 774 établissements, une des lignes appariées est telle que les noms de commune sont exactement égaux. Pour ces établissements, on retire les autres lignes issues de l'appariement avec la base de passage. Puis on ajoute deux mesures de similarité textuelle entre libellés de commune (Damerau-Levenshtein et longueur cumulée des mots en commun, voir annexe 2 pour une explication des mesures). La similarité vaut 1 pour au moins une ligne pour 65 établissements parmi ceux qui n'avaient pas encore eu de correspondance exacte, ce qui correspond à 3,7 % des 1 774 établissements de départ. Enfin, 36 établissements ont un code postal retrouvé mais pas de similitude avec le nom de la commune (soit 2,0 % des 1 774 établissements). Ces établissements sont supprimés, ainsi que les 18 qui n'avaient pas de code commune Insee (soit 54 établissements au total), pour ne garder que les 1 720 (dont 4 en DROM) établissements avec un *codgeo* correctement apparié.

Repérage des doublons

Nous cherchons à repérer les couples d'identifiant correspondant à des établissements qui sont à la fois dans « sanitaire et social » et dans « cap résidence ». Pour cela nous repartons de la base concaténée, pour la couper en deux (la base d'origine « sanitaire et social », de 781 établissements, et la base d'origine « cap résidence », de 939). Puis les deux bases sont appariées en utilisant le code géographique de la commune : chaque établissement « sanitaire et social » est associé à tous les établissements « cap résidence » de la même commune. Nous mesurons ensuite la similitude de libellé entre les noms des résidences, et entre leurs adresses (distance Damerau-Levenshtein et distance longueur cumulée des mots en commun, voir en annexe 2 pour plus de détails). Pour chaque établissement « sanitaire et social », on note, le maximum de la similitude (pour l'adresse et le nom respectivement) parmi l'ensemble des établissements « cap résidence » de la même commune. Lorsque ces deux maximums sont supérieurs à 0,8, les établissements sont notés comme doublon. 47,8 % des établissements d'origine « sanitaire et social » ont un doublon dans « cap résidence », et 39,7 % des établissements d'origine « cap résidence » ont un doublon dans « sanitaire et social ». Les couples pour lesquels seul l'adresse ou le nom a une forte similitude sont regardés au cas par cas (46 couples avec nom proche et adresse différente, et 17 couples avec adresse proche et nom

différent), et ils sont également considérés comme doublons. Les couples de doublon (436 couples) sont pour l'instant traités comme les autres résidences-services et ils seront retirés après l'appariement avec Fantoir.

Fidéli

Nous utilisons ici Fidéli 2020, qui porte sur les revenus de l'année 2019. Fidéli se présente en 5 tables, dont 3 seront mobilisées dans cette étude :

- Individu : **une ligne pour chaque individu** (identifié par *id_num*), sauf pour les individus bilocalisés (il y a alors deux lignes). Les individus décédés en 2019 sont aussi présents. Chaque individu a une adresse brute (adresse sous forme de code). Pour chaque individu, des données individuelles telles que l'année de naissance, le sexe, la situation conjugale, ou la pension de retraite reçue sont disponibles ;
- Revenus ménage : **une ligne par logement** (*id_log*), apporte des informations complémentaires sur les revenus au niveau de l'ensemble des occupants du logement. Les revenus ne sont disponibles que pour les ménages qui habitent dans un logement assujéti à la taxe d'habitation
- Adresses hors TH : **une ligne par adresse** (identifiée par l'adresse brute), pour certaines adresses hors taxe d'habitation et/ou soupçonnées d'accueillir une communauté. La variable *id_scat* vaut 11 lorsqu'un EHPAD est à l'adresse.

Dans les tables Individu et adresses hors TH, l'adresse brute au 1^{er} janvier 2020 est codée au format Rivoli, avec les variables suivantes :

- Code commune : 5 chiffres (code commune Insee) ;
- Rue : 4 chiffres (identifiant de la voie) ;
- Noimm : 4 chiffres (numéro dans la rue). Il vaut 0000 lorsque le numéro est manquant ;
- Btq : 1 lettre (indicateur de répétition, « bis », « ter » ...).

Fantoir

Nous utilisons la base Fantoir de l'année 2020 pour faire le passage entre libellés d'adresse des tables Finess et RSS d'un côté et les codes Rivoli de Fidéli de l'autre (tableau 1). Il y a une ligne par voie. Chaque ligne comporte un code commune Insee, un identifiant de la voie (4 chiffres), un code de la nature de la voie (ex : rue, av...) et un libellé de la voie (ex : Général Leclerc).

Nous nettoyons les textes de tout accent ou ponctuation.

Tableau 1 • Exemple d'adresse dans les différentes bases

Finess ou RSS	Fidéli			Fantoir		
	Rue	Noimm	Btq	Id voie	Code nature	Libelle voie
Libellé adresse (avant re-traitement)						
3 bis avenue du Maine	2594	0003	B	2594	Av	Du Maine

■ MÉTHODE D'APPARIEMENT

Les bases des RSS et des établissements Finess sont appariées à la base Fantoir pour ajouter l'adresse brute (ou code Rivoli) de chaque établissement. Puis les deux bases ainsi constituées sont mises en cohérence entre elles, en retirant de la liste des RSS les établissements qui sont également dans Finess. Ensuite, d'une part la base RSS est appariée à la base des individus de Fidéli via l'adresse, et d'autre part la base des établissements Finess est appariée à la base des adresses brutes de Fidéli (base intermédiaire). Cette base intermédiaire permet de relier les établissements Finess à plus d'adresses et donc d'augmenter les chances d'appariement. Enfin, la base Finess-adresses de Fidéli est appariée à la base des individus de Fidéli, et une mise en cohérence des deux bases est à nouveau mise en œuvre. La méthode est synthétisée dans le schéma 1 en fin de partie.

Appariement des établissements et des résidences-services seniors avec Fantoir

Étapes de l'appariement

Les bases des établissements et résidences-services seniors sont appariées à la base Fantoir, l'une après l'autre, en utilisant le code commune Insee et le libellé d'adresse (sans le numéro).

En 1^{ère} étape, l'appariement se fait sur le libellé de voie complet (code commune Insee, code nature voie et libellé voie). Lorsqu'il y a des doublons, c'est que le libellé Fantoir est le même, mais avec différents identifiants de la voie. Ces doublons sont conservés. Les établissements pour lesquels une ligne correspondante a été trouvée dans Fantoir à cette première étape sont notés comme appariés de manière satisfaisante à Fantoir.

En 2^{ème} étape, les établissements qui ne sont pas appariés de manière satisfaisante à l'issue de la première étape sont traités. Pour un établissement **l'ensemble** des voies Fantoir de la même commune sont associés. Puis les textes sont nettoyés (transformation de chiffre en lettre, suppression de mots non porteurs de sens, uniformisation de certains groupes de mots : « docteur », « doc », « dc » en « dr » par exemple). Puis pour chaque couple de libellé de voie (libellé de l'établissement et libellé Fantoir dans la même commune), deux mesures de similarité textuelle sont calculées : Damerau-Levenshtein et longueur cumulée des mots communs (annexe 2 pour plus de détails). Ces mesures ont des valeurs comprises entre 0 et 1. Puis :

- Les établissements qui n'ont aucune ligne avec au moins une des deux similarités supérieure à 0,6 sont considérés en échec d'appariement ;
- Pour tous les autres, la ou les lignes qui correspondent au mieux au libellé de la voie doivent être sélectionnées, et les autres lignes supprimées. Nous commençons par supprimer toutes les lignes qui ont les deux similarités inférieures à 0,6, puis :
 - Lorsqu'un établissement a une ou plusieurs lignes avec une des similarités égale à 1, nous gardons ces lignes et supprimons les autres (ces établissements sont appariés de manière satisfaisante avec Fantoir) ;
 - Lorsqu'un établissement n'a pas de doublon, nous gardons la ligne unique. L'établissement est jugé apparié de manière satisfaisante ou non après une étude au cas par cas, qui sera différente dans Finess et dans RSS ;
 - Ensuite, pour les établissements qui présentent des doublons (et dont aucune ligne n'avait de similarité à 1), une mesure de similarité est ajoutée : longueur de la chaîne (liée à la longueur de la chaîne de caractère en commun la plus longue, voir en annexe 2 pour plus de détails). Si un établissement compte une ou plusieurs lignes avec une des 3 similarités supérieure à 0,8, nous gardons seulement ces lignes, et on note ces établissements comme appariés de manière satisfaisante ;
 - Les établissements qui n'ont toujours pas été traités (ils ne répondaient à aucun des critères cités jusqu'à présents) sont considéré comme appariés de manière non satisfaisante.

Au terme de l'appariement, il est possible qu'un même libellé d'adresse soit associé à plusieurs codes brutes différents. Certains établissements sont appariés à ces adresses brutes de manière satisfaisante et d'autres n'ont pas un appariement satisfaisant, soit car aucune adresse n'a été associée (échec d'appariement), soit car la similarité des adresses retrouvées est trop faible.

Résultats RSS

Appariement Fantoir

On procède à l'appariement avec Fantoir des 1 720 résidences-services seniors qui ont un code géographique de la commune correctement apparié (dont 4 en DROM).

En 1^{ère} étape, 1 282 des établissements trouvent une correspondance (dont deux établissements en DROM).

En 2^{ème} étape, sur les 438 RSS non appariées de façon satisfaisante à l'étape 1 :

- 17 établissements sont considérés comme en échec d'appariement ;
- 134 ont une ou plusieurs lignes avec une similarité égale à 1 ;
- 117 résidences ont une unique ligne associée, et après une étude au cas par cas, seuls deux d'entre eux sont en appariement non satisfaisant, les autres sont appariés de manière satisfaisante ;
- 150 résidences ont une ou plusieurs lignes avec une des 3 similarités supérieure à 0,8 ;
- 20 établissements⁸ n'ont pas trouvé d'adresse satisfaisante dans Fantoir.

Finalement, 1 681 résidences sont appariées de manière satisfaisante, soit 94,8 % des 1 774 établissements de départ. En métropole, elles sont 1 677 sur les 1 768 lignes de départ, soit 94,9 %.

La ou les adresses brutes de chaque résidence sont construites en ajoutant le numéro dans la rue au code brut de la voie retrouvée dans Fantoir. Lorsque le numéro renseigné dans la base RSS est de la forme « 3-7 », nous formons une ligne par numéro compris dans l'intervalle (par exemple 3, 5 et 7 ici).

Traitements supplémentaires

De nouveaux doublons d'établissements sont détectés lorsque plusieurs établissements partagent une même adresse brute. Sur les 1 640 adresses brutes de la base, 371 (22,6%) sont associées à deux établissements. 337 couples de résidences partagent une ou plusieurs adresses brutes, avec une résidence provenant de « sanitaire et social » et l'autre de « cap résidence ». Après une étude au cas par cas de ces couples, il s'avère qu'il s'agit en réalité de la même résidence. Nous concaténons enfin cette liste de 337 couples à celle des 436 doublons précédemment établie. Il y a 99 couples de la liste précédente qui ne sont pas dans cette liste (doublon de nom et/ou d'adresse, mais pas d'adresse brute), et aucun couple qui est dans cette liste mais pas dans la précédente (doublon d'adresse brute mais pas de nom et/ou d'adresse). La liste des doublons finale contient donc 436 couples⁹.

Nous pouvons finalement construire la base finale, avec une ligne par adresse brute pour chaque résidence. Si la résidence fait partie de la liste des doublons, les informations « sanitaire et social » et « cap résidence » sont renseignées en différentes colonnes. De plus, nous renseignons qu'elle est appariée de manière satisfaisante à Fantoir si elle l'était soit dans l'appariement entre son adresse « sanitaire et social » et Fantoir, soit dans l'appariement entre son adresse « cap résidence » et Fantoir. Si aucun de ces deux appariements n'était satisfaisant, la résidence est jugée appariée de manière non satisfaisante dans la base finale.

En enlevant les doublons à la base de 1 774 établissements du départ, on obtient au total une base de 1 338 établissements. **93,7 % de ces établissements (soit 1 253) sont appariés de manière satisfaisante¹⁰**. En métropole, ils sont 1 249 établissements sur 1 332 (soit 93,8 %).

35 établissements sont suspectés d'être des résidences autonomie (présence de « résidence autonomie » dans le nom), dont 33 appariés de manière satisfaisante avec Fantoir¹¹. 82,6 % des établissements appariés de manière satisfaisante ont une ligne unique, et il y a au maximum 14 lignes (pour deux établissements). Les 3 établissements qui ont plus de 10 lignes sont des établissements qui avaient une adresse avec un numéro du type « 3-19 ». Il n'y a pas d'adresse brute associée à différents établissements.

Ce chiffre de 1 338 RSS est plus élevé que les 972 RSS dénombrées par la publication du think tank *Matières Grises* (Kuhn Lafont, Troutot, 2022) en 2021. Cet écart peut s'expliquer car les sites internet recensent parfois sous le libellé de RSS des

⁸ Pour beaucoup d'établissements en échec d'appariement ou appariés de manière non satisfaisante, l'adresse est du type « rues exemple1, exemple2, ... » c'est-à-dire une liste de rues, par exemple « RUES DES PENSEES DES TULIPES DU JASMIN DU MUGUET DES PIVOINES ET DU HAINAUT ».

⁹ 54,0% des 808 établissements provenant de « sanitaire et social » partagent une adresse avec un établissement provenant de « cap résidence », et 45,1% des 966 établissements provenant de « cap résidence » partagent une adresse avec un établissement provenant de « sanitaire et social ».

¹⁰ Sur les 436 doublons, 428 étaient appariés de manière satisfaisante, et donc le nombre total d'établissement apparié de manière satisfaisante est 1 681-428 = 1 253.

¹¹ Les appariements suivants montreront que 27 de ces 33 établissements sont en effet des résidences autonomie.

établissements qui sont en réalité des résidences autonomie, ou répertorient des RSS dont l'ouverture est prévue au courant de l'année 2022.

Résultats Finess

L'appariement avec Fantoir est effectué sur les 10 175 établissements Finess avec une adresse hors Saint Barthelemy et Saint Martin (dont 0,85 % sont en DROM).

En 1^{ère} étape, 7 552 établissements trouvent une correspondance. Parmi les établissements qui ne trouvent pas de correspondance à cette première étape et qui seront traités en étape 2, il y a une sur-représentation des DROM (1,7 %, soit deux fois plus que dans l'ensemble des 10 175 établissements à appairer).

En 2^{ème} étape, sur les 2 623 établissements non appariés de façon satisfaisante à l'étape 1 :

- 122 établissements sont considérés en échec d'appariement ;
- 888 établissements ont une ou plusieurs lignes avec une similarité égale à 1 ;
- 637 établissements ont une unique ligne associée, et après une étude au cas par cas, nous décidons de classer ceux dont les deux similarités sont inférieures à 0,7 en appariement non satisfaisant (soit 42 établissements), les autres sont appariés de manière satisfaisante avec Fantoir ;
- 905 établissements ont une ou plusieurs lignes avec une des 3 similarités supérieure à 0,8. Ils sont appariés de manière satisfaisante.
- 71 établissements n'ont pas trouvé d'adresses satisfaisante dans Fantoir. Une analyse des établissements non appariés de manière satisfaisante est disponible en annexe.

Finalement, **9 940 établissements sont appariés de manière satisfaisante, soit 96,6 % des 10 289 établissements de départ**, et 96,7 % en métropole (9 858 sur 10 199). Parmi ces établissements au niveau métropole et DROM, 92,8 % ont une unique adresse Fantoir associée, et 10 ont plus de 10 lignes (jusqu'à 44 lignes).

La ou les adresses brutes de chaque établissement sont construites en ajoutant le numéro dans la rue au code brut de la voie retrouvée dans Fantoir. Dans Finess, ce numéro est toujours unique : il n'y a pas de numéros de la forme « 3-7 ».

Tableau 2 • Résultats d'appariement Finess-Fantoir

	Métropole et DROM	Métropole
Établissements Finess	10 289	10 199
<i>dont : sans adresse</i>	112	110
<i>dont : St Barthélémy ou St Martin</i>	2	0
<i>dont : échec d'appariement avec Fantoir</i>	122	122
<i>dont : appariement non satisfaisant</i>	113	109
% d'établissements Finess appariés à Fantoir de manière satisfaisante	96,6	96,7

Champ > Établissements pour personnes âgées, Finess 2020, France métropolitaine et DROM.

Source > Finess 2019, Fantoir 2020, calculs DREES.

La répartition des établissements Finess appariés de manière satisfaisante à Fantoir entre les différentes catégories d'établissement est similaire à la répartition des établissements de départ (voir tableau 7 pour le niveau métropole et DROM et tableau A11 en annexe pour la métropole).

Synthèse

Au sein du répertoire Finess, 96,6 % des établissements (soit 9 940 sur 10 289) sont appariés de manière satisfaisante, et 96,7 % en métropole (9 858 sur 10 199).

Dans les RSS, 93,7 % des établissements (soit 1 253 sur 1 338) sont appariés de manière satisfaisante, et 93,8 % en métropole (1 249 sur 1 332).

Tableau 3 • Synthèse des résultats d'appariement avec Fantoir

Etat d'appariement	Finess				RSS			
	Métropole et DROM		Métropole		Métropole et DROM		Métropole	
	Nombre	%	Nombre	%	Nombre	%	Nombre	%
Apparié en 1 ^{ère} étape	7 552	73,4	7 510	73,6	980	73,2	978	73,3
Apparié en 2 ^{ème} étape	2 388	23,2	2 348	23,0	273	20,4	271	20,4
Appariement non satisfaisant	349	3,4	341	3,3	85	6,4	83	6,2
Ensemble	10289	100	10199	100	1 338	100	1 332	100

Champ > Établissements pour personnes âgées, Finess 2020 et RSS issues du recensement web, France métropolitaine et DROM.

Source > Finess 2019, recensement des RSS « sanitaire et social » et « cap résidence » 2022, Fantoir 2020, calculs DREES.

Retrait des établissements Finess dans la base RSS

Il arrive que certaines RSS issues du recensement web aient un nom suggérant qu'il s'agit en réalité de résidences autonomie. Afin d'être sûr que nos résidences RSS n'incluent pas d'établissements sanitaires et sociaux, nous retirons de la base que nous venons de construire les RSS qui sont à une adresse brute identifiée dans Finess (suite à l'appariement Finess-Fantoir). La base RSS-Fantoir est appariée à la base Finess-Fantoir, en utilisant les adresses brutes. Nous ne traitons que les RSS et les Finess appariées de manière satisfaisante. Ainsi nous procédons à un appariement entre la base des 1 253 RSS appariées de manière satisfaisante, et les 9 940 établissements Finess appariés de manière satisfaisante.

Parmi ces 1 253 RSS, 181 sont appariées à Finess, soit **14,5 %** des RSS appariées de manière satisfaisante à Fantoir. Les 181 RSS appariées à Finess sont toutes en métropole :

- 172 sont associées à un unique établissement Finess ;
 - o 54 % à une résidence autonomie
 - o 43 % à un EHPAD
 - o 3 % à un EHPA
- 8 sont associés à deux établissements Finess (avec toujours un EHPAD) ;
- 1 est associée à trois établissements Finess (deux résidences autonomie et un EHPAD).

Il est donc relativement fréquent d'avoir une RSS à la même adresse brute qu'un établissement Finess. Cela peut s'expliquer dans deux situations :

- Il y a effectivement une RSS au même endroit qu'un établissement Finess (il est tout à fait possible qu'une RSS soit à la même adresse qu'un EHPAD, tout comme des résidences autonomie sont parfois à la même adresse qu'un EHPAD). Dans ce cas, il est impossible de déterminer dans quelle structure vivent effectivement les individus à ces adresses brutes dans Fidéli.
- La RSS et l'établissement Finess au même endroit sont en réalité le même établissement. Ceci semble probable lorsque l'établissement Finess associé est une résidence autonomie. Compte tenu du décalage des dates, il est possible que la RSS ait remplacé un ancien établissement Finess.

Afin d'éviter de prendre le risque d'inclure des RSS qui sont en réalité des établissements sanitaires, sociaux ou médico-sociaux, nous ne travaillons dans la suite qu'avec les RSS appariées de manière satisfaisante à Fantoir et qui n'ont pas été appariées à Finess, soit 1 072 RSS au niveau métropole et DROM, 1 068 au niveau métropole. Il faut noter que compte tenu du décalage de deux ans entre l'adresse du recensement web et l'adresse dans Finess (fin 2019 pour Finess, début 2022 pour le recensement web), il est possible que des établissements qui en 2019 étaient des résidences autonomie, mais sont devenues des RSS par la suite soient retirés. Inversement, les RSS qui n'étaient pas des résidences autonomie en 2019, mais le sont début 2022, sont conservées.

Appariement des établissements et résidences enrichis de Fantoir avec Fidéli

Nous avons d'une part la liste des établissements Finess appariés à Fantoir et d'autre part la liste des RSS appariées à Fantoir. Ces deux tables sont à appairer aux individus de Fidéli, vivant à la fin de l'année 2019¹². La stratégie ne sera pas la même avec ces deux bases.

Pour les résidences-services seniors, l'adresse brute issue de l'appariement RSS-Fantoir est recherchée dans la table des individus de Fidéli. Cela permet d'obtenir les individus qui vivent à l'adresse exacte de la RSS.

Pour les établissements Finess, deux approches sont associées. D'une part, les adresses déclarées dans la table des Individu de Fidéli sont analysées afin de déterminer si elles sont des adresses potentielles d'établissements pour personnes âgées ou non. Pour cela nous utilisons une variable disponible dans les tables Fidéli (*id scat*) qui indique si l'adresse est repérée comme susceptible d'être un EHPAD ou non. Elle est alors identifiée par le « critère Fidéli ». Un « critère d'accumulation » est également utilisé ; il identifie les adresses avec une concentration de seniors (ce critère est détaillé plus loin) Ce critère permet notamment d'identifier d'autres types d'établissements sanitaires que les EHPAD (en particulier les résidences autonomie). Ces deux critères identifient un très grand nombre d'adresses susceptibles d'être des établissements pour personnes âgées. D'autre part, les adresses brutes de Finess appariées à Fantoir sont recherchées dans la liste des adresses brutes de Fidéli. Ces deux approches se combinent comme suit : les adresses brutes de Finess sont recherchées dans la liste des adresses brutes de Fidéli, puis pour chaque établissement, nous regardons si des adresses proches dans Fidéli (dans la même rue notamment) sont identifiées par un des deux critères. Si oui, elles sont associées à l'établissement. Ceci permet de palier aux erreurs légères des adresses Finess, et surtout aux erreurs de déclaration des résidents. Par exemple, pour un établissement situé au 1 rue des Lilas, il est possible qu'une part non négligeable de ses résidents oublie de déclarer le numéro dans la rue (ils déclarent seulement *rue des Lilas*), leur numéro dans Fidéli sera donc 0000. À cette adresse brute (au 0000), le critère d'accumulation des seniors identifie un établissement (car ce sont principalement des seniors qui s'y sont déclarés). Lorsque nous recherchons l'établissement Finess dans Fidéli, nous recherchons d'abord l'adresse brute avec le numéro 1. Elle sera retrouvée en première étape. Puis nous allons regarder les adresses brutes proches, et le numéro 0000, identifié par le critère d'accumulation, sera retrouvé et associé à l'établissement. Avec cette approche, un établissement Finess est jugé apparié de manière sûre si au moins une des adresses qui lui a été associée dans Fidéli est identifiée par un des deux critères comme un établissement pour personnes âgées. Une fois l'appariement avec les adresses brutes de Fidéli terminé, les individus de Fidéli vivant à chaque établissement peuvent être associés. Le schéma 1 permet de visualiser l'ensemble de la méthode d'appariement.

Après les appariements avec la table des individus de Fidéli, une attention particulière doit être portée sur la situation de bilocalisation : un même individu (identifié par la variable *id num*) peut avoir deux lignes dans la table individu lorsqu'il est bilocalisé. Une de ses deux lignes voit alors sa variable *poids* nulle, et sa variable *biloc final* indique qu'il s'agit de sa seconde résidence. Nous faisons le choix de ne pas retirer les lignes avec un poids nul, car en pratique 2,6 % des résidents Finess et 3,2 % des résidents de RSS sont retrouvés grâce à la ligne correspondant à la seconde résidence. Les individus en doublet seront traités spécifiquement à la fin des appariements.

Établissements RSS

Nous n'adoptons pas une approche d'identification des adresses brutes Fidéli avec les critères Fidéli et/ou d'accumulation, utilisée pour appairer des adresses proches de l'adresse de départ, pour plusieurs raisons :

- Les tables Fidéli ne comportent pas de variable pouvant être utilisée pour identifier les résidences-services seniors. En effet la variable *id scat* ne permet de repérer que des EHPAD ;
- Le critère d'accumulation des seniors à une adresse n'est pas adapté car les seniors des RSS sont plus jeunes, et ces résidences sont parfois intergénérationnelles. Nous avons tenté de déterminer des seuils à adopter pour le critère d'accumulation des seniors sur le champ des RSS, les critères obtenus étaient trop peu discriminants ;
- Les annuaires web ont été recensés début 2022 et Fidéli porte sur les adresses au 1^{er} janvier 2020, ce qui peut poser problème. En effet, il se peut que nous allions chercher des individus qui vivent à un endroit où il n'y avait pas encore de RSS en 2019. Ce décalage se réduira avec une base Fidéli plus récente disponible.

Nous utilisons seulement l'adresse exacte des établissements RSS.

Nous appariions la base RSS-Fantoir, après retrait des établissements appariés à Finess (1 072 RSS), avec la table Individu de Fidéli, en utilisant l'adresse brute. 389 RSS sur les 1 072 (soit 36,3 %) n'ont pas d'adresse brute retrouvée dans Fidéli (il

¹² Nous faisons le choix de retirer les individus décédés au cours de l'année 2019, afin de travailler avec les seniors vivant au 31 décembre 2019. Dans les adresses appariées aux établissements Finess à la fin de la méthode, il y a 15% d'individus décédés au cours de l'année, et donc non pris en compte. Il est possible de reproduire la méthode construite ici mais en incluant ces individus.

n'y a aucun individu vivant dans la base Fidéli à s'être fiscalement déclaré à une des adresses brutes de la RSS). Il est possible qu'une partie de ces RSS n'existait pas encore en 2019 car selon l'étude du think tank *Matières Grises* (Kuhn Lafont, Troutot, 2022), autour de 200 RSS ont été créés entre 2019 et 2021. Il reste donc 683 RSS, parmi lesquelles 45 n'ont aucun individu de plus de 60 ans retrouvé dans Fidéli. **Ainsi, 638 RSS sont appariées de manière satisfaisante avec Fantoir, non identifiées comme établissement sanitaire et social, et avec au moins un individu de plus de 60 ans retrouvé dans Fidéli. Elles sont considérées appariées de manière sûre à Fidéli.** Il y en a 636 au niveau métropole. Ce sont ces RSS qui constitueront la base d'étude RSS dans la suite.

Ces 638 RSS regroupent **32 911 (32 866 en métropole) individus**, dont 117 sont présents deux fois (ils sont bilocalisés). Sur ces 117 individus, 109 sont à deux adresses brutes différentes (mais seulement 4 individus sur les 109 sont à deux RSS différentes). Lorsqu'un individu est bilocalisé, une seule ligne est conservée. Lorsque l'adresse est la même, nous retirons aléatoirement une des deux lignes. Lorsque l'adresse est différente mais correspond à la même RSS, la ligne conservée est celle associée à l'adresse avec le plus d'individus de plus de 60 ans. Lorsque les adresses sont différentes et associées à deux RSS différentes, la ligne conservée est celle associée à la RSS avec l'identifiant le plus petit (choix aléatoire).

Sur l'ensemble des individus résidant dans une RSS identifiée dans les bases fiscales, 16,7 % a strictement moins de 60 ans (16,6% en métropole), ce qui est possible du fait de l'absence de condition d'âge minimal pour intégrer une RSS.

Établissements Finess

L'appariement entre les établissements Finess et les individus de Fidéli passe par les adresses brutes présentes dans la table des individus de Fidéli. La base des adresses des individus de Fidéli est enrichie avec des informations indiquant si l'adresse est identifiée par le critère Fidéli et/ou d'accumulation. Ainsi, après l'appariement sur l'adresse exacte, les adresses candidates (identifiées comme établissement) qui peuvent également correspondre à un établissement Finess seront recherchées. Une fois que chaque établissement Finess est associé à une liste d'adresse brutes présentes dans Fidéli, la base des individus de Fidéli peut être appariée afin de retrouver l'ensemble des résidents.

Préparation des adresses de Fidéli

Une base avec une ligne par adresse brute présente dans la table Individu de Fidéli est constituée. Tout d'abord sont retirés de la base Individu les individus qui ne sont présents que l'année 2018, ainsi que les individus décédés à la fin de l'année 2019. Pour chaque adresse brute, nous renseignons le nombre d'individus présents dans la table, et également le nombre à être plus âgés que 60, 70, 80 ans. Lorsque nous comptons les individus par adresse, chaque identifiant individuel (*id num*) n'est compté qu'une fois (il arrive en effet dans de très rares cas qu'un même *id num* soit deux fois à la même adresse brute ; 0,02 % des lignes de la table Individu). En revanche, un même individu (bilocalisé) peut être compté simultanément dans plusieurs adresses brutes. Ceci n'est pas traité à cette étape.

Afin d'identifier les adresses Fidéli qui sont susceptibles d'être des établissements de type EHPAD, EHPA, résidence autonomie et USLD, la variable *id scat* présente dans la table « adresses hors TH » de Fidéli (une ligne par adresse brute) est utilisée. Cette variable vaut 11 si l'adresse est susceptible d'être un EHPAD selon Fidéli, et d'autres valeurs si l'adresse est susceptible d'être une communauté¹³. Lorsqu'*id scat* vaut 11, l'adresse est considérée comme identifiée comme un établissement par Fidéli (**critère Fidéli**). Malheureusement ce critère n'est pas suffisant pour correctement identifier les établissements Finess dans la base Fidéli, pour deux raisons :

- Parmi l'ensemble des établissements Finess recherchés, *id scat* n'identifie que les EHPAD (elle ne permet pas d'identifier les résidences autonomie, EHPA et USLD) ;
- *id scat* identifie parfois comme EHPAD des adresses qui ne sont pas réellement des EHPAD en soi, mais qui ont été déclarées par un individu qui vit en EHPAD. Ceci est possible car lors de la construction des tables Fidéli le libellé de complément d'adresse est utilisé (par exemple « EHPAD beau soleil, 3 avenue du Maine ») – non présent dans les tables accessibles par la DREES – pour déterminer si une adresse est un EHPAD ou non. Ainsi, un senior qui vit dans un EHPAD au 3 avenue du Maine peut se tromper et déclarer vivre au 2 avenue du Maine (erreur de déclaration), tout en déclarant vivre dans l'« EHPAD beau soleil ». La variable *id scat* identifie alors l'adresse au numéro 2 comme un EHPAD. Ainsi, toutes les adresses identifiées par *id scat* ne sont pas forcément les véritables adresses d'EHPAD.

¹³ Cette variable est construite dans Fidéli grâce au répertoire des communautés et au complément d'adresse, qui peut comporter le mot « EHPAD ».

Un autre critère construit indique si l'adresse est potentiellement un établissement pour personnes âgées. Il s'agit du « critère d'accumulation ». Il est défini de la manière suivante, pour chaque adresse : **plus de 10 individus de plus de 70 ans s'y sont déclarés, et 70 % au moins des individus de l'adresse ont plus de 70 ans.**

Les choix des seuils de ce critère sont expliqués en annexe. Ils ont été déterminés en fonction de la structure des établissements du type EHPAD, EHPA, USLD et résidences autonomie¹⁴. Ils visent donc à identifier les adresses qui pourraient être de ce type. On ne cherche pas à identifier des adresses qui pourraient être une RSS, car ces établissements n'ont pas les mêmes caractéristiques (il y a moins d'individus et la population est plus jeune).

Avec Fidéli 2020 (portant sur les revenus 2019) au niveau métropole et DROM, il y a 69 305 036 individus dans la table Individu (pour 70 284 366 lignes) et 17 233 501 adresses brutes. Le critère Fidéli identifie 17 163 adresses brutes comme EHPAD (variable *id scat*) et le critère d'accumulation identifie 12 421 adresses brutes comme établissement pour personnes âgées (dont 7 227 adresses brutes sont identifiées à la fois par le critère Fidéli et par le critère d'accumulation).

Le critère Fidéli reconnaît beaucoup plus d'adresses car lorsqu'un senior se trompe dans sa déclaration (il oublie par exemple de déclarer le numéro dans la rue), le critère Fidéli peut considérer qu'il s'agit tout de même d'un EHPAD grâce au libellé du complément d'adresse (dont nous ne disposons pas). Ainsi un même établissement peut avoir été détecté plusieurs fois si ses résidents ont fait des erreurs de déclaration dans l'adresse. Il n'y a cependant pas beaucoup d'individus à ces adresses issues d'erreurs de déclarations (à la fin des appariements, la moitié des adresses appariées à un des établissements Finess, identifiées par *id scat* et qui ne sont pas une traduction directe de l'adresse Finess ont moins de deux individus de plus de 60 ans). Aussi, parmi les adresses reconnues par le critère Fidéli et pas par le critère d'accumulation, 46,8 % (au niveau Métropole et DROM) ont un numéro dans la rue qui vaut 0000 (c'est-à-dire qu'il n'a pas été renseigné dans Fidéli), ce qui suggère bien que ces adresses sont en grande partie des erreurs de déclaration.

12,8 % des adresses reconnues par le critère d'accumulation et non par le critère Fidéli (variable *id scat* vide ou différente de 11) sont tout de même présentes dans la base des adresses hors TH, et donc reconnues par Fidéli comme des communautés (mais pas des EHPAD). Lorsqu'*id scat* est renseigné, elles sont principalement reconnues comme des « structures pour personnes nécessitant des soins médicaux » ou des « communauté religieuse ».

À l'échelle de l'individu, 300 925 individus sont identifiés à une adresse type établissement selon le critère Fidéli et le critère d'accumulation. Le critère Fidéli en compte 123 858 de plus, et le critère d'accumulation en compte 172 423 de plus. 3,4 % des plus de 70 ans sont dans un EHPAD selon le critère Fidéli, et 4,4 % sont dans un établissement pour personnes âgées selon le critère d'accumulation.

¹⁴ En pratique, les établissements de type EHPA et USLD ne seront que très rarement repérés par ce critères (voir annexe).

Tableau 4 • Adresses et individus identifiés à un établissement selon les critères Fidéli et d'accumulation

	Métropole et DROM	Métropole
Individus de Fidéli (sans doublon)	69 305 036	67 108 880
<i>Lignes dans la table Individu de Fidéli</i>	70 284 366	68 051 496
Adresses brutes	17 233 501	16 768 882
Adresses brutes identifiées par le critère Fidéli comme EHPAD	17 163	16 996
Adresses brutes identifiées par le critère d'accumulation comme établissement pour personnes âgées	12 421	12 374
<i>dont adresses brutes identifiées par le critère Fidéli ET par le critère d'accumulation</i>	7 227	7 193
Adresses brutes identifiées par le critère Fidéli et/ou le critère d'accumulation	22 357	22 177
Individus à une adresse identifiée par le critère Fidéli comme EHPAD	424 783	415 545
Individus à une adresse identifiée par le critère d'accumulation comme établissement pour personnes âgées	473 348	471 715
<i>dont individus à une adresse identifiée par le critère Fidéli et par le critère d'accumulation</i>	300 925	299 804
Individus à une adresse identifiée par le critère Fidéli et/ou le critère d'accumulation	597 142	587 393
% d'individus de plus de 70 ans en EHPAD selon le critère Fidéli	3,4	3,4
% d'individus de plus de 70 ans en établissement selon le critère d'accumulation	4,4	4,5

Champ > Individus Fidéli, vivant à la fin de l'année 2019, France métropolitaine et DROM.

Source > Fidéli 2020, calculs DREES.

Finalement, nous obtenons une base avec une ligne par adresse brute déclarée dans Fidéli. Pour chaque adresse, nous savons s'il s'agit d'un EHPAD selon Fidéli (avec *id scat*) et/ou d'un établissement pour personnes âgées selon le critère d'accumulation. Au niveau métropole et DROM, 22 357 adresses sont identifiées par le critère Fidéli et/ou d'accumulation comme un établissement pour personnes âgées sur les 17 233 501 adresses brutes.

Étape d'appariements

L'appariement des établissements Finess aux adresses brutes de Fidéli est conduit seulement pour les établissements appariés de façon satisfaisante avec Fantoir (9 940 sur les 10 289 de départ). Pour chacun de ces établissements (qui peut comporter plusieurs lignes suite à l'appariement avec Fantoir), nous appariions avec la table des adresses de Fidéli, qui contient une variable indiquant si l'adresse brute est un établissement selon le critère Fidéli et/ou d'accumulation. Ce critère est utilisé pour aller au-delà d'un appariement exact de l'adresse, et appairer également les adresses brutes candidates proches.

Les appariements se déroulent en plusieurs étapes. À chaque étape, les traitements se font pour tous les établissements Finess, de sorte à associer toutes les adresses brutes de Fidéli candidates. À la fin de ces appariements, un établissement Finess est apparié de manière sûre à Fidéli si au moins une adresse brute correspondant à un établissement **selon le critère Fidéli et/ou d'accumulation** lui a été associée.

En première étape, l'appariement utilise l'adresse brute totale (code commune Insee, identifiant de la voie, numéro dans la voie et indicateur de répétition « bis », « ter » ...). Après cette première étape :

- 85,9 % des établissements à traiter ont été associés à une adresse brute Fidéli identifiée comme établissement pour personnes âgées par le critère d'accumulation ou comme EHPAD par le critère Fidéli (85,9 % en métropole). Plus précisément :
 - 73,5 % ont été associés à au moins une adresse brute Fidéli identifiée comme établissement pour personnes âgées par le critère d'accumulation (73,9 % en métropole).
 - 12,4 % n'ont pas été associés à une adresse brute Fidéli identifiée par le critère d'accumulation, mais ont été associés à au moins une adresse brute Fidéli identifiée par le critère Fidéli (12,0 % en métropole).
- 9,5 % ont une adresse qui a été retrouvée dans Fidéli, mais ils n'ont pas été associés à une adresse considérée comme un établissement par le critère d'accumulation ou par le critère Fidéli (9,4 % en métropole).
- 4,6 % n'a pas du tout été retrouvé dans Fidéli (4,6 % également en métropole).

Les établissements qui sont appariés avec Fidéli mais pas avec une adresse considérée comme un établissement selon le critère d'accumulation (21,9 % en métropole et DROM, 21,5 % en métropole) sont un peu plus petits (médiane des capacités installées de 56 contre 72 pour l'ensemble des établissements Finess appariés à Fantoir de manière satisfaisante) : il est donc

possible que les adresses brutes correspondantes n'aient pas été repérées par le critère d'accumulation, car trop peu de résidents s'y sont déclarés fiscalement.

En seconde étape, l'appariement se base sur l'adresse brute mais en retirant l'indice de répétition (bis, ter). Les nouvelles adresses ainsi appariées ne sont conservées que si elles correspondent à un établissement selon le critère Fidéli et/ou d'accumulation.

En troisième étape, l'appariement se base sur l'adresse brute mais en retirant le numéro dans la voie. Les informations sur lesquelles se basent l'appariement sont donc le code commune Insee et l'identifiant de la voie. Pour chaque établissement Finess, nous ne gardons que les lignes d'adresse brutes Fidéli considérées comme établissement par le critère Fidéli et/ou d'accumulation et qui sont proches de l'établissement Finess. Deux adresses sont proches lorsque leur distance (x,y) est inférieure à 100 mètres¹⁵ (200 mètres dans les DROM) ou alors lorsqu'un des deux numéros des adresses est nul (non renseigné). En effet, il est possible qu'un résident d'établissement n'ait pas déclaré de numéro dans sa déclaration fiscale, ainsi son (x, y) est situé au numéro 0 de la rue, qui peut être à plus de 100 mètres du véritable numéro de son établissement.

En quatrième et dernière étape, l'appariement se base sur le code de la commune uniquement. Pour chaque établissement Finess, nous ne gardons que les lignes d'adresses brutes Fidéli considérées comme établissement par le critère Fidéli et/ou d'accumulation et qui sont à moins de 150 mètres de l'établissement Finess (1000 mètres pour les DROM).

Réduction des doublons

Au terme de ces appariements, nous obtenons une base qui a pour clé le couple adresse brute + numéro Finess (une ligne unique pour chaque couple).

Il y a cependant, par établissement Finess, différentes adresses brutes associées. Au niveau métropole et DROM, un même établissement Finess a jusqu'à 44 adresses brutes. Parmi les adresses brutes associées à un établissement Finess :

- Certaines n'ont pas été retrouvées du tout dans Fidéli ;
- Certaines sont présentes dans Fidéli mais ne correspondent pas à un établissement (selon les deux critères) ;
- Certaines sont présentes dans Fidéli et correspondent à un établissement selon le critère Fidéli mais pas selon le critère d'accumulation ;
- Certaines sont dans Fidéli et correspondent à un établissement selon le critère d'accumulation.

Idéalement, il faudrait avoir pour chaque établissement Finess un et un seul établissement identifié par le critère d'accumulation. Il arrive cependant que pour un même établissement Finess plusieurs adresses différentes correspondent à un établissement selon le critère d'accumulation. Lorsque cela est le cas, les adresses repérées par le critère d'accumulation qui ont été ajoutées en étape 4 (lorsque la recherche se faisait à proximité dans la commune) sont retirées. Aussi, lorsque la rue est la même, si deux adresses brutes identifiées par le critère d'accumulation sont appariées au même établissement Finess, mais que ces adresses ont un numéro dans la rue renseigné (différent de 0) et différent, n'est conservée que l'adresse qui a été appariée en première.

Il arrive aussi que pour une même adresse brute, différents établissements Finess soient associés. Nous dirons qu'ils sont en conflit d'intérêt d'adresse. En général, ceci s'explique car les différents établissements Finess appariés à la même adresse brute avaient le même libellé d'adresse dans Finess¹⁶. Les établissements Finess ont donc été appariés aux mêmes adresses brutes, aux mêmes étapes. Ces situations sont conservées. En revanche, lorsque l'étape des appariements n'est pas la même, seul l'appariement qui a eu lieu en premier est conservé (il est plus probable que l'adresse brute lui corresponde). Pour les cas d'adresses brutes associées à différents établissements Finess restantes, la similarité Damerau-Levenshtein entre les adresses Finess des différents établissements est calculée. Si la similarité est supérieure à 0,6, les doublons d'établissements sont conservés. Si la similarité est inférieure à 0,6, un dernier nettoyage est effectué : si les établissements Finess concernés avaient déjà été appariés avec des adresses considérées comme établissement selon le critère d'accumulation ou le critère Fidéli dans une étape antérieure, nous supprimons les lignes problématiques.

¹⁵ Les données Fidéli et les données Finess comportent toutes deux des coordonnées (x, y) au format Lambert 93 en métropole et UTM dans les DROM. Cependant, les coordonnées (x,y) ne sont pas tout à fait robustes : pour des adresses considérées comme appariées sans ambiguïté entre Finess et Fidéli, la médiane de la distance géographique entre les adresses des deux bases est de 52 mètres.

¹⁶ Par exemple la résidence autonomie Saint André se situe au 4^{ème} étage de la maison Sainte Marie (un EHPAD) au 35 avenue Peboue à Pau (64000). Ces deux établissements Finess partagent exactement le même libellé d'adresse.

Appariements sûrs

À l'issue de ces appariements et suppressions de doublon, un établissement Finess est apparié de manière sûre à Fidéli si au moins une adresse brute correspondant à un établissement **selon le critère Fidéli et/ou d'accumulation** lui a été associée.

Au niveau métropole et DROM, sur les 9 940 établissements Finess à traiter (ceux appariés de manière satisfaisante à Fantoir), 924 ne sont pas appariés de manière sûre à Fidéli, soit 9,3 %. Au niveau métropole, 913 sur 9 858 ne sont pas appariés de manière sûre à Fidéli, soit 9,3 %. Sur les 9 016 établissements appariés de manière sûre à Fidéli, 3 n'ont que des individus âgés de moins de 60 ans (ce sont deux EHPAD privés à but non lucratif et un EHPA, avec moins de 3 individus retrouvés dans Fidéli, et identifiés par Fidéli seulement). Nous ne les considérons finalement pas comme des appariements sûrs à Fidéli. Les 9 013 autres ont au moins un senior de plus de 60 ans. Ils représentent 90,7 % des établissements appariés de manière satisfaisante à Fantoir. Parmi eux, 1 016 ne sont pas associés à une adresse brute identifiée comme établissement par le critère d'accumulation.

Ainsi, sur l'ensemble des établissements Finess, 96,6 % sont appariés à Fantoir de manière satisfaisante et peuvent être cherchés dans Fidéli. Parmi eux, 90,7 % sont appariés avec au moins une adresse brute de Fidéli qui correspond à un établissement selon le critère Fidéli ou le critère d'accumulation (et au moins un senior de plus de 60 ans). Ils sont appariés de manière sûre à Fidéli. **Finalement, 87,6 % des établissements Finess de départ sont appariés de manière sûre à Fidéli (87,7 % pour la métropole).** Parmi ces établissements, 88,7 % ont au moins une adresse reconnue par le critère d'accumulation (89,1 % en métropole).

Les résultats sont satisfaisant du point de vue des répétitions d'adresse. Sur l'ensemble des établissements Finess appariés de manière sûre à Fidéli, 95,2 % ont moins de trois adresses retrouvées dans Fidéli et identifiées comme établissement par le critère Fidéli et/ou d'accumulation. Ils ont tous moins de deux adresses brutes différentes identifiées par le critère d'accumulation. On compte 411 établissements (tous en métropole) avec deux adresses brutes différentes correspondant à un établissement selon le critère d'accumulation. Ces situations paraissent crédibles car à chaque fois une de ces deux adresses a un numéro manquant (il est possible que plusieurs seniors aient oublié de déclarer le numéro dans la rue).

De plus, 3 adresses brutes sont associées à 4 établissements Finess différents (ils ont sensiblement la même adresse dans Finess), 13 adresses brutes sont associées à 3 établissements Finess différents, et 636 adresses brutes (628 en métropole) sont associées à deux établissements Finess différents. Les établissements Finess qui partagent une même adresse brute sont désignés comme étant en conflit d'intérêt sur une adresse, ils devront être mis de côté pour la production de certaines statistiques.

Tableau 5 • Résultats d'appariement Finess-Fidéli

	Métropole et DROM	Métropole
Établissements Finess appariés de manière sûre (au moins une adresse identifiée comme établissement par le critère Fidéli et/ou d'accumulation, et au moins un senior de plus de 60 ans)	9013	8942
<i>dont : non associés à une adresse brute identifiée comme établissement par le critère d'accumulation</i>	1016	974
% d'établissements Finess appariés de manière sûre, par rapport à tous les établissements Finess de départ	87,6	87,7
% d'établissements Finess appariés de manière sûre, par rapport aux établissements Finess appariés à Fantoir de manière satisfaisante	90,7	90,7
% d'établissements Finess appariés de manière sûre avec au moins un établissement identifié par le critère d'accumulation, par rapport aux établissements Finess appariés de manière sûre	88,7	89,1

Champ > Établissements Finess appariés de manière sûre à Fidéli, France métropolitaine et DROM.

Source > Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Appariement aux individus et traitements supplémentaires

La base des adresses brutes de chaque établissement Finess est appariée à la table des individus de Fidéli, afin de retrouver la population des résidents d'établissement. Puis des derniers nettoyages sont mis en œuvre.

Mise en cohérence des deux bases

15 adresses brutes sont associées à la fois à un établissement Finess apparié de manière sûre à Fidéli et à une RSS appariée de manière sûre à Fidéli. Ceci est possible malgré l'étape de retrait des RSS qui étaient en réalité des établissements Finess, car à cette étape nous avons seulement considéré les adresses associées à Finess à l'issue de l'appariement Finess-Fantoir (donc les adresses issues d'une traduction directe du libellé d'adresse Finess). N'étaient donc pas incluses dans cette étape les adresses issues de l'appariement avec les adresses proches. Ces 15 adresses présentes dans les deux bases comportent

662 individus, qui se répartissent dans 15 établissements Finess. Ces adresses sont retirées de la liste des adresses associées aux établissements Finess. Pour 3 de ces 15 établissements Finess, l'adresse ainsi retirée était la seule identifiée par le critère Fidéli et/ou d'accumulation. Ainsi, 3 établissements Finess ne sont plus considérés comme appariés de manière sûre à Fidéli. Finalement 9 010 établissements Finess sont appariés de manière sûre à Fidéli (au moins une adresse identifiée comme établissement par le critère Fidéli et/ou d'accumulation, et au moins un individu de plus de 60 ans), et non plus 9 013. En métropole, 8 939 établissements sont appariés de manière sûre à Fidéli. 88,7 % de ces établissements ont au moins un établissement repéré par le critère d'accumulation (89,0 % en métropole).

Identification des établissements en conflit d'intérêt sur la même adresse

Sur l'ensemble des adresses brutes associées aux établissements Finess, un certain nombre sont associées à plusieurs établissements Finess différents. Ceci s'explique lorsque le libellé d'adresse dans Finess est très proche car les établissements sont en réalité au même endroit et ils ne peuvent pas être différenciés (cf. exemple de la note de bas de page précédente – n.16). Au total, 747 établissements Finess sont en conflit d'intérêt sur une de leurs adresses brutes, soit 8,3 % des 9 010 établissements appariés de manière sûre. Ces établissements regroupent 6,9 % des individus vivant dans les établissements Finess appariés de manière sûre. En prenant en compte l'ensemble des établissements appariés de manière sûre, on compte 437 292 individus, dont 394 820 ont plus de 60 ans (90,3 %). Si on compte seulement les établissements hors conflit d'intérêt, il y a 406 937 individus, dont 368 325 de plus de 60 ans (90,5 %). Il faudra être très vigilant aux établissements en conflit d'intérêt lors de la réalisation des statistiques descriptives, car les adresses étant dupliquées, les individus qui y vivent le sont également. La répartition dans les différentes catégories d'établissement est très proche, selon que l'on considère l'ensemble des établissements ou seulement ceux qui ne sont pas en conflit d'intérêt sur une adresse (tableaux A9 et A10 en annexe).

Tableau 6 • Établissements en conflit d'intérêt et individus dans ces adresses

	Métropole et DROM	Métropole
Établissements Finess appariés de manière sûre	9 010	8 939
<i>dont : en conflit d'intérêt sur une adresse</i>	747	739
<i>dont : hors conflit d'intérêt sur une adresse</i>	8 263	8 200
Individus dans les établissements appariés de manière sûre	437 292	428 769
<i>dont plus de 60 ans (%)</i>	90,3	91,3
Individus hors établissements en conflit d'intérêt d'adresse	406 937	399 861
<i>dont plus de 60 ans (%)</i>	90,5	91,4

Champ > Individus dans les établissements Finess appariés de manière sûre à Fidéli, France métropolitaine et DROM.
Source > Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Élimination des doublons d'individus

Certains individus (identifié grâce à la variable *id num* de Fidéli) sont présents plusieurs fois dans la base des résidents. Ceci est en partie une conséquence directe des doublons d'adresses brutes vu précédemment (les individus des adresses présentes plusieurs fois sont eux aussi présents plusieurs fois). En revanche il arrive que des individus soient en doublon, même lorsque nous nous restreignons aux établissements qui ne sont pas en conflit d'intérêt sur une adresse. Ainsi, sur les 406 937 individus de ces établissements, 497 (soit 0,12 %) sont présents deux fois. À chaque fois, ce sont des individus bilocalisés. 96,2 % de ces individus sont à deux adresses différentes, mais seulement 27,4 % (136) sont à deux établissements Finess différents (dont 9 dans deux départements différents). Il y a donc 342 de ces individus (83,8 %) qui sont à deux adresses différentes, mais associées au même établissement.

Pour ces individus bilocalisés, une seule ligne est conservée. Lorsque l'adresse est la même, nous retirons aléatoirement une des deux lignes. Lorsque l'adresse est différente mais correspond au même établissement Finess, la ligne conservée est celle associée à l'adresse avec le plus d'individus de plus de 60 ans. Lorsque les adresses sont différentes et associées à deux établissements Finess différents, la ligne conservée est celle associée à l'établissement avec la fin du numéro Finess (4 derniers chiffres) le plus petit (choix aléatoire).

Résultats des appariements

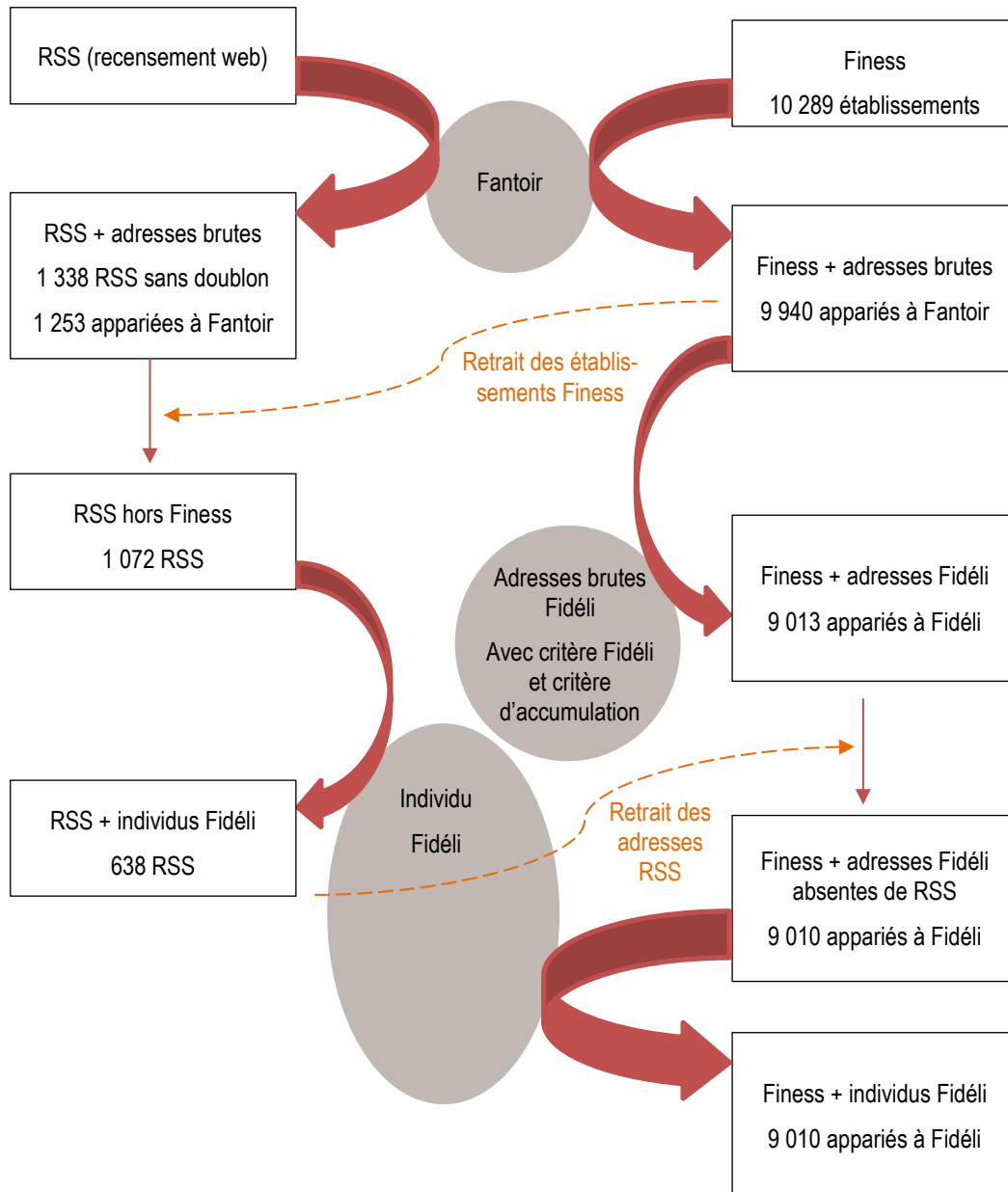
Résultats généraux

Sur l'ensemble des 17 233 501 adresses brutes présentes dans Fidéli au niveau métropole et DROM, 15 833 sont associées de manière sûre à un établissement Finess et 850 sont associées de manière sûre à une RSS.

En termes d'individus, Fidéli compte 69 305 036 individus, dont 17 728 620 ont plus de 60 ans (soit 25,6 %), et 9 691 710 ont plus de 70 ans (soit 14,0 %). Dans les établissements Finess, 437 292 individus sont retrouvés, dont 394 820 ont plus de 60 ans (soit 90,3 %), et 362 487 ont plus de 70 ans (soit 82,9 %). Dans les RSS, 32 911 individus sont retrouvés, dont 27 410 ont plus de 60 ans (soit 83,3 %), et 25 552 ont plus de 70 ans (soit 77,6 %). Ainsi nous identifions 2,2 % des individus de plus de 60 ans dans des établissements de type Finess, et 0,2 % des individus de plus de 60 ans dans des RSS. De la même manière, nous identifions 3,7 % des individus de plus de 70 ans dans des établissements de type Finess, et 0,3 % des individus de plus de 70 ans dans des RSS.

Sur les 10 289 établissements Finess de départ, 9 940 sont appariés de manière satisfaisante avec Fantoir (soit 96,6 %), et 9 010 sont appariés de manière sûre avec Fidéli (soit 87,6 %). Pour les RSS, l'appariement avec Fantoir et la mise au propre de la base conduisent à une table de 1 338 RSS. Parmi elles, 1 253 sont appariées de manière satisfaisante avec Fantoir (soit 93,6 %), dont 1 072 ne sont pas assimilées à des établissements Finess (soit 80,1 % des 1 338 RSS de départ), et 638 sont présentes dans Fidéli avec au moins un individu de plus de 60 ans, c'est-à-dire appariées à Fidéli de manière sûre (soit 47,7 % des 1 338 RSS de départ).

Schéma 1 • Schéma récapitulatif de la méthode d'appariement



Scores d'appariement Finess par catégorie d'établissement

La répartition des établissements dans les différentes catégories d'établissement est similaire avant et après l'appariement avec Fantoir (seulement 3,4 % des établissements sont perdus à cette étape, ce qui déforme très peu la répartition). Après l'appariement avec Fidéli, les EHPAD sont sur-représentés parmi les établissements correctement appariés, et les résidences autonomie sont sous-représentées. Les EHPA et les USLD sont les établissements que l'on retrouve le moins à l'issue de l'appariement avec Fidéli. Les EHPA représentent par exemple 3,1 % du champ des établissements Finess de départ, mais seulement 2,1 % du champ des établissements appariés de manière sûre (tableau 7 pour le niveau métropole et DROM, tableau A11 en annexe pour la métropole).

Tableau 7 • Répartition dans les catégories d'établissement au départ, après appariement avec Fantoir, et après appariement avec Fidéli, France métropolitaine et DROM

Catégorie	Nombre départ	Répartition départ (%)	Nombre appariés avec Fantoir	Répartition appariés avec Fantoir (%)	Nombre appariés avec Fidéli	Répartition appariés avec Fidéli (%)
EHPAD privés à but lucratif	1 817	17,7	1 767	17,8	1 733	19,2
EHPAD privés à but non lucratif	2 359	22,9	2 275	22,9	2 185	24,3
EHPAD publics hospitaliers	1 210	11,8	1 160	11,7	1 091	12,1
EHPAD publics non hospitaliers	2 133	20,7	2 066	20,8	2 033	22,6
Résidences autonomie	2 290	22,3	2 225	22,4	1 700	18,9
EHPA	314	3,1	287	2,9	192	2,1
USLD	166	1,6	160	1,6	76	0,8
Ensemble	10 289	100,0	9 940	100,0	9 010	100,0

Champ > Établissements Finess de départ (établissements pour personnes âgées), établissements Finess appariés de manière satisfaisante avec Fantoir et établissements Finess appariés de manière sûre avec Fidéli, France métropolitaine et DROM.

Source > Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Il y a donc des catégories d'établissements qui sont mieux appariées que d'autre, comme l'indique le tableau 8 (le tableau A12 pour la métropole est disponible en annexe). Les résultats sont moins satisfaisants pour les USLD et EHPA, ce qui peut s'expliquer par différents facteurs. Pour les EHPA, une part plus importante des seniors est en hébergement temporaire ou en accueil de jour qu'au sein des EHPAD et résidences autonomie, et donc une part plus importante ne s'y déclare pas fiscalement. En effet dans les EHPA, seulement 83 % des résidents sont en hébergement permanent, contre 96 % au sein des EHPAD et 99 % dans les résidences autonomie¹⁷. Dans les USLD, bien que 99 % des résidents soient en hébergement permanent, il semble qu'ils se déclarent fiscalement moins fréquemment à leur établissement. Ceci peut en partie s'expliquer par une durée de séjour légèrement plus faible¹⁸ que dans les autres types d'établissements, et donc une plus faible probabilité de se déclarer fiscalement à l'adresse de l'établissement. En revanche pour les établissements pour personnes âgées qui ne sont pas de ces deux catégories (soit 95,3% des établissements pour personnes âgées), les résultats sont satisfaisants. 93,7 % des EHPAD sont appariés de manière sûre à Fidéli (avec un score supérieur à 90 % dans toutes les catégories d'EHPAD), et 74,2 % des résidences autonomie le sont.

Tableau 8 • Scores d'appariement par catégorie d'établissement

Catégorie	Nombre	Répartition (%)	Part à être appariée de manière sûre (%)
EHPAD privés à but lucratif	1 817	17,7	95,4
EHPAD privés à but non lucratif	2 359	22,9	92,6
EHPAD publics hospitaliers	1 210	11,8	90,2
EHPAD publics non hospitaliers	2 133	20,7	95,3
Résidences autonomie	2 290	22,3	74,2
EHPA	314	3,1	61,1
USLD	166	1,6	45,8
Ensemble	10 289	100	87,6

Champ > Établissements pour personnes âgées dans Finess, France métropolitaine et DROM.

Source > Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Pour chaque catégorie d'établissement le tableau 9 détaille la part à ne pas avoir été retrouvée à chaque étape (appariement avec Fantoir, puis appariement avec Fidéli) et la part à avoir été appariée grâce à chacun des critères (critère Fidéli et/ou d'accumulation). Le tableau A13 pour la métropole est disponible en annexe. Par exemple, pour les EHPAD publics non hospitaliers, 0,6 % est apparié uniquement grâce au critère d'accumulation, contre 59,9 % des résidences autonomie. Les EHPA sont ceux qui sont le plus perdus à la fois dans l'appariement avec Fantoir, et dans l'appariement avec Fidéli. Les USLD

¹⁷ Enquête EHPA 2019, data.drees.

¹⁸ Au 31 décembre 2019, 30% des résidents d'USLD sont dans leur établissement depuis moins de 10 mois, contre 1 an et 2 mois dans les EHPAD, enquête EHPA 2019, DREES.

sont un peu mieux retrouvés dans Fantoir et dans Fidéli, mais les adresses Fidéli auxquelles ils sont associés ne sont pas identifiées comme des adresses d'établissement dans 40,4 % des cas.

Tableau 9 • Étapes d'appariement par catégorie d'établissement

Catégorie	Non apparié avec Fantoir	Pas dans Fidéli	Aucun établissement	Accumulation et Fidéli	Accumulation seulement	Fidéli seulement	Total
EHPAD privés à but lucratif	2,8	0,9	0,9	86,1	0,9	8,4	100
EHPAD privés à but non lucratif	3,6	1,8	1,9	76,0	2,7	14,0	100
EHPAD publics hospitaliers	4,1	1,5	4,2	68,4	7,4	14,4	100
EHPAD publics non hospitaliers	3,1	0,8	0,8	84,4	0,6	10,3	100
Résidences autonomie	2,8	4,9	18,0	11,9	59,9	2,4	100
EHPA	8,6	12,1	17,8	28,3	8,9	24,2	100
USLD	3,6	10,2	40,4	20,5	15,1	10,2	100
Ensemble	3,4	2,5	6,5	62,0	15,6	9,9	100

Note > Ne sont pas retirés les 3 établissements qui avaient tous leurs individus plus jeunes que 60 ans, et les 3 établissements dont l'appariement n'est plus sûr suite au retrait des adresses en conflit avec la base RSS.

Lecture > Parmi les 2 290 résidences autonomie dans Finess, 2,8 % ont un appariement avec Fantoir non satisfaisant, 4,9 % ne sont pas retrouvées dans Fidéli du tout (pas de correspondance d'adresse brute), 18,0 % sont retrouvées dans Fidéli mais aucune adresse ne correspond à un établissement, ni selon le critère Fidéli (*id scat*), ni selon le critère d'accumulation. 11,9 % ont au moins une adresse retrouvée dans Fidéli qui avait été identifiée à la fois par le critère Fidéli et par le critère d'accumulation. 59,9 % ont au moins une adresse retrouvée identifiée par le critère d'accumulation, et aucune adresse identifiée par le critère Fidéli. 2,4 % ont au moins une adresse retrouvée identifiée par le critère Fidéli, et aucune adresse identifiée par le critère d'accumulation.

Champ > Établissements pour personnes âgées dans Finess, France métropolitaine et DROM.

Source > Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Les établissements qui ont été appariés de manière satisfaisante avec Fidéli mais dont aucune adresse n'a été reconnue par le critère d'accumulation (9,9 % des établissements) ont une médiane de capacité installée de 56, contre 74 pour l'ensemble des établissements Finess appariés. Leur plus faible capacité peut expliquer le fait qu'ils n'ont pas été repérés par le critère d'accumulation.

Aussi, 2,4 % des EHPAD ont été appariés de manière sûre à Fidéli, mais ils n'ont pas du tout été identifiés comme EHPAD par la variable *id scat* (critère Fidéli), soit 183 établissements. 171 de ces EHPAD ont été appariés à Finess dès la première étape d'appariement (avant de procéder à l'appariement des adresses proches), c'est-à-dire que l'adresse brute Fantoir de leur libellé d'adresse dans Finess était reconnue par le critère d'accumulation. Les 12 EHPAD restant ont été obtenus grâce aux appariements d'adresses proches et à l'identification du critère d'accumulation. Parmi ces 183 EHPAD, 54,1 % sont identifiés par *id scat* comme une structure pour personnes nécessitant des soins médicaux, 13,1 % comme une communauté religieuse, et 11,5 % sont dans la base des communautés mais sans précision de type de communauté.

Adresses identifiées par un des critères

Les 22 357 adresses brutes de Fidéli identifiées par l'un des deux critères n'ont pas toutes été appariées à un établissement Finess ou à une RSS. La probabilité pour une adresse d'être appariée à un établissement Finess dépend de si elle est identifiée par le critère d'accumulation, Fidéli ou les deux. 87,4 % des adresses brutes identifiées par le critère Fidéli et par le critère d'accumulation sont présentes dans Finess de manière sûre (c'est à dire qu'elles ont été associées à un établissement Finess de manière sûre). En revanche, cette proportion tombe à 64,6 % pour les adresses brutes seulement identifiées par le critère Fidéli. Cela peut s'expliquer car parmi les adresses brutes identifiées par le critère Fidéli, beaucoup sont des erreurs de déclaration des résidents (qui ne sont donc pas appariées avec de véritables adresses Finess). Cependant ces adresses comptent peu d'individus. En effet, la moitié des adresses brutes identifiées par Fidéli seulement, associées à un établissement Finess de manière sûre, mais pas en première étape (l'adresse est proche de l'adresse exacte) compte moins de deux seniors de plus de 60 ans (trois si l'on compte l'ensemble des individus).

De plus, sur l'ensemble des adresses brutes seulement identifiées par le critère d'accumulation, 34,6 % sont appariées avec un établissement Finess. Ces adresses correspondent donc à des établissements appariés à Finess, mais qui n'avaient pas du tout été repérés par Fidéli (colonne « Accumulation seulement » du tableau 9 précédent). 65,4 % des adresses brutes seulement identifiées par le critère d'accumulation ne sont donc pas appariées à Finess, soit 3 399 adresses. Parmi elles, 394 (soit 11,6 %) sont présentes dans RSS, et ont donc été associées à une résidence-service seniors. Ces 394 adresses présentes dans RSS et identifiées par le critère d'accumulation représentent 46,4 % des 850 adresses brutes présentes dans RSS. Au total sur les 638 RSS, 60,5 % ont au moins une adresse identifiée par le critère d'accumulation des seniors.

À l'échelle des individus :

- 91,1 % des individus de plus de 60 ans qui vivent à une adresse identifiée par les deux critères sont présents dans un établissement Finess ;
- 68,3 % des individus de plus de 60 ans qui vivent à une adresse identifiée par Fidéli seulement sont présents dans un établissement Finess.
- 47,7 % des individus de plus de 60 ans qui vivent à une adresse identifiée par le critère d'accumulation seulement sont présents dans Finess, et 13,9 % sont présents dans RSS.

Aussi, 26,6 % des individus de plus de 60 ans et vivant à une adresse identifiée par le critère d'accumulation et absente de Finess sont dans une RSS. Enfin, 68,7 % des individus en RSS sont à une adresse identifiée par le critère d'accumulation et absente de Finess.

Le fait que les parts d'individus soient plus élevées que les parts d'adresse suggère que ce sont les adresses avec le plus d'individus qui ont été retrouvées dans Finess et dans RSS.

Tableau 10 • Adresses brutes identifiées par un des critères, et individus qui y sont déclarés

	Métropole et DROM	Métropole
% des adresses brutes identifiées par le critère Fidéli ET par le critère d'accumulation à être présentes dans Finess de manière sûre	87,4	87,4
% des adresses brutes identifiées par le critère Fidéli uniquement à être présentes dans Finess de manière sûre	64,6	64,3
% des adresses brutes identifiées par le critère d'accumulation uniquement à être présentes dans Finess de manière sûre	34,6	34,6
% des adresses brutes identifiées par le critère d'accumulation uniquement à être présentes dans RSS de manière sûre	7,6	7,6
% des individus de plus de 60 ans, vivant à une adresse brute identifiée par le critère Fidéli ET par le critère d'accumulation à être présent dans Finess de manière sûre	91,1	91,2
% des individus de plus de 60 ans, vivant à une adresse brute identifiée par le critère Fidéli uniquement à être présent dans Finess de manière sûre	68,3	67,4
% des individus de plus de 60 ans, vivant à une adresse brute identifiée par le critère d'accumulation uniquement à être présent dans Finess de manière sûre	47,7	47,8
% des individus de plus de 60 ans, vivant à une adresse brute identifiée par le critère d'accumulation uniquement à être présent dans RSS de manière sûre	13,9	14,0

Champ > Adresses brutes déclarées dans la table Individu de Fidéli, France métropolitaine et DROM.
Source > Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Apport de la technique d'appariement des adresses proches pour Finess

À l'échelle des établissements, 9 010 établissements (soit 87,6 % des établissements Finess de départ) sont appariés de manière sûre à la fin des appariements. Ci-dessous les résultats si nous arrêtons l'appariement avant la fin (la dernière étape étant celle qui apparie les adresses proches dans la commune) :

- À l'issue de l'appariement avec l'adresse exacte : 8 532 établissements sont appariés de manière sûre, soit 94,7 % des 9 010 établissements appariés de manière sûre à la fin des appariements.
- À l'issue de l'appariement avec l'adresse exacte, et indice de répétition qui peut différer : 8 559 établissements appariés de manière sûre, soit 95,0 % des établissements appariés de à la fin des appariements.
- À l'issue de l'appariement dans la même rue : 8 870 établissements appariés de manière sûre soit 98,4 % des établissements appariés à la fin des appariements.

À l'échelle des individus de plus de 60 ans, il y a 394 820 seniors dans les établissements Finess appariés de manière sûre (9 010 établissements). Si nous arrêtons l'appariement avant la fin, nous comptons les individus de plus de 60 ans dans les adresses obtenues en cours d'appariement:

- À l'issue de l'appariement avec l'adresse exacte : 85,8 % des individus de plus de 60 ans et appariés à la fin des appariements sont retrouvés.
- À l'issue de l'appariement avec l'adresse exacte, et indice de répétition qui peut différer : 86,6 % des individus de plus de 60 ans et appariés à la fin des appariements sont retrouvés.

- À l'issue de l'appariement avec les adresses dans la même rue : 96,5 % des individus de plus de 60 ans et appariés à la fin des appariements sont retrouvés.

■ EXPERTISE QUALITÉ DE LA MÉTHODE ET PREMIÈRES STATISTIQUES SUR LA POPULATION IDENTIFIÉE

Nous avons construit d'une part une base des individus déclarés dans une RSS en 2019 (pour 638 RSS), et d'autre part une base des individus déclarés dans un des établissements Finess pour personnes âgées en 2019 (pour 9 010 établissements). Dans cette partie, la qualité de la méthode est expertisée en analysant les caractéristiques des individus retrouvés.

Pour les données Finess, le nombre d'individus déclarés et retrouvés dans Fidéli par établissement peut être comparé au nombre de résidents effectivement installés. Cela permet de jauger du taux de déclaration des établissements, et donc de voir s'il y a trop ou trop peu d'individus retrouvés dans Fidéli pour certains établissements. Les données peuvent également être confrontées à d'autres jeux de données. Les individus de Finess-Fidéli peuvent être comparés avec ceux de l'enquête EHPA, qui se base aussi sur le champ des établissements Finess pour personnes âgées. Avec Finess, la catégorie de l'établissement est également disponible. Il est ainsi possible d'analyser le niveau des pensions en fonction de la catégorie de l'établissement, et de conclure si les résultats vont dans le sens attendu : des niveaux plus élevés dans les EHPAD privés à but lucratif. Enfin, les pensions des résidents peuvent être analysées et comparées aux résultats obtenus dans l'enquête Care-I.

Pour la base RSS, malheureusement, étant donné le manque de données concernant les RSS dans le champ de la statistique publique, la structure de la population retrouvée ne peut pas être confrontée à une autre source. La méthode présentée ici permet pour la première fois de proposer une analyse des principales caractéristiques de ces individus (âge, sexe, situation matrimoniale, pension et autres revenus). Les résultats sont disponibles dans l'étude *Résidences-services seniors : des résidents au niveau de vie supérieur à celui des seniors en logements ordinaires* (Esteban, 2023).

Enfin, l'identification des individus vivant dans un établissement Finess et des individus vivant en RSS permet de produire une description comparée de ces populations, en termes de structure et de revenu. Ces populations peuvent également être confrontées à l'ensemble des seniors de plus de 60 ans, tout type de ménage confondu.

Taux de déclaration

Le taux de déclaration d'un établissement est défini comme le rapport entre d'une part le nombre **d'individus déclarés et retrouvés** dans Fidéli au sein de l'ensemble des adresses brutes associées, non décédés au 1^{er} janvier 2020, et âgés de plus de 60 ans¹⁹ et d'autre part le **nombre de résidents installés** à l'établissement²⁰. Afin de calculer ce taux de déclaration, il est nécessaire de connaître le nombre de résidents effectivement installés à l'établissement. Les données Finess renseignent la capacité d'installation de chaque établissement (le nombre de places offertes). Pour se ramener au nombre de résidents effectivement installés, cette capacité est multipliée par un taux d'occupation. L'enquête EHPA 2019 permet d'obtenir un taux d'occupation moyen selon la catégorie de l'établissement (tableau 11). Le nombre de résidents installés est donc calculé pour tous les établissements, selon leur capacité d'installation et leur catégorie.

¹⁹ Ce critère d'âge permet d'éviter de compter le personnel soignant habitant sur place.

²⁰ Il s'agit donc d'une combinaison du taux de déclaration des résidents (la part à se déclarer fiscalement à son établissement) et du taux de recouvrement de la méthode d'appariement (la part des individus retrouvés), sachant qu'il est possible que la méthode assigne à un établissement des individus qui n'y sont pas en réalité (si l'établissement est à la même adresse qu'une habitation classique).

Tableau 11 • Taux de d'occupation par catégorie d'établissement, EHPA 2019

Catégorie d'établissement	Taux d'occupation (%)
EHPAD, dont :	97
EHPAD privés à but lucratif	94
EHPAD privés à but non lucratif	99
EHPAD publics	98
Résidences autonomie, dont :	87
Résidences autonomie privées à but lucratif	89
Résidences autonomie privées à but non lucratif	87
Résidences autonomie publiques	87
EHPA, dont :	91
EHPA privés à but lucratif	83
EHPA privés à but non lucratif	99
EHPA publics	59
Établissements expérimentaux pour personnes âgées	125
USLD	96
Ensemble	96

Note > Le taux d'occupation est calculé en prenant en compte l'ensemble des résidents (y compris ceux de moins de 60 ans). Le taux supérieur à 100 % pour les établissements expérimentaux peut s'expliquer car ce sont des établissements un peu particuliers (peu nombreux, méconnus, peu encadrés).

Champ > France métropolitaine, et DROM, hors Mayotte, structures d'hébergement pour personnes âgées, hors centres d'accueil de jour.

Source > Enquête EHPA 2019, calculs DREES.

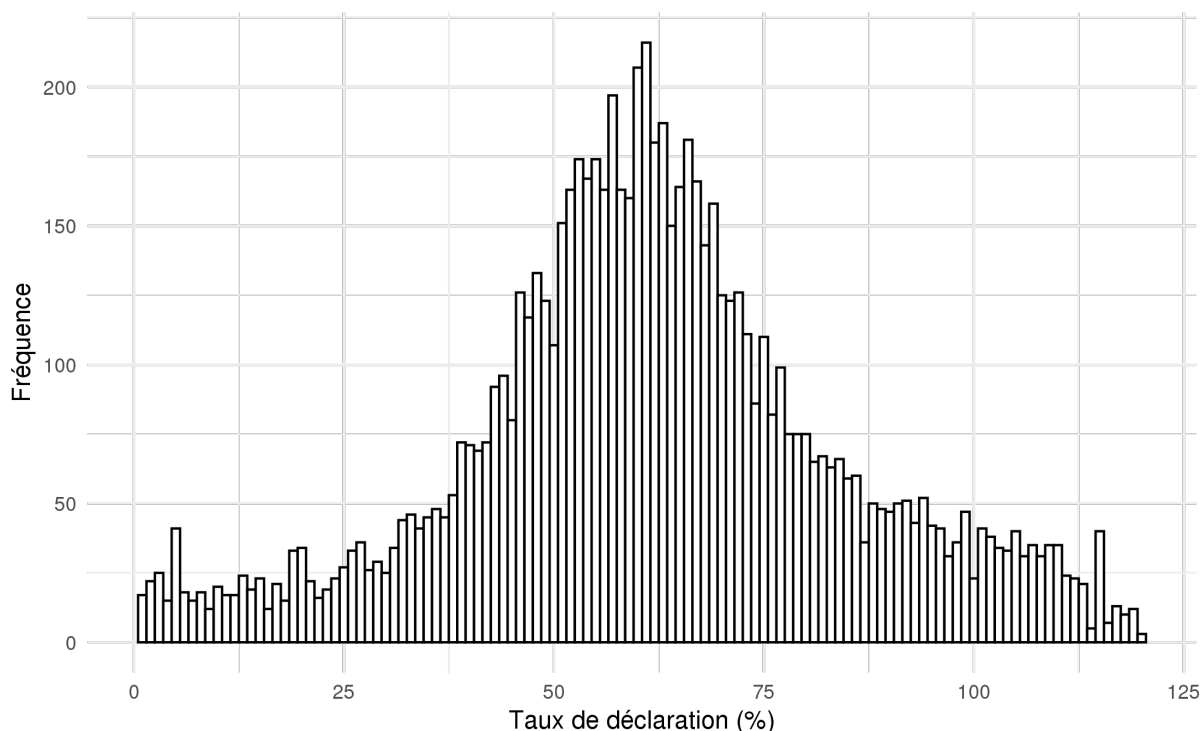
Le calcul du taux de déclaration nécessite de compter les individus à chaque adresse, il faut donc être vigilant quant au champ sur lequel il est calculé : parmi les établissements Finess appariés de manière sûre, un certain nombre sont appariés à une adresse brute qui est elle-même associée à plusieurs établissements Finess. Ainsi, il n'est pas possible de savoir à quel établissement Finess sont rattachés les individus qui vivent à cette adresse brute. Dans le cadre de statistiques individuelles, nous faisons le choix de retirer ces établissements Finess en conflit d'intérêt sur une même adresse brute (il y en a 747, regroupant 30 362 individus). Le champ est donc composé de 8 263 établissements au niveau métropole et DROM, regroupant 406 937 individus. Seuls les 368 325 individus de plus de 60 ans sont conservés (soit 90,5 % des individus). Le taux de déclaration est calculé en faisant le rapport entre le nombre d'individus de plus de 60 ans retrouvés à l'établissement dans Fidéli, et le nombre de résidents installés dans l'établissement²¹.

Sur l'ensemble de ces établissements, la médiane du taux de déclaration est de 61,5 %. 8,8 % des établissements ont un taux de déclaration strictement supérieur à 100 %. Cela ne peut pas s'expliquer par la rotation des résidents au cours de l'année (la place d'une personne décédée est prise par un autre senior dans la même année), car seuls les individus vivant au 31 décembre 2019 sont comptabilisés. En revanche, un taux de déclaration strictement supérieur à 100 % peut s'expliquer si l'établissement se situe à une adresse qui héberge de nombreuses personnes, en dehors de l'établissement lui-même, par exemple un hôpital ou un immeuble. L'appariement entre les individus de Fidéli et ces établissements Finess avec un très fort taux de déclaration n'est donc pas tout à fait satisfaisant. Parmi les 203 établissements avec un taux de déclaration supérieur à 120 %, il y a une sur-représentation des EHPA, USLD et résidences autonomie : 12,3 % de ces 203 établissements sont des EHPA, 4,9 % sont des USLD et 34,0 % sont des résidences autonomie, contre respectivement 1,6 %, 0,3 % et 18,1 % dans l'ensemble des établissements.

Afin d'éviter de prendre en compte ces cas, le graphique 1 est restreint au champ des établissements avec un taux de déclaration inférieurs à 120 %, soit 8 027 établissements au niveau métropole et DROM (97,1 % des établissements étudiés dans cette partie). Pour ces établissements la médiane du taux de déclaration est 61,1 %.

²¹ Ce nombre de résidents installés regroupe tous les résidents, y compris ceux de moins de 60 ans.

Graphique 1 • Fréquences du taux de déclaration



Lecture > 27 établissements ont un taux de déclaration entre 24.5% et 25.5%.

Champ > Établissements Finess à être appariés de manière sûre à Fidéli, sans conflit d'intérêt d'adresse, et avec un taux de déclaration inférieur à 120% (8027 établissements), France métropolitaine et DROM.

Source > Enquête EHPA 2019, Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

399 établissements ont un taux de déclaration strictement inférieur à 20 % (ce qui représente 5,0 % des 8 027 établissements étudiés dans l'histogramme précédent). Il y a parmi ces établissements une forte surreprésentation des EHPAD privés à but non lucratif (45,4 % contre 24,8 % dans l'ensemble des établissements du champ dans cette partie).

Il y a un bien meilleur taux de déclaration dans les résidences autonomie, probablement car il y a plus de logements privatifs dans ces établissements. Dans les USLD, la médiane du taux de déclaration est anormalement élevée, ce qui indique une fois encore que la méthode d'appariement n'est pas pertinente pour ce type d'établissements (tableau 12).

Tableau 12 • Taux de déclaration par catégorie d'établissement

Catégorie	Médiane du taux de déclaration (%), sans retrait des taux supérieurs à 120%	Médiane du taux de déclaration (%), avec retrait des taux supérieurs à 120%
EHPAD privés à but lucratif	54,6	54,5
EHPAD privés à but non lucratif	60,6	60,2
EHPAD publics hospitaliers	55,2	54,6
EHPAD publics non hospitaliers	60,7	60,6
Résidence autonomie	92,0	90,8
EHPA	74,1	67,3
USLD	214,8	57,3
Ensemble	61,5	61,1

Champ > Établissements Finess à être appariés de manière sûre à Fidéli, sans conflit d'intérêt d'adresse, France métropolitaine et DROM.

Source > Enquête EHPA 2019, Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Comparaison avec EHPA 2019

Dans l'enquête auprès des établissements d'hébergement pour personnes âgées (EHPA) 2019, une fiche individuelle permet d'obtenir des renseignements sur chaque résident des établissements (notamment l'âge, la situation familiale, le sexe). Une pondération individuelle pour une représentativité des individus est disponible à l'échelle nationale et régionale, à condition de

ventiler par catégories d'établissement. Les établissements et la population issus de l'appariement Finess-Fidéli peuvent donc être comparés avec les données de l'enquête EHPA 2019 par catégorie d'établissement, sur les données d'âge, de sexe et de situation familiale²².

Structure générale

Dans l'ensemble, les individus présents dans EHPA en hébergement permanent et de plus de 60 ans et les individus retrouvés grâce aux appariements Finess-Fidéli de plus de 60 ans présentent la même distribution dans les différentes catégories d'établissements. En proportion sur le total des individus étudiés, Fidéli contient plus de seniors en résidence autonomie, et moins de seniors en EHPAD public hospitalier (tableau 13).

Tableau 13 • Comparaison de la répartition des individus dans les catégories d'établissement (hors EHPA et USLD), entre EHPA et Finess-Fidéli

Catégorie d'établissement	Nombre d'individus EHPA	% EHPA	Nombre d'individus Finess-Fidéli	% Finess-Fidéli
EHPAD privés à but lucratif	123 470	18,9	65 741	18,0
EHPAD privés à but non lucratif	162 264	24,8	89 317	24,5
EHPAD publics hospitaliers	119 157	18,2	57 045	15,6
EHPAD publics non hospitaliers	154 810	23,6	87 421	24,0
Résidences autonomie	95 088	14,5	65 063	17,8
Ensemble	654 789	100	364 587	100

Champ > Pour EHPA, individus en hébergement permanent dans l'établissement (93,5 % des individus) et de plus de 60 ans (99,2 % des individus en hébergement permanent), hors USLD et EHPA (95,2 % des individus en hébergement permanent de plus de 60 ans) soit 654 789 individus. Pour Finess-Fidéli, individus en établissement hors conflit d'intérêt d'adresse, de plus de 60 ans, hors EHPA et USLD, soit 364 587 individus au niveau métropole et DROM, France métropolitaine et DROM.

Source > EHPA 2019, Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Le tableau 14 permet de comparer la composition des individus, entre EHPA et Finess-Fidéli. Les moyennes d'âge sont assez proches entre la population identifiée dans Fidéli et celle présente dans l'enquête EHPA 2019, pour toutes les catégories d'établissement sauf EHPA. Dans toutes les catégories d'établissement, la proportion de femmes est légèrement plus grande au sein des individus Fidéli qu'au sein des individus EHPA. De manière beaucoup plus marquée, pour toutes les catégories, la proportion de personnes seules est largement plus grande au sein des individus Fidéli qu'au sein des individus EHPA. Ces deux observations sont liées car les femmes sont plus souvent seules (veuves) que les hommes. En effet, dans le champ EHPA, 86,0 % des femmes sont seules contre 71,7 % des hommes, et dans le champ Fidéli, 95,4 % des femmes sont seules contre 85,0 % des hommes. Il y a davantage de personnes seules dans les données Fidéli car la méthode utilisée pour le repérage des individus est basée sur la déclaration fiscale à l'établissement d'hébergement. Or au sein des personnes en couple, il peut arriver que le conjoint soit resté à la résidence familiale, et que donc la déclaration fiscale du couple soit basée à l'adresse de la résidence familiale, et non de l'EHPAD. Les individus en couple sont ainsi plus souvent manqués dans les données Finess-Fidéli. Une autre observation va dans ce sens : la proportion de personnes seules entre Fidéli et EHPA est beaucoup plus proche au sein des résidences autonomie qu'au sein des autres types d'établissement. Cela s'explique car les résidences autonomie proposent des logements privés (complets) pouvant très facilement accueillir des couples, et donc les résidents y sont moins séparés de leur conjoint que dans les autres catégories d'établissements. On constate que 2 % seulement des résidents ont un conjoint en dehors de l'établissement, contre 10 % pour l'ensemble des établissements de l'enquête EHPA²³. Ces résidences ne sont pas destinées à des personnes en perte d'autonomie, mais à des seniors désirant vivre dans une structure en communauté avec un certain nombre de services. Elles sont proches des logements ordinaires, et donc les individus ont une plus grande probabilité de s'y déclarer fiscalement (notamment lorsqu'ils sont en couple), comme cela avait été montré en partie « taux de déclaration » (tableau 12).

²² Il faut noter toutefois une différence de définition concernant la situation familiale : tandis que la situation dans Fidéli est connue grâce aux données fiscales (marié, pacsé, célibataire, divorcé ou veuf), elle est obtenue par déclaration par le personnel dans l'enquête EHPA (sans conjoint, conjoint dans l'établissement ou conjoint hors de l'établissement).

²³ Enquête EHPA 2019, data.drees.

Tableau 14 • Composition des individus dans EHPA et dans Finess-Fidéli

Catégorie	Part à avoir moins de 60 ans (%)		Moyenne d'âge chez les plus de 60 ans		Part de femmes chez les plus de 60 ans (%)		Part de personnes seules chez les plus de 60 ans (%)	
	EHPA	Finess-Fidéli	EHPA	Finess-Fidéli	EHPA	Finess-Fidéli	EHPA	Finess-Fidéli
EHPAD privés à but lucratif	0,5	12,1	88,3	87,5	74,8	77,3	80,9	91,4
EHPAD privés à but non lucratif	0,5	9,6	87,3	86,5	76,1	77,6	82,7	93,6
EHPAD publics hospitaliers	0,9	10,5	85,8	84,2	70,6	71,0	79,6	94,6
EHPAD publics non hospitaliers	0,5	8,3	87,1	86,1	73,9	75,0	83,2	93,2
Résidences autonomie	1,2	6,6	82,9	82,9	72,5	74,0	88,7	91,0
EHPA	1,8	22,7	85,2	83,8	67,0	71,1	84,0	89,0
USLD	2,9	14,5	83,6	83,8	63,4	69,6	69,3	93,7
Ensemble	0,8	9,5	86,4	85,6	73,3	75,2	82,2	92,8

Note > Les personnes seules regroupent les célibataires, divorcés et veufs dans Fidéli, et les personnes sans conjoint dans l'enquête EHPA.

Lecture > La moyenne d'âge parmi les individus vivant en EHPAD privés à but lucratif est de 88,3 ans pour les individus de EHPA de plus de 60 ans et en hébergement permanent, contre 87,5 ans pour les individus de plus de 60 ans dans Fidéli. Parmi ces individus en EHPAD privés à but lucratif, 74,8 % sont des femmes dans les données EHPA, et 77,3 % sont des femmes dans les données Fidéli.

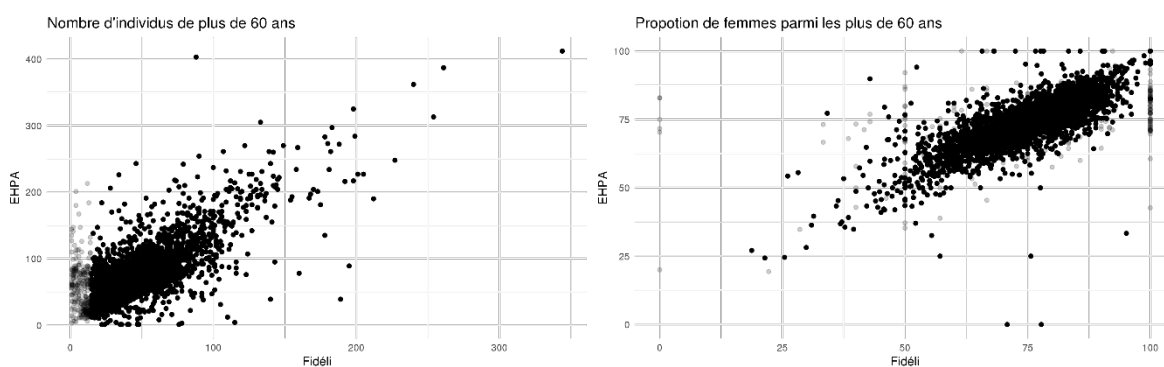
Champ > Pour EHPA, individus en hébergement permanent dans l'établissement. Pour Fidéli, individus en établissements hors conflit d'intérêt d'adresse, France métropolitaine et DROM.

Source > EHPA 2019, Finess 2029, Fantoir 2020, Fidéli 2020, calculs DREES.

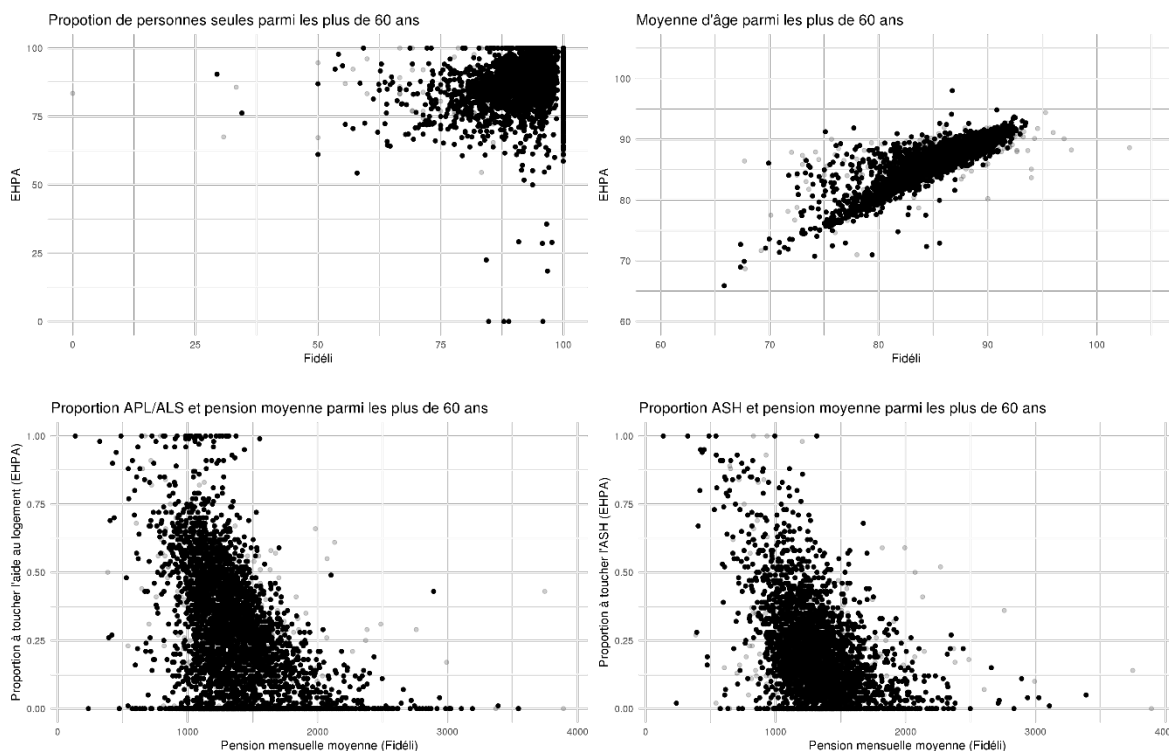
Établissement par établissement

Parmi les établissements Finess appariés de manière sûre avec Fidéli, hors établissements en conflit d'intérêt d'adresse, hors EHPA et USLD (8 104 établissements), 4 100 sont retrouvés dans la base EHPA 2019 portant sur les résidents²⁴, (soit 50,6 %). Pour chacun de ces établissements, la composition des résidents au sein de Fidéli est comparée à la composition au sein d'EHPA.

Graphiques 2 • Nuage de points, pour les établissements présents dans Fidéli-Finess et dans EHPA.



²⁴ Ne sont conservés que les établissements de la base EHPA 2019 qui ont une fiche individuelle suffisamment bien remplie, c'est-à-dire tel que nous avons au moins 80% de personnes décrites par rapport à la capacité installée.



Note > Les petits établissements (strictement moins de 15 individus) sont représentés avec un point gris, car leur faible nombre d'individus ne permet pas d'obtenir des statistiques robustes à l'échelle de l'établissement.

Lecture > Un des 4100 établissements à être présent dans EHPA et Finess-Fidéli compte 227 individus de plus de 60 ans dans Fidéli, et 248 individus en hébergement permanent dans EHPA.

Champ > Établissements Finess appariés de manière sûre à Fidéli, hors conflit d'intérêt d'adresse, hors USLD et EHPA et présents dans la base EHPA 2019 fiche individuelle, avec au moins 80 % des résidents décrits dans cette fiche et avec au moins un individu en hébergement permanent dans EHPA, soit 4100 établissements. Pour l'étude de la part des individus qui touchent une aide sociale à l'hébergement (ASH) dans EHPA, établissements qui ont une habilitation partielle ou totale à toucher l'aide sociale départementale (3 298 sur les 4 100 de départ).

Source > EHPA 2019, Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Dans l'ensemble, les nuages de points forment bien une diagonale, ce qui suggère que les appariements Finess-Fidéli sont en général corrects. Les statistiques issues de la population retrouvée par l'appariement sont similaires à celles obtenues dans EHPA, établissement par établissement. Les corrélations (sans prise en compte des petits établissements) sont particulièrement satisfaisantes pour le nombre d'individus de plus de 60 ans (0,75), la proportion de femmes parmi les plus de 60 ans (0,79) et la moyenne d'âge parmi les plus de 60 ans (0,87).

Au-delà de corrélations satisfaisantes, les résultats sont cohérents avec les attentes. Tout d'abord, il y a davantage de seniors dans EHPA que dans Finess-Fidéli (le taux de déclaration dans Fidéli est autour de 62 %, cf. graphique 1). De plus, la relation est décroissante entre la pension moyenne (*encadré 2*) par établissement et la part des résidents qui bénéficient d'une aide au logement (APL, ALS) d'une part, et la part des résidents qui reçoivent l'aide sociale à l'hébergement (ASH) d'autre part. Plus la pension moyenne des résidents de l'établissement est faible, plus la proportion de seniors qui touchent une de ces aides est élevée. En revanche, il y a une très faible corrélation en ce qui concerne la part des personnes seules (0,31 sans les petits établissements). Ceci peut être expliqué car la définition du statut matrimonial diffère entre Fidéli et EHPA et que les couples sont sous-représentés dans Fidéli.

Analyse de la pension selon la catégorie de l'établissement Finess

Si les appariements sont corrects, les pensions au sein des établissements publics devraient être inférieures aux pensions au sein des établissements privés, comme cela s'observe dans les données Care-I (Boneschi, 2023). Parmi les résidents retrouvés avec la méthode d'appariement, il y a bien une pension plus élevée dans les EHPAD privé à but lucratif que dans les autres catégories d'établissement. D'un point de vue du statut juridique, c'est au sein des organismes privés commerciaux que la pension est la plus élevée (tableau 15).

Tableau 15 • Pension mensuelle par catégorie d'établissement

Catégorie	Répartition (%)	Pension mensuelle moyenne, parmi ceux qui ont une pension (€)	Pension mensuelle médiane, parmi ceux qui ont une pension (€)	Part à avoir une pension nulle (%)
EHPAD privés à but lucratif	17,8	1 926	1 745	2,8
EHPAD privés à but non lucratif	24,2	1 476	1 358	3,4
EHPAD publics hospitaliers	15,5	1 264	1 217	5,8
EHPAD publics non hospitaliers	23,7	1 330	1 252	3,9
Résidences autonomie	17,7	1 395	1 331	3,4
EHPA	0,7	1 455	1 307	5,0
USLD	0,3	1 359	1 254	7,6
Établissement Public	51,4	1 312	1 251	4,4
Etat & Col.Territoriales	0,8	1 406	1 352	3,4
Org. Privé Commercial	18,4	1 914	1 732	2,8
Org.Privé non Lucraif.	29,4	1 484	1 372	3,4
Ensemble	100	1 476	1 346	3,8

Champ > Individus de plus de 60 ans en établissements Finess appariés de manière sûre à Fidéli, hors établissement en conflit d'intérêt d'adresse, France métropolitaine et DROM.
Source > Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Encadré 2 • La pension dans Fidéli

La « pension » fait référence à la case « pensions, retraites, rentes » de la déclaration des revenus 2019. Sont déclarées :

- Les sommes perçues au titre des retraites publiques ou privées y compris les majorations de retraite ou de pension pour charges de famille (*les pensions de réversion sont incluses*) ;
- Les rentes viagères à titre gratuit ;
- Les rentes versées à la sortie d'un PERP, du régime PREFON, d'un contrat Madelin, d'un régime obligatoire de retraite supplémentaire d'entreprise (« article 83 ») ou d'un nouveau plan d'épargne retraite (PER individuel, Pereco, Pero) lorsque les versements ont été déduits d'un revenu catégoriel ou du revenu global ;
- Les prestations de retraite (de source française ou étrangère) versées sous forme de capital, par exemple le capital perçu à l'échéance d'un plan d'épargne pour la retraite populaire (PERP).

Les montants de pension étudiés ici sont donc les montants imposables, y compris CSG non déductible et CRDS, légèrement plus élevés que la valeur nette de la pension perçue.

Source > Notice du formulaire 2042 de la déclaration des revenus 2019.

Comparaison avec Care-Institution

L'enquête Capacité, Aides et REssources des seniors en Institutions (Care-I) a été réalisée par la DREES de septembre à décembre 2016. Le champ de l'enquête est constitué des personnes âgées de 60 ans ou plus hébergées de façon permanente dans les établissements pour personnes âgées de France métropolitaine, de type EHPAD, EHPA et USLD. Les revenus individuels perçus par les seniors enquêtés sont obtenus après un appariement avec les données issues des déclarations de revenus.

Afin d'avoir la population la plus comparable possible, le champ Fidéli est restreint aux individus de France métropolitaine vivant en EHPAD, car les appariements des établissements du type EHPA et USLD sont moins bons, et que les résidences autonomie ne sont pas dans l'enquête Care-I. Les établissements en conflit d'intérêt sur une adresse sont donc retirés afin d'être sûr que l'individu vit bien en EHPAD. Enfin sont seulement comparés les individus seuls (célibataires, divorcés, veufs), car ce sont ceux qui sont le mieux repérés dans Fidéli. Les seniors de plus de 60 ans seuls représentent 93,3 % des seniors de plus de 60 ans dans un EHPAD hors conflit d'intérêt en France métropolitaine. Les écarts de pension moyenne entre les

deux sources sont assez faibles : 5 % dans l'ensemble (tableau 16). Une partie de cet écart peut s'expliquer par l'augmentation générale des pensions entre 2016 et 2019, estimée à 1,8 %²⁵.

Tableau 16 • Pension mensuelle des seniors qui perçoivent une pension, par situation familiale et sexe, dans Care-I et Finess-Fidéli

Statut matrimonial x sexe	Pension mensuelle moyenne dans Care-I (€)	Pension mensuelle moyenne dans Fines-Fidéli (€)	Écart de pension (%)	Taux de détention Care-I (%)	Taux de détention Finess-Fidéli (%)
Veuve	1 460	1 570	8	100	100
Veuf	1 950	2 070	6	99	100
Femme seule non veuve	1 110	1 200	8	92	91
Homme seul non veuf	1 150	1 140	-1	88	90
Ensemble	1 430	1 500	5	97	97

Lecture > Dans Care-I, 99 % des hommes veufs perçoivent une pension, et ceux qui en perçoivent une ont une pension mensuelle moyenne de 1 950 euros. Dans Finess-Fidéli, 100 % des hommes veufs perçoivent une pension, avec un montant de 2 070 euros mensuel en moyenne, soit 6 % de plus que dans Care-I.

Champ > Pour Care-I : individus seuls, en EHPAD. Pour Finess-Fidéli : individus seuls de plus de 60 ans, dans un établissement hors conflit d'intérêt, de catégorie EHPAD, France métropolitaine.

Source > Care-I 2016, Fines 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Données sur les revenus du ménage

La table « revenus ménage » de Fidéli donne des informations sur les revenus et les prestations sociales à l'échelle d'un logement, pour l'année qui vient de s'écouler (2019) et l'année précédente (2018). Dans le millésime 2020 utilisé ici, ces données ne sont malheureusement disponibles que pour 2018²⁶.

Les revenus du ménage ne sont disponibles que pour les ménages qui habitent dans un logement assujéti à la taxe d'habitation. Or, les résidents d'établissements Finess ne sont que très rarement dans des logements assujéti à la taxe d'habitation (il faut qu'ils aient une utilisation privative de leur logement)²⁷. **Ainsi pour les individus des établissements Finess, il n'est pas possible d'apparier avec la table des revenus du ménage.**

En revanche, les résidents de RSS sont dans des logements ordinaires, assujéti à la taxe d'habitation. Ils ont donc un identifiant du logement et il est possible de les apparier à la base des revenus des ménages. Plus précisément, 78 % des individus de RSS de plus de 60 ans disposent d'un identifiant du logement pour 2019, et 84,6 % disposent d'un identifiant du logement pour 2018 (celui nécessaire ici), et ils peuvent donc être apparés à la base des revenus du ménage. Les données étudiées sont donc des données de 2018 portant sur des individus de plus de 60 ans fin 2019 et qui sont résidents en RSS en 2019. Parmi la population dont nous pouvons retrouver les revenus du ménage, 3 % n'a pas de renseignement complété pour l'année 2018. Au total, on dispose de l'information sur les revenus du ménage pour 82 % des résidents de RSS de plus de 60 ans. Ils sont représentatifs de l'ensemble des résidents en termes d'âge, de sexe et de situation matrimoniale. Il faut noter cependant que leur situation matrimoniale a évolué entre 2018 et 2019 ; il y avait en 2018 plus de résidents mariés/pacsés en proportion, et moins de résidents veufs en proportion (tableau 17).

²⁵ Calcul issu des données de la fiche 5 du panorama de la DREES *Les retraités et les retraites, édition 2022*.

²⁶ La documentation de Fidéli indique : « Suite aux exonérations effectuées par le gouvernement pour les revenus de 2019, le revenu disponible tel que mesuré par les sources fiscales, a été sensiblement affecté. Ce problème a conduit à ne pas diffuser les revenus au niveau de l'ensemble des occupants du logement (Filosofi 2019) dans le millésime 2020 de Fidéli. Les variables de revenus relatives à cette source seront donc renseignées à 0. »

²⁷ Parmi les résidents retrouvés avec notre apparier, hors établissement en conflit d'intérêt et de plus de 60 ans, 80.1% ne sont pas assujéti à la taxe d'habitation et n'ont pas d'identifiant logement pour 2019, variable qui permet de retrouver les revenus dans la table « revenus ménage » de Fidéli.

Tableau 17 • Composition du champ des résidents de plus de 60 an en RSS et du champ dont les revenus 2018 sont disponibles

Catégorie	Résidents RSS		Résidents dont les revenus 2018 sont disponibles			
	Effectif	Répartition	Effectif (avec situation matrimoniale de 2019)	Répartition (avec situation matrimoniale de 2019)	Effectif (avec situation matrimoniale de 2018)	Répartition (avec situation matrimoniale de 2018)
0-79	6 134	22,4	5 031	22,4	Même chose qu'avec la situation matrimoniale de 2019.	
80-84	5 043	18,4	4 123	18,4		
85-89	7 869	28,7	6 413	28,5		
90-94	6 235	22,8	5 124	22,8		
95-	2 129	7,8	1 780	7,9		
Célibataire/ divorcé	5 725	20,9	4 576	20,4	4 550	20,3
Marié/pacsé	5 904	21,5	5 096	22,7	5 703	25,4
Veuf	15 711	57,3	12 758	56,8	12 177	54,2
Autre	70	0,3	41	0,2	41	0,2
Femme	20 029	73,1	16 432	73,1	Même chose qu'avec la situation matrimoniale de 2019.	
Homme	7 367	26,9	6 025	26,8		
F célibataire/ divorcée	3 902	14,2	3 146	14,0	3 137	14,0
F mariée/ pacsée	2 920	10,7	2 546	11,3	2 966	13,2
Veuve	13 172	48,1	10 723	47,7	10 312	45,9
F autre	35	0,1	17	0,1	17	0,1
H célibataire/ divorcé	1 823	6,7	1 430	6,4	1 413	6,3
H marié/pacsé	2 984	10,9	2 550	11,4	2 737	12,2
Veuf	2 539	9,3	2 035	9,1	1 865	8,3
H autre	21	0,1	10	0,00	10	0,00
Ensemble	27 410	100	22 471	100	22 471	100

Lecture > 20,9 % des résidents de RSS sont célibataires ou divorcés. Parmi les résidents dont les revenus de 2018 sont disponibles, 20,4 % sont célibataires ou divorcés en 2019, et ils étaient 20,3 % en 2018.

Champ > Individus de plus de 60 ans retrouvés à l'adresse d'une RSS en 2019, France métropolitaine et DROM.

Source > Recensement des RSS « sanitaire et social » et « cap résidence » 2022, Fantoir 2020, Fidéli 2020, calculs DREES.

Les niveaux de pension sont également comparables : 1 928 euros par mois en médiane pour les individus dont nous disposons des revenus 2018, contre 1 919 euros par mois en médiane pour l'ensemble des individus en RSS. 2,8 % des individus dont nous disposons des revenus de 2018 ont une pension manquante, contre 3,1 % parmi l'ensemble des individus de RSS.

Pour une analyse de la pension et des autres types de revenus des résidents de RSS, voir l'étude *Résidences-services seniors : des résidents au niveau de vie supérieur à celui des seniors en logements ordinaires* (Esteban, 2023).

Description de la population identifiée

La composition de la population identifiée est comparée entre les différents champs :

- L'ensemble des seniors de plus de 60 ans dans Fidéli ;
- Les résidents d'établissements Finess (hors EHPA et USLD), de plus de 60 ans ;
- Les résidents de RSS, de plus de 60 ans.

Tableau 18 • Composition de la population des 60 ans et plus, et de la population identifiée dans Finess et en RSS

	Seniors de 60 ans et plus		Finess		RSS	
	Nombre	%	Nombre	%	Nombre	%
Ensemble	17 728 524	100	393 708	100	27 410	100
Femme	9 824 914	55,4	295 680	75,1	20 029	73,1
Marié/pacsé	10 128 069	57,1	28 148	7,2	5 904	21,5
Veuf	3 418 701	19,3	239 337	60,8	15 711	57,3
Célibataire	1 635 307	9,2	72 763	18,5	2 036	7,4
Divorcé	2 516 638	14,2	53 050	13,5	3 689	13,5
Autre	29 809	0,2	410	0,1	70	0,3
60-64 ans	4 136 375	23,3	12 119	3,1	699	2,6
65-69 ans	3 900 500	22,0	20 077	5,1	1 159	4,2
70-74 ans	3 460 066	19,5	27 265	6,9	1 786	6,5
75-79 ans	2 189 976	12,4	29 356	7,5	2 490	9,1
80-84 ans	1 835 274	10,4	54 247	13,8	5 043	18,4
85-89 ans	1 338 175	7,6	94 435	24,0	7 869	28,7
90-94 ans	654 334	3,7	100 439	25,5	6 235	22,8
95- et plus	213 824	1,2	55 770	14,2	2 129	7,8
Titulaire carte invalidité	845 840	4,8	56 557	14,4	2 863	10,5
Titulaire manquant	66 713	0,4	759	0,2	263	1,0
Perception d'une pension	15 373 848	86,7	378 498	96,1	26 574	97,0
Perception d'un revenu d'activité	3 076 016	17,4	4 955	1,3	607	2,2
Perception d'une pension alimentaire.	112 383	0,6	7 350	1,9	352	1,3

Lecture > 55,4 % des seniors de plus de 60 ans sont des femmes. Elles représentent 75,1 % des résidents en établissement Finess, et 73,1 % des résidents en RSS.

Champ > Individus de 60 ans et plus vivant à la fin de l'année 2019, en France métropolitaine et DROM, en établissements Finess hors USLD et EHPA, et en RSS.

Source > Fines 2019, recensement des RSS « sanitaire et social » et « cap résidence » 2022, Fantoir 2020, Fidéli 2020, calculs DREES.

Les résidents d'établissements Finess sont plus âgés que ceux en RSS, eux-mêmes plus âgés que les seniors de 60 ans et plus en population générale. En lien direct avec cette observation, la composition en termes de situation familiale est elle aussi différente : il y a moins de mariés/pacsés et plus de veufs dans Finess que dans RSS, et la différence est encore plus marquée avec la population générale²⁸. Parmi les plus de 60 ans en population générale, 86,7 % ont une pension, contre 96,1 % dans Finess et 97,0 % dans les RSS.

Dans les établissements Finess, la pension est en médiane plus faible que dans la population générale. En revanche elle est beaucoup plus élevée pour les individus en RSS (Esteban 2023).

²⁸ Une partie de cet écart est expliquée par la méthode elle-même, qui identifie moins bien les individus en couple en établissement.

Tableau 19 • Niveaux de revenus et dépenses de la population des plus de 60 ans, et de la population identifiée dans Finess et en RSS

	Q1	Médiane	Moyenne	Q3	Part nulle (%)
Seniors de 60 ans et plus (17 728 524 individus)					
Pension	950	1 455	1 610	2 031	13,3
Pension alimentaire	200	305	514	600	99,4
Revenu total individuel	1 009	1 531	1 801	2 163	4,7
Revenu d'activité	321	1 098	1 825	2 260	82,7
Âge	65	71	72,3	79	0
Finess (393 708 individus)					
Pension	953	1 343	1 470	1 826	3,9
Pension alimentaire	200	377	484	660	98,1
Revenu total individuel	963	1 355	1 488	1 844	3,4
Revenu d'activité	110	507	1 234	1 640	98,7
Âge	81	88	85,5	92	0
RSS (27 410 individus)					
Pension	1 336	1 919	2 143	2 676	3,1
Pension alimentaire	250	498	668	990	98,7
Revenu total individuel	1 353	1 931	2 166	2 693	2,2
Revenu d'activité	157	827	1 467	1 850	97,8
Âge	80	86	84,7	90	0

Lecture > La moitié des seniors de 60 ans et plus ont une pension mensuelle supérieure à 1 455 euros, contre 1 343 euros parmi les résidents d'établissements Finess et 1 919 euros parmi les résidents de RSS.

Champ > Individus de plus de 60 ans vivant à la fin de l'année 2019, en France métropolitaine et DROM, en établissements Finess hors USLD et EHPA, et en RSS.

Source > Fines 2019, recensement des RSS « sanitaire et social » et « cap résidence » 2022, Fantoir 2020, Fidéli 2020, calculs DREES.

■ CONCLUSION

La méthode d'appariement permet d'identifier dans la table Individu de Fidéli les seniors qui résident en établissements sanitaires et sociaux pour personnes âgées et en résidences-services seniors. Du côté des établissements sanitaires et sociaux pour personnes âgées, 87,6 % des établissements sont retrouvés dans les tables Fidéli, et regroupent 437 300 individus, dont 90,3 % ont plus de 60 ans. Leur identification permet de produire des statistiques annuelles et robustes sur leur composition en termes d'âge, de sexe et de situation familiale et de décrire leurs revenus individuels, en particulier leur pension.

Concernant les résidences-services seniors, le recensement des sites internet sanitaire et social et cap résidences seniors, couplé à la méthode d'appariement à Fidéli, conduit à identifier 32 900 résidents répartis dans 638 résidences-services seniors, dont 83,3 % ont plus de 60 ans. La constitution d'une base individuelle des résidents de résidences-services seniors permet de décrire pour la première fois la composition de cette population, en termes d'âge, de sexe, de situation matrimoniale et de revenus individuels. De plus, les seniors vivant en résidences-services seniors étant en logement ordinaire, les données Fidéli permettent d'avoir accès aux revenus à l'échelle de leur ménage, notamment leurs prestations sociales, ou leurs revenus fonciers et mobiliers.

Les caractéristiques des seniors en établissements sanitaires et sociaux pour personnes âgées et des seniors en résidences-services seniors peuvent être confrontées aux caractéristiques des seniors de plus de 60 ans résidant en ménage ordinaire. Ces statistiques peuvent être produites en ventilant par exemple par département ou par catégorie d'établissement, et offrent donc une vaste possibilité d'analyse de ces populations. L'information sur leurs revenus individuels et sur les revenus de leur ménage (pour les seniors en résidences-services seniors) permet d'avoir un premier aperçu des inégalités économiques entre les seniors selon leur lieu de vie. Cependant ces informations de revenus ne sont pas suffisantes pour comparer le réel niveau de vie des seniors, après leurs dépenses pour l'hébergement. Il faudrait pour cela connaître leur patrimoine, car les seniors ont davantage de patrimoine que le reste de la population, et une partie de leurs dépenses d'hébergement peut être financée par l'accumulation de ce patrimoine (par exemple lorsque la mise en vente d'un bien immobilier permet de payer les frais d'EHPAD). L'Insee estime en effet que la moitié des ménages dont la personne de référence est âgée de plus de 70 ans ont un patrimoine net²⁹ supérieur à 209 900 euros en 2021, contre 124 800 euros pour l'ensemble des ménages (Cheptitski, A., Cheloudko, P., Hagège, C., Hubert, O., 2023), soit 68 % de plus. Il s'agit donc d'un facteur important à prendre en compte, en plus des revenus, si l'on souhaite analyser le niveau de vie des seniors. Cette donnée n'est malheureusement pas disponible dans Fidéli.

Il faut enfin noter que la population retrouvée est la population des résidents fiscalement déclarés à leur établissement ou à leur résidence-service seniors, vivant à la fin de l'année 2019. Dans le cas des établissements sanitaires et sociaux, les seniors qui se déclarent à leur établissement sont plus probablement des seniors sans conjoint à l'ancien domicile, qui ne sont plus propriétaires de leur ancien domicile, qui ont une ancienneté supérieure à un an dans leur établissement et qui sont suffisamment accompagnés pour que la déclaration fiscale soit bien mise à jour à la nouvelle adresse. Autour de 38 % des résidents d'établissements sanitaires et sociaux pour personnes âgées ne sont pas couverts par l'appariement. Ce taux est beaucoup plus faible dans les résidences autonomes, qui se rapprochent plus d'un logement ordinaire. Dans les résidences-services seniors, qui sont des logements ordinaires, le taux de résidents non déclarés fiscalement à la résidence est probablement encore plus faible, mais il n'est pas possible de l'estimer. La population identifiée représente donc un champ particulier au sein des résidents d'établissements pour personnes âgées. De plus il est possible que la méthode manque certains seniors qui sont pourtant bien déclarés à l'adresse de leur établissement (s'il y a une erreur de déclaration non repérée). Il serait donc idéal de procéder à un redressement de la population identifiée. Malheureusement, nous ne disposons pas de variable de calage, car aucun jeu de données ne permet d'avoir une description de la population fiscalement déclarée en établissements pour personnes âgées.

La plupart des données utilisées dans cette méthode d'appariement sont mises à jour annuellement (Finess, Fidéli, Fantoir), ce qui permet de produire les mêmes statistiques chaque année afin de mettre en œuvre des études longitudinales et de regarder l'évolution de la composition de la population identifiée.

²⁹ Le patrimoine net est le montant total des actifs détenus par un ménage (patrimoine brut) duquel est déduit le montant du capital qu'il doit encore au titre de ses emprunts contractés pour acquérir un bien immobilier, un bien d'équipement, ou pour tout autre motif personnel ou professionnel.

■ POUR EN SAVOIR PLUS

Balavoine, A. (2022, juillet). Des résidents de plus en plus âgés et dépendants dans les établissements d'hébergement pour personnes âgées. DREES, *Études et Résultats*, 1237.

Boneschi, S. (2023). La moitié des seniors en institution ont un niveau de vie inférieur à 1 400 euros par mois. DREES, *Études et Résultats*, à paraître.

Cheptitski, A., Cheloudko, P., Hagège, C., Hubert, O. (2023, janvier). Début 2021, 92 % des avoirs patrimoniaux sont détenus par la moitié des ménages. Insee, *Insee Focus*, 287.

EHPA 2019, data.drees (https://data.drees.solidarites-sante.gouv.fr/explore/dataset/587_l-enquete-aupres-des-etablissements-d-hebergement-pour-personnes-agees-ehpa/information/)

Esteban, L. (2023, mars). Résidences-services seniors : des résidents au niveau de vie supérieur à celui des seniors en logements ordinaires. DREES, *Études et Résultats*, 1261.

Kuhn Lafont, A., Troutot, R. (2022, mars). Résidences services seniors : appréhender une offre devenus incontournable pour les territoires. *Les études de Matières Grises*, mars 2022.

Marino, A. (2022, mai). *Les retraités et les retraitées* – Édition 2022. DREES, coll. Panoramas de la DREES-Social.

Annexe 1. Caractéristiques des établissements Finess non appariés

Caractéristiques des établissements Finess non appariés à Fantoir

Outre les DROM, collectivités d'Outre-Mer et la Corse qui ont très peu d'établissements Finess dès le départ, c'est en Occitanie que les résultats sont les moins bons (9,7 % des établissements ne sont pas appariés à Fantoir). Inversement, c'est en Île de France que les résultats sont les meilleurs (1,2 % des établissements ne sont pas appariés à Fantoir). Dans l'ensemble au niveau métropole et DROM, 97,6 % des établissements Finess sont appariés de manière satisfaisante à Fantoir.

Tableau A1 • Part des établissements Finess à ne pas être appariés à Fantoir de manière satisfaisante, par région

Région	Nombre dans Finess	Nombre non apparié à Fantoir	% non apparié à Fantoir
Collectivités d'Outre-Mer	2	2	100
Corse	30	10	33,3
Martinique	34	4	11,8
Occitanie	1 011	98	9,7
La Réunion	24	2	8,3
Bourgogne-Franche-Comté	535	26	4,9
Provence-Alpes-Côte d'Azur	748	32	4,3
Normandie	638	27	4,2
Nouvelle-Aquitaine	1 222	45	3,7
Auvergne-Rhône-Alpes	1 280	36	2,8
Centre-Val de Loire	426	9	2,1
Hauts-de-France	829	13	1,6
Grand Est	863	12	1,4
Pays de la Loire	798	11	1,4
Bretagne	610	8	1,3
Île-de-France	1 209	14	1,2
Guadeloupe	24	0	0
Guyane	6	0	0
Ensemble	10 289	349	3,4

Champ > Établissements Finess pour personnes âgées, France métropolitaine et DROM.

Source > Finess 2019, Fantoir 2020, calculs DREES.

Ce sont les EHPA qui ont les plus mauvais résultats pour cette première étape d'appariement (8,6 % des EHPA ne sont pas appariés de manière satisfaisante à Fantoir), tandis que les EHPAD privés à but lucratif ont les meilleurs résultats (2,8 % de ces EHPAD ne sont pas appariés de manière satisfaisante à Fantoir).

Tableau A2 • Part des établissements Finess à ne pas être appariés à Fantoir de manière satisfaisante, par catégorie d'établissement

Région	Nombre dans Finess	Nombre non apparié à Fantoir	% non apparié à Fantoir
EHPA	314	27	8,6
EHPAD publics hospitaliers	1 210	50	4,1
USLD	166	6	3,6
EHPAD privés à but non lucratif	2 359	84	3,6
EHPAD publics non hospitaliers	2 133	67	3,1
Résidences autonomie	2 290	65	2,8
EHPAD privés à but lucratif	1 817	50	2,8
Ensemble	10 289	349	3,4

Champ > Établissements Finess pour personnes âgées, France métropolitaine et DROM.

Source > Fines 2019, Fantoir 2020, calculs DREES.

Caractéristiques des établissements Finess non appariés à Fidéli

En ce qui concerne l'appariement avec Fidéli, c'est en Île de France que les résultats sont les moins bons (14,8 % des établissements ne sont pas appariés à Fidéli), tandis qu'ils sont très bons en Bourgogne-Franche-Comté (5,3 % d'échec seulement).

Tableau A3 • Part des établissements Finess de à ne pas être appariés à Fidéli, par région

Région	Nombre dans Finess	Nombre non apparié à Fidéli	% non apparié à Fidéli
Guyane	6	3	50
Guadeloupe	24	4	16,7
Île-de-France	1 195	177	14,8
Normandie	611	87	14,2
Martinique	30	4	13,3
Corse	20	2	10
Nouvelle-Aquitaine	1 177	114	9,7
Hauts-de-France	816	75	9,2
Occitanie	913	80	8,8
Pays de la Loire	787	64	8,1
Auvergne-Rhône-Alpes	1 244	101	8,1
Grand Est	851	68	8
Provence-Alpes-Côte d'Azur	716	57	8
Centre-Val de Loire	417	25	6
Bretagne	602	36	6
Bourgogne-Franche-Comté	509	27	5,3
La Réunion	22	0	0
Ensemble	9 940	924	9,3

Champ > Établissements Finess pour personnes âgées correctement appariés à Fantoir, France métropolitaine et DROM.

Source > Fines 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Les USLD et EHPA ont les résultats d'appariement avec Fidéli les moins satisfaisants. Viennent ensuite les résidences autonomies, parmi lesquelles 23,5 % ne sont pas appariées à Fidéli. Les meilleurs résultats sont pour les EHPAD publics hospitaliers, avec 1,6 % d'échec seulement.

Tableau A4 • Part des établissements Finess de à ne pas être appariés à Fidéli, par catégorie d'établissement

Catégorie	Nombre dans Finess	Nombre sans adresse	% non satisfaisant dans la catégorie
USLD	160	84	52,5
EHPA	287	93	32,4
Résidences autonomie	2 225	523	23,5
EHPAD publics hospitaliers	1 160	69	6
EHPAD privés à but non lucratif	2 275	88	3,9
EHPAD privés à but lucratif	1 767	34	1,9
EHPAD publics non hospitaliers	2 066	33	1,6
Ensemble	9940	924	9,3

Champ > Établissements Finess pour personnes âgées correctement appariés à Fantoir, France métropolitaine et DROM.

Source > Fines 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Annexe 2. Mesures de similarité textuelle

L'ensemble des mesures présentées ici permettent d'obtenir la distance entre deux chaînes de caractère. La distance est ensuite ramenée à une similarité, dont la valeur est comprise en 0 (chaînes pas similaires du tout) et 1 (chaînes très similaires).

Damerau-Levenshtein

La distance entre deux chaînes est définie comme le nombre minimal d'opérations (insertion, suppression, substitution ou transposition) nécessaire pour passer d'une chaîne à l'autre.

La « full » distance Damerau-Levenshtein est utilisée. Il s'agit de l'« optimal string alignment distance » avec autorisation d'effectuer plusieurs modifications sur la même sous-chaîne de caractère.

Par exemple, la distance entre « CA » et « ABC » est 2 : CA → transposition → AC → insertion → ABC

Pour ramener cette distance à une mesure de similarité, nous mesurons l'écart entre la distance et la longueur de la plus longue chaîne, et nous le rapportons à la longueur de la plus longue chaîne. Par exemple entre les deux chaînes de caractère « rue des clés » et « rue des clefs » la distance est de 2 (changement de « é » en « e » et ajout de « f »), et la plus longue chaîne de caractère est de longueur 13. La similarité est donnée par $\frac{13-2}{13} = 0.85$.

Longueur des mots

Il s'agit de la somme des longueurs des mots en commun. Nous regardons les mots en commun entre deux chaînes de caractère, puis nous faisons la somme des longueurs de ces mots. Par exemple « rue Hugonot » et « rue Robert Hugonot » ont pour mots communs « rue » et « Hugonot », la mesure sera donc de 3 + 7 = 10.

La similarité est calculée en divisant cette mesure par le nombre de caractère (en retirant les espaces) de la chaîne la plus courte. Ici, la similarité vaut donc $\frac{10}{10} = 1$.

Plus longue chaîne de caractère commune

Entre deux chaînes de caractère, nous identifions toutes les sous-chaînes possibles, et nous regardons les sous-chaînes communes. La mesure est la longueur de la plus grande sous-chaîne. Par exemple « rue maison vieille » et « rue maison vieillie » auront pour plus longue sous-chaîne de caractère « rue maison vieill » de longueur 17.

Afin de se ramener à une mesure de similarité, nous divisons cette mesure par la longueur de la chaîne la plus courte, ici 18. La similarité est donc $\frac{17}{18} = 0.94$.

Annexe 3. Détermination du critère d'accumulation

Le critère d'accumulation vise à identifier si une adresse brute de Fidéli est un établissement pour personnes âgées (de type EHPAD, EHPA, USLD, résidence autonomie) ou non. Il se base sur les caractéristiques des adresses concernant le nombre des individus qui y vivent et leur âge. Il y a donc certains choix de seuils à faire :

- La barrière d'âge à considérer pour savoir quel seuil d'âge appliquer lorsqu'on cherche un certain nombre d'individus plus âgés que ce seuil (qu'est-ce qu'un « senior » ?) ;
- Le nombre d'individus déclarés à l'adresse plus âgés que ce seuil. Ce critère permet de repérer des adresses avec une accumulation de seniors ;
- La concentration des seniors : la part que représentent les seniors sur l'ensemble des personnes déclarées à l'adresse. Ce critère permet d'éviter de repérer des immeubles qui contiennent beaucoup d'individus, dont beaucoup de seniors.

Construction de la base pour produire les statistiques

La première étape consiste à constituer la base des adresses qui seront analysées pour déterminer les seuils appropriés. Les données de la table Individu de Fidéli sont utilisés³⁰. Plus précisément, la composition des adresses où il n'y a aucun doute sur la présence d'un établissement Finess est étudiée. Afin d'isoler ces adresses et les individus qui y vivent, la base d'étude est construite comme suit.

À partir de la base des adresses Finess appariées à Fantoir, seuls les établissements dont l'appariement a réussi en première étape sont conservés, c'est-à-dire que le libellé d'adresse est exactement le même entre Finess et Fantoir. Il reste donc 7 552 établissements. Les établissements qui partagent une même adresse (conflit d'intérêt d'adresse) sont retirés, car il n'est pas possible de déterminer dans quel établissement vivent les individus de l'adresse. Ceci retire 436 établissements, pour obtenir une base de 7 116 établissements. Cette base est ensuite appariée à la base Individu de Fidéli, en utilisant l'adresse brute complète. 296 établissements ne sont pas du tout retrouvés dans Fidéli, ce qui laisse une base de 6 820 établissements. Enfin, pour 45 établissements, aucun individu de plus de 60 ans n'est retrouvé dans Fidéli. Ils sont retirés, ce qui laisse une base avec 6 775 établissements. Pour chacun de ces établissements, une seule adresse est conservée : celle où il y a le plus d'individus de plus de 60 ans, afin de viser les potentiels résidents (un individu de moins de 60 ans est plus probablement quelqu'un qui travaille dans l'établissement)³¹. Enfin les doublons d'identifiants sont retirés, lorsque le même individu est présent plusieurs fois dans la base (0,03 % des individus). Il s'agit d'individu bilocalisés. Si l'individu est deux fois à la même adresse, nous gardons aléatoirement une des deux lignes, et s'il est à deux adresses différentes (et donc deux établissements différents), la ligne conservée est celle qui se situe à l'établissement avec la fin de l'identifiant Finess le plus petit (choix aléatoire). La base finale contient 6 775 établissements et 291 119 individus.

Une base avec une adresse unique par établissement Finess, et tous les individus qui y vivent est donc constituée. Cette adresse présente les caractéristiques suivantes :

- Le libellé d'adresse Finess est exactement identique au libellé d'adresse dans Fantoir ;
- L'établissement à cette adresse n'a pas de conflit d'intérêt d'adresse (aucune de ses adresses potentielles n'est associée à un autre établissement apparié à Fantoir en première étape) ;
- L'appariement avec Fidéli est exact (l'adresse brute Fantoir est exactement égale à l'adresse brute Fidéli) ;
- Il y a au moins un individu de plus de 60 ans à cette adresse ;
- L'adresse est celle qui comporte le plus d'individus de plus de 60 ans dans Fidéli (dans le cas où un établissement Finess présente plusieurs adresses qui cochent les critères précédents).

78,9 % des adresses de cette base sont dans la table des adresses hors TH, et 73,4 % des adresses sont reconnues par le critère *id scat*. Ce statut de reconnaissance dépend énormément de la catégorie de l'établissement ; 93,6 % des EHPAD sont identifiés par le critère Fidéli, contre 6,3 % des résidences autonomie.

³⁰ Étant donné que le critère sera ensuite appliqué aux adresses de Fidéli, il est cohérent que sa construction se base sur la structure des données Fidéli.

³¹ Pour un de ces établissements, il y a deux adresses avec un nombre d'individus de plus de 60 ans égal au maximum. On prend aléatoirement la première de ces adresses.

Détermination des seuils

La composition de ces établissements en termes de nombre et d'âge des résidents est étudiée. Cette étude est faite par catégorie d'établissement, car les résultats diffèrent beaucoup selon la catégorie. En particulier, les résultats en USLD et EHPA ne sont pas satisfaisants (nous y trouvons très peu de seniors, les résidents semblent plus jeunes). Ceci s'explique notamment car il y a beaucoup de résidents en hébergement temporaire dans les EHPA, et car l'ancienneté des résidents en hébergement permanent au sein des USLD est plus courte. Dans l'ensemble dans ces catégories établissements, les résidents ont une plus faible probabilité de se déclarer fiscalement à l'établissement. Le choix est fait de ne pas prendre en compte leurs statistiques de composition dans le choix des seuils. Cette exclusion se justifie d'autant mieux que ces catégories d'établissement ne représentent qu'une très faible part des établissements de la base d'étude constituée pour cette partie. En effet, 75,6 % des établissements sont des EHPAD, 21,6 % sont des résidences autonomie et seulement 2,8 % des établissements de la base sont des USLD et EHPA.

Seuil d'âge

Pour chaque adresse, on calcule le premier quartile de l'âge (Q1, qui indique l'âge minimum de 75% des individus à adresse), puis par catégorie d'établissement on regarde les premiers déciles (D1 et D2) et le premier quartile (Q1) de ce premier quartile (tableau A5).

Tableau A5 • Répartition des âges dans chaque adresse de la sous-base d'étude, par catégorie d'établissement

Catégorie	Nombre d'établissements	D1 du 1 ^{er} quartile	D2 du 1 ^{er} quartile	Q1 du 1 ^{er} quartile
EHPAD privés à but lucratif	1 269	69,8	81	82
EHPAD privés à but non lucratif	1 623	67	76,3	79
EHPAD publics hospitaliers	766	55	67,5	70
EHPAD publics non hospitaliers	1 465	67,3	76	78,5
Résidences autonomie	1 465	34	67	69
EHPA	130	13	24,5	29,3
USLD	57	ns	ns	ns
Ensemble	6 775	59	71	73

Lecture > Parmi les 1 269 EHPAD privés à but lucratif, 80 % ont un premier quartile de plus de 81 ans, c'est-à-dire que 80 % de ces établissements ont des résidents tels que 75 % d'entre eux ont plus de 81 ans.

Champ > Individus vivant dans la sous-base d'étude.

Source > Fines 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Avec le premier décile, les résultats sont peu satisfaisants car les âges sont très faibles, surtout dans les résidences autonomie. Ces chiffres faibles au premier décile peuvent peut-être s'expliquer par la présence d'adresses qui sont soit des adresses qui correspondent également à un autre immeuble que l'établissement pour personnes âgées en question (comme pour les EHPAD hospitaliers), soit des adresses qui correspondent à des établissements « intergénérationnels » (comme certaines résidences autonomie). Il n'y a pas tant d'adresses avec ce type de problématique, car les âges obtenus au second décile deviennent plus cohérents avec les attentes. Enfin, le premier quartile permet de dire que dans la catégorie où les âges sont le plus faible (les résidences autonomie), **75% des établissements ont un quartile de plus de 69 ans, c'est-à-dire que 75% de leurs déclarants ont plus de 69 ans**. Le seuil d'âge indiqué par ces résultats est donc 70 ans.

Une analyse complémentaire confirme ce choix de seuil. Les individus de toutes les adresses de la même catégorie d'établissement sont regroupés pour éviter de donner trop de poids aux établissements potentiellement problématiques. Dans les résidences autonomie, **75% des individus fiscalement déclarés ont plus de 71 ans** (tableau A6). Dans les autres catégories d'établissements, les résidents sont toujours plus âgés (sauf en EHPA et USLD). Le seuil d'âge pour comptabiliser les individus pour les points d'accumulations à une adresse donnée est ainsi fixé à 70 ans.

Tableau A6 • Répartition des âges dans les catégories d'établissement dans la sous-base d'étude

Catégorie	Nombre d'établissements	D1	D2	Q1
EHPAD privés à but lucratif	1 269	70	80	83
EHPAD privés à but non lucratif	1 623	68	78	81
EHPAD publics hospitaliers	766	60	69	73
EHPAD publics non hospitaliers	1 465	68	77	81
Résidences autonomie	1 465	52	68	71
EHPA	130	22	37	45
USLD	57	14	26	29,5
Ensemble	6 775	63	73	77

Champ > Individus vivant dans la sous-base d'étude.

Source > Fines 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Nombre de seniors à l'adresse

Par catégorie d'établissement, les premiers déciles et le premier quartile du nombre d'individus de plus de 70 ans déclarés à l'adresse sont détaillés dans le tableau A7.

Tableau A7 • Répartition du nombre de seniors de plus de 70 ans dans les adresses de la sous-base d'étude, par catégorie d'établissement

Catégorie	Nombre d'établissements	D1	D2	Q1
EHPAD privés à but lucratif	1 269	11	20	24
EHPAD privés à but non lucratif	1 623	6	14	19
EHPAD publics hospitaliers	766	7	16	20
EHPAD publics non hospitaliers	1 465	9	19,8	23
Résidences autonomie	1 465	5	12	15
EHPA	130	1	3	4,3
USLD	57	0	1	2
Ensemble	6 775	6	15	19

Champ > Individus de plus de 70 ans vivant dans un établissement de la sous-base d'étude.

Source > Fines 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Avec le deuxième décile, le nombre d'individus de plus de 70 ans déclarés à l'adresse est supérieur à 10 pour toutes les catégories. Si nous retenons le premier décile, nous devrions fixer un nombre de seniors à 5, ce qui paraît trop peu exigeant. Nous choisissons donc de garder le seuil indiqué par le deuxième décile, quitte à exclure plus d'établissements (car ils auraient moins de 10 personnes de 70 ans ou plus déclarées fiscalement à une même adresse), plutôt que de prendre le risque d'identifier trop facilement une adresse comme étant un établissement (car il y aurait plus de 5 personnes de 70 ans ou plus).

Concentration des seniors

Pour chaque établissement, on calcule le taux de senior à 70 ans, c'est-à-dire la part des plus de 70 ans sur l'ensemble des individus de l'adresse. Ensuite par catégorie d'établissement, les premiers déciles et le premier quartile sont donnés dans le tableau A8.

Tableau A8 • Répartition du taux de concentration des seniors dans les adresses, par catégorie d'établissement

Catégorie	Nombre d'établissements	D1	D2	Q1
EHPAD privés à but lucratif	1 269	82,1	90,9	92,7
EHPAD privés à but non lucratif	1 623	73,5	86,7	89,1
EHPAD publics hospitaliers	766	56,2	75	79,0
EHPAD publics non hospitaliers	1 465	75	86,2	88,7
Résidences autonomie	1 465	52,9	69,9	73,6
EHPA	130	14,8	26,5	36,5
USLD	57	0	2,1	2,8
Ensemble	6 775	64,7	80	83,8

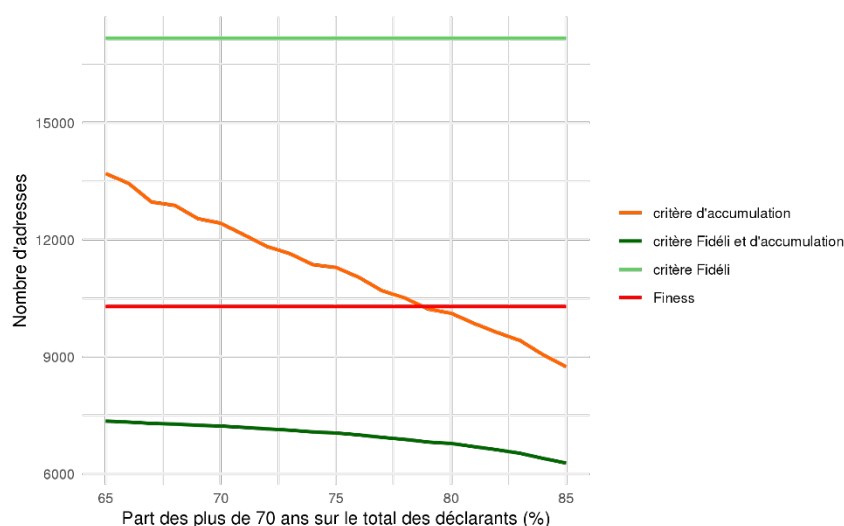
Champ > Individus vivant dans la sous-base d'étude.

Source > Fines 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Avec un seuil d'âge à 70 ans pour le taux de concentration, le premier décile indique qu'il faut prendre un taux de 50%, ce qui paraît très faible. Avec le deuxième décile, le seuil candidat est 69.9%.

Afin de valider ce seuil, une méthode complémentaire est mise en place. **Nous fixons comme premier critère « avoir au moins 10 individus de plus de 70 ans »** (c'est ce qui est conclu pour le choix des autres seuils) et nous cherchons à déterminer le taux de concentration. Nous pouvons faire varier ce taux, puis appliquer le critère d'accumulation avec ce taux sur l'ensemble des adresses déclarées dans la table Individu de Fidéli, pour regarder le nombre d'adresses qui sont identifiées.

Graphique A1 • Effets du seuil du taux de senior sur le nombre d'adresses identifiées



Champ > Adresses brutes de la table Individu de Fidéli, France métropolitaine et DROMI.

Source > Fines 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Nous cherchons à avoir un nombre d'adresses identifiées proche du nombre d'adresse Finess, mais supérieur car nous identifions également des lieux de vie accueillant des seniors sans être des établissements pour personnes âgées (comme les résidences-services seniors). Le taux doit donc être inférieur à 78 %. De plus selon le think tank *Matières Grises* (Kuhn Lafont, Troutot, 2022), il y a autour de 750 résidences-services seniors en 2019. Si on les somme au nombre d'établissements Finess pour personnes âgées, la barre du nombre d'établissements à viser croise la courbe orange autour de 74 %.

Nous choisissons un taux à 70%.

Synthèse

Le critère choisi est :

- Au moins 10 individus de plus de 70 ans à l'adresse ;
- Au moins 70 % des individus de l'adresse ont plus de 70 ans.

Ce critère est appliqué à l'ensemble des adresses de Fidéli, et à la sous-base d'étude qui a été utilisée pour le déterminer. Sur la sous base, 79,3 % des établissements sont identifiés par ce critère (pour 84,1 % des individus), ce qui est cohérent avec le fait que les seuils choisis correspondent à des niveaux indiqués par les 2^{ème} déciles (donc autour de 20 % des adresses sont perdues). Sur la base des adresses déclarées dans la table Individu de Fidéli, le critère identifie 12 421 adresses, dont 58,2 % sont également identifiées par *id scat* (critère Fidéli).

Annexe 4. Autres tableaux

Tableau A9 • Répartition dans les catégories d'établissement, établissements appariés de manière sûre avec Fidéli et établissements appariés de manière sûre avec Fidéli sans conflit d'intérêt d'adresse, niveau métropole et DROM

Catégorie	Nombre appariés à Fidéli	Répartition appariés à Fidéli (%)	Nombre sans conflit d'intérêt d'adresse	Répartition sans conflit d'intérêt d'adresse (%)
EHPAD privés à but lucratif	1 733	19,2	1 671	20,2
EHPAD privés à but non lucratif	2 185	24,3	2 050	24,8
EHPAD publics hospitaliers	1 091	12,1	968	11,7
EHPAD publics non hospitaliers	2 033	22,6	1 919	23,2
Résidences autonomie	1 700	18,9	1 496	18,1
EHPA	192	2,1	134	1,6
USLD	76	0,8	25	0,3
Ensemble	9 010	100,0	8 263	100,0

Champ > Établissements Finess appariés de manière sûre à Fidéli et établissements Finess appariés de manière sûre à Fidéli sans conflit d'intérêt d'adresse, France métropolitaine et DROM.

Source > Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Tableau A10 • Répartition dans les catégories d'établissement, établissements appariés de manière sûre avec Fidéli et établissements appariés de manière sûre avec Fidéli sans conflit d'intérêt d'adresse, niveau métropole

Catégorie	Nombre appariés à Fidéli	Répartition appariés à Fidéli (%)	Nombre sans conflit d'intérêt d'adresse	Répartition sans conflit d'intérêt d'adresse (%)
EHPAD privés à but lucratif	1 717	19,2	1 656	20,2
EHPAD privés à but non lucratif	2 163	24,2	2 030	24,8
EHPAD publics hospitaliers	1 075	12,0	953	11,6
EHPAD publics non hospitaliers	2 029	22,7	1 916	23,4
Résidences autonomie	1 699	19,0	1 496	18,2
EHPA	180	2,0	124	1,5
USLD	76	0,0	25	0,3
Ensemble	8 939	100,0	8 200	100,0

Champ > Établissements Finess appariés de manière sûre à Fidéli et établissements Finess appariés de manière sûre à Fidéli sans conflit d'intérêt d'adresse, France métropolitaine.

Source > Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Tableau A11 • Répartition dans les catégories d'établissement au départ, après appariement avec Fantoir, et après appariement avec Fidéli, métropole

Catégorie	Nombre départ	Répartition départ (%)	Nombre appariés avec Fantoir	Répartition appariés avec Fantoir (%)	Nombre appariés avec Fidéli	Répartition appariés avec Fidéli (%)
EHPAD privés à but lucratif	1 800	17,6	1 750	17,8	1 717	19,2
EHPAD privés à but non lucratif	2 332	22,9	2 251	22,8	2 163	24,2
EHPAD publics hospitaliers	1 191	11,7	1 142	11,6	1 075	12,0
EHPAD publics non hospitaliers	2 127	20,9	2 062	20,9	2 029	22,7
Résidences autonomie	2 284	22,4	2 219	22,5	1 699	19,0
EHPA	300	2,9	275	2,8	180	2,0
USLD	165	1,6	159	1,6	76	0,0
Ensemble	10 199	100,0	9 858	100,0	8 939	100,0

Champ > Établissements Finess de départ (établissements pour personnes âgées), établissements Finess appariés de manière satisfaisante avec Fantoir et établissements Finess appariés de manière sûre avec Fidéli, France métropolitaine.

Source > Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Tableau A12 • Scores d'appariement par catégorie d'établissement, métropole

Catégorie	Nombre	Répartition (%)	Part à être appariée de manière sûre (%)
EHPAD privés à but lucratif	1 800	17,6	95,4
EHPAD privés à but non lucratif	2 332	22,9	92,8
EHPAD publics hospitaliers	1 191	11,7	90,3
EHPAD publics non hospitaliers	2 127	20,9	95,4
Résidences autonomie	2 284	22,4	74,4
EHPA	300	2,9	60,0
USLD	165	1,6	46,1
Ensemble	10 199	100	87,6

Champ > Établissements pour personnes âgées dans Finess, France métropolitaine.

Source > Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

Tableau A13 • Étapes d'appariement par catégorie d'établissement, métropole

Catégorie	Non apparié avec Fantoir	Pas dans Fidéli	Aucun établissement	Les deux	Accumulation seulement	Fidéli seulement	Total
EHPAD privés à but lucratif	2,8	0,9	0,9	86,6	0,9	7,9	100
EHPAD privés à but non lucratif	3,5	1,8	1,8	76,3	2,7	13,8	100
EHPAD publics hospitaliers	4,1	1,4	4,2	69,3	7,5	13,5	100
EHPAD publics non hospitaliers	3,1	0,8	0,8	84,5	0,6	10,2	100
Résidences autonomie	2,8	4,9	17,9	11,9	60,1	2,4	100
EHPA	8,3	12,7	18,7	28,0	9,3	23,0	100
USLD	3,6	10,3	40,0	20,6	15,2	10,3	100
Ensemble	3,3	2,5	6,4	62,3	15,8	9,6	100

Champ > Établissements pour personnes âgées dans Finess, France métropolitaine.

Source > Finess 2019, Fantoir 2020, Fidéli 2020, calculs DREES.

DREES MÉTHODES

N° 9 • mars 2023

Identification des seniors en hébergements dédiés aux
personnes âgées dans les données Fidéli

Directeur de la publication

Fabrice Lenglard

Responsable d'édition

Valérie Bauer-Eubriet

ISSN

2495-120X

Ministère des Solidarités et de la Santé
Direction de la recherche, des études, de l'évaluation et des statistiques (DREES)

14 avenue Duquesne - 75 350 paris 07 SP
Retrouvez toutes nos publications sur drees.solidarites-sante.gouv.fr et nos données sur www.data.drees.sante.fr
