

L'enquête Feti'i e fenua (Enquête Famille, territoire et relations intergénérationnelles en Polynésie française) : Apurement et imputation des données

Leïla Fardeau, Éva Lelièvre et l'équipe ATOLLs
(Archipels, Territoires et MObilités FamiLiaLes)

Leïla Fardeau, Éva Lelièvre et l'équipe ATOLLS (Archipels, Territoires et MObilités FamiLiaLes), *L'enquête Feti'i e fenua (Enquête Famille, territoire et relations intergénérationnelles en Polynésie française) : Apurement et imputation des données*, Paris, Ined, Document de travail, 262



L'enquête *Feti'i e fenua* (Enquête Famille, territoire et relations intergénérationnelles en Polynésie française) : Apurement et imputation des données

Leïla Fardeau, Éva Lelièvre et l'équipe ATOLs (Archipels, Territoires et MObilités FamiLiaLes)¹

Résumé

Ce document de travail présente les opérations statistiques post-collecte relevant principalement de l'apurement et de l'imputation. Ces deux opérations essentielles de la production de données d'enquêtes quantitatives, relativement peu connues et documentées dans la littérature de sciences sociales, conditionnent la production du Fichier de Production et de Recherche (FPR), mis à disposition de la communauté scientifique pour l'analyse et la valorisation des résultats de l'enquête. A l'occasion des opérations de préparation des données de la première enquête *Famille* collectée sur le territoire de la Polynésie française (*Feti'i e fenua*, Ined-Ispf 2020), ce document de travail présente la codification des questions ouvertes (localisations et professions), l'apurement dont le principe est d'éliminer les incohérences internes aux différents modules et l'imputation qui pallie la non-réponse. L'enquête *Feti'i e fenua* inspirée de l'enquête *Famille et logements* associée au recensement réalisée en France métropolitaine, a été collectée dans 31 îles du territoire auprès de 5 139 répondants avec un remarquable taux de réponse de 86%. Les utilisateurs des grandes sources de données quantitatives disposent rarement du détail des choix qui ont prévalu à la création du fichier rendu disponible pour l'analyse. Pourtant la robustesse de ces analyses dépend autant de la qualité de la collecte que des traitements nécessaires pour homogénéiser l'ensemble des réponses.

Mots clefs

Enquête Famille, apurement, Polynésie française, imputation multiple

Summary :

This working paper presents the post-collection statistical operations relating mainly to the cleaning of the data and imputation of missing data. Relatively ignored and rarely documented in the social science literature, these essential operations in the production of quantitative survey data condition the Production and Research File made available to the scientific community for the analysis and promotion of the survey results. During the data preparation operations of the first Family survey collected on the territory of French Polynesia (*Feti'i e fenua*, Ined-Ispf 2020), this working paper presents the codification of open questions (locations and professions), the data cleaning which is undertaken to eliminate internal inconsistencies between the different modules and the imputation which compensates for non-response. The *Feti'i e fenua* survey, inspired by the *Family and Housing Survey* associated to the census carried out in metropolitan France, was collected in 31 islands in the territory from 5,139 respondents with a remarkable response rate of 86%. Users of large quantitative

¹ Wilfried Rault, Célio Sierra Paycha, Loïc Trabut, Géraldine Vivier, Patricia Thauvin (sur cette tâche).

data sources rarely have the details of the choices that prevailed when creating the file made available for analysis. However, the robustness of the analyses depends as much on the quality of the data collection as on the processing necessary to standardize all the responses.

Keywords:

Family survey, data cleaning, French Polynesia, multiple imputation

Les opérations de nettoyage de données, de correction de la non-réponse et de codification des questions saisies en clair (ici les professions et les localisations) donnent rarement lieu à des publications ; et si les méthodes statistiques de redressement trouvent leurs lecteurs dans les revues spécialisées de statistiques², les utilisateurs des grandes sources de données quantitatives disposent rarement du détail des choix qui ont prévalu à la création du fichier rendu disponible pour l'analyse. Pourtant la robustesse de ces analyses dépend autant de la qualité de la collecte (non-réponse, répartition de l'échantillon, qualité du questionnaire, nature des entretiens, etc.) que des traitements nécessaires pour homogénéiser l'ensemble des réponses. C'est donc dans cet objectif que ce document a été conçu et que la nature des différentes opérations a été détaillée.

A l'occasion des opérations de préparation des données de la première enquête *Famille* collectée sur le territoire de la Polynésie française (*Feti'i e fenua*, Ined-Ispf 2020), ce document de travail présente la codification des questions ouvertes, l'apurement dont le principe est d'éliminer les incohérences internes aux différents modules et l'imputation qui pallie la non-réponse.

L'ensemble des analyses qui seront produites à partir du FPR sont en effet conditionnées par ces opérations et les choix qui y ont présidé. Souvent invisibles, la robustesse des analyses produites en dépend pourtant. Aussi la description du recueil des informations, de leur traitement et des opérations d'apurement et le cas échéant d'imputation nous semblent devoir faire l'objet d'une présentation à laquelle les utilisateurs pourront se référer pour juger de la solidité des résultats produits à partir des données de la première enquête *Famille* conduite en Polynésie française.

Après avoir présenté (I) l'enquête et ses différents modules, une deuxième section (II) rendra compte des stratégies adoptées pour produire une codification harmonisée des variables de localisation collectées au long du questionnaire selon différents procédés. La section suivante (III) sera consacrée au traitement des données concernant les professions et l'activité en général. Enfin, concernant les dates qui jalonnent le questionnaire (IV), nous aborderons la question des données manquantes et de leur imputation.

² L'imputation, nécessaire pour redresser les biais de non-réponse, repose sur des procédures statistiques bien documentées en statistique voir par exemple Ford, 1983 ; Little, 1986 ; Deville et al., 1993 ; Caron, 2005 ; Haziza et Beaumont, 2007 ; Aldridge et Little, 2010 ; Chauvet et al., 2011.

I. Présentation de l'enquête

I- 1. Objectifs de l'enquête et caractéristiques de la collecte

Plusieurs grandes enquêtes ont été élaborées au début des années 2010 permettant l'étude des situations familiales et la dispersion spatiale des familles en France. Les équipes de l'INED ont été ainsi mobilisées en collaboration avec l'INSEE sur la conception, la collecte et l'analyse de la dernière enquête *Famille et logements* (<http://lili-efl2011.site.ined.fr/>) enquête associée au recensement réalisée en France métropolitaine ainsi que sur les enquêtes *Migrations famille et vieillissement* (<https://www.ined.fr/fr/recherche/projets-recherche/P0819>) portant sur quatre des départements d'Outre-mer. La Polynésie française n'étant pas dans le champ de ces dispositifs d'enquête, l'ISPF³ et le gouvernement de Polynésie française ont exprimé le besoin de disposer d'un éclairage sur la dispersion spatiale des familles polynésiennes, sur les relations entretenues entre leurs membres, ainsi que sur la mobilité des habitants du territoire, au sein de la Polynésie française et au-delà.

Un projet d'enquête a ainsi vu le jour, fondé sur la collecte d'informations géographiques, socio-économiques et relationnelles concernant les membres de la famille des enquêtés et le cas échéant, en symétrie, de la famille de leur conjoint. L'État et le Pays⁴ ont soutenu le projet d'enquête *Feti'i e fenua* (littéralement « Liens familiaux et territoire ») mis en œuvre par l'Ined et l'ISPF en 2017-2018 (tests) pour une réalisation effective en 2019-2020, en apportant la contribution financière décisive à l'opération.

L'enquête *Feti'i e fenua* (ou enquête *Famille, territoire et relations intergénérationnelles en Polynésie française*) s'appuie en partie sur l'enquête *Famille et logements* (Insee, 2011) et en propose une adaptation au territoire de Polynésie Française. Elle vise à éclairer la composition et la dispersion spatiale des familles à travers les lieux de résidences de leurs membres, y compris en dehors du territoire, ainsi que leurs occupations et les liens qu'ils entretiennent avec les lieux d'origine (lieu de naissance, île ou région de socialisation avant 6 ans, lieux de scolarité, etc.). La description de l'espace résidentiel de la famille est en effet un outil essentiel à la compréhension des dynamiques familiales aussi bien en termes de mobilités de ses membres que de solidarités entre eux. Il s'agit de décrire l'espace formé par les lieux de résidence de l'ensemble des membres de la famille qui ne sont pas nécessairement corésidents ainsi que la fréquence des contacts au sein de cet espace familial.

Ceci permet (1) d'obtenir une "cartographie" exhaustive de la répartition spatiale de la famille des enquêtés, en Polynésie et au-delà ; au moment de l'enquête, (2) d'identifier la situation de chacun en termes d'occupation et de situation familiale sur trois générations : parents, collatéraux, enfants de l'enquêté mais aussi de son conjoint le cas échéant. On a ainsi le portrait spatialisé de ces familles qui résulte et résume les migrations effectuées.

Le questionnaire est individuel et il y a un seul enquêté par ménage. Les enquêté.e.s sont âgé.e.s de 40 à 59 ans, des deux sexes. Ils constituent les « adultes pivots »⁵ entre des parents vieillissant et des

³ Institut de la Statistique de Polynésie française.

⁴ A la suite de la création du statut d'autonomie en 1984, le gouvernement de la Polynésie française constitue l'organe exécutif de la collectivité d'outre-mer (le Pays) de la République française (l'État).

⁵ Par adulte pivot, nous faisons référence aux générations enquêtées qui sont nées entre 1960 et 1979, soit en pleine transition démographique. Leurs parents sont potentiellement encore en vie car ils ont bénéficié des gains d'espérance de vie. Leurs frères et sœurs sont plus nombreux que leurs enfants en raison de la fécondité élevée de leurs parents par contraste avec leur propre fécondité.

enfants en âge de décohabiter. Le périmètre de la famille qui fait l'objet du recueil porte donc sur 3 générations.

La collecte s'est déroulée sur l'ensemble des 5 archipels dans 31 îles⁶ (sur les 74 habitées) auprès d'individus âgés de 40 à 59 ans tirés au hasard dans un échantillon de ménages issus du recensement de 2017. Sur 5 964 ménages échantillonnés, l'information sur les familles de 5 139 répondants a été collectée, ce qui représente un taux de réponse de 86% (72% à Tahiti).

La collecte de l'enquête *Feti'i e fenua* s'est déroulée en deux vagues : d'octobre à décembre 2019, une saison qui permet la tournée en bateau pour les îles isolées des Tuamotu, puis de février à mars 2020 où elle s'est interrompue lors de la première phase épidémique du covid-19.

Les données ont été recueillies grâce à un système de collecte assistée par ordinateur (CAPI pour *Computer-Assisted Personal Interview*) consistant à équiper les enquêtrices d'ordinateurs portables munis d'une application informatique (le DataModel développé par Laurent Rouquier de l'ISPF) permettant de saisir immédiatement les informations recueillies en face à face. Elles ont ensuite été transmises à l'ISPF via un système de télétransmission sécurisé.

Les données brutes ont été livrées en deux fois à la fin des deux vagues. Leïla Fardeau a été chargée de l'apurement sous la direction des membres de l'équipe : Éva Lelièvre, Celio Sierra Paycha et Loïc Trabut. La codification des nombreuses activités professionnelles collectées a été préparée par Laurent Rouquier et Laurent Pellet de l'ISPF et supervisée à l'Ined par Géraldine Vivier et Patricia Thauvin du Service des Enquêtes et des Sondages.

I- 2. Prise de contact et tirage de la personne enquêtée parmi les éligibles

Les 32 enquêtrices disposaient de fiches-adresse résumant les coordonnées de contact⁷ des ménages échantillonnés à partir du recensement de 2017, et dont au moins un membre était âgé de 40 à 59 ans l'année de la collecte. L'introduction du formulaire CAPI permettait d'identifier les personnes éligibles du ménage et de tirer aléatoirement selon une méthode Kish⁸, la personne à enquêter.

Exceptionnellement, si la personne désignée ne pouvait répondre à l'enquête (indisponible, absente le temps de l'enquête...), son conjoint – *et exclusivement son conjoint* – pouvait répondre à sa place. Effectivement, du fait de la « symétrie » du questionnement sur les informations familiales (cf. figure 1), il a pu être toléré qu'un conjoint plus disponible –souvent une femme- réponde à la place de la personne initialement tirée au sort. Cela s'est produit dans 824 cas. Une variable permet d'identifier ces questionnaires (cf. Annexe A1).

I- 3. Structure du questionnaire

Le questionnaire est structuré en modules thématiques et/ou correspondant aux membres de la famille (voir le questionnaire reproduit en Annexe).

On procède à la collecte des informations en commençant par celles qui concernent la personne enquêtée, puis son conjoint. Le module suivant porte sur la description du logement puis des biens

⁶ <http://www.ispf.pf/bases/recensements/2012/Presentation/Geographie/Guide.aspx>

⁷ Il n'y a pas d'adressage en Polynésie française, le courrier est distribué via des boîtes postales.

⁸ Cette méthode de sélection aléatoire de l'individu à enquêter par tirage parmi les individus éligibles, porte le nom de son concepteur, le statisticien américain Leslie Kish (1965).

fonciers de chacun des membres du couple (ces éléments apportant également des informations sur les lieux d'origine et les migrations anciennes). On procède ensuite à la description des autres membres de la famille avec un module parents et parents du conjoint, puis un module dédié aux collatéraux (frères et sœurs) et enfin aux enfants de l'enquêté-e. Un court module sur les autres personnes considérées comme proches complète le panorama. Enfin, le questionnaire se termine par la reconstitution des membres du ménage, sous forme d'un tableau de composition du ménage (TCM).

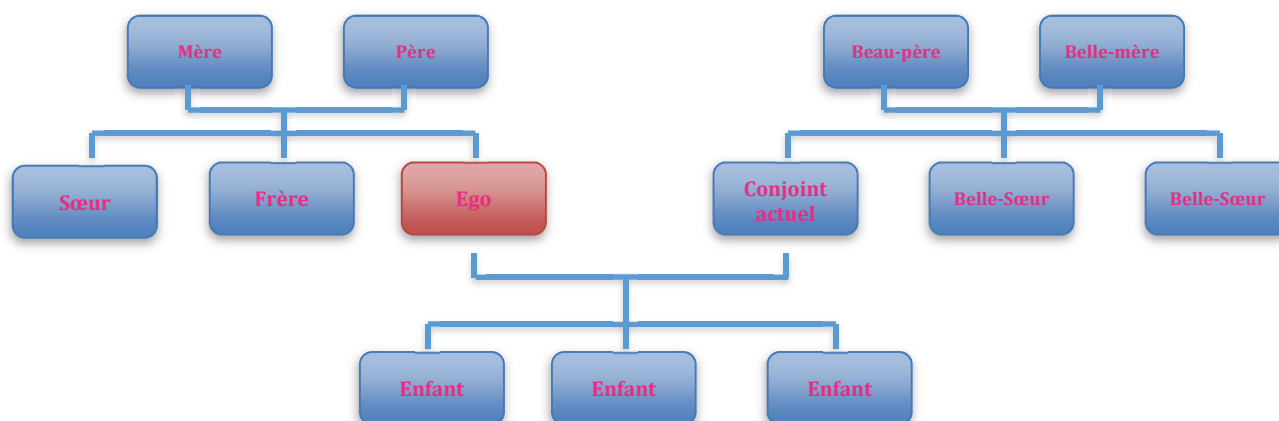


Figure 1 Univers familial parcouru au cours du questionnaire

I- 3.1 Le repérage des localisations

La localisation de différents lieux est une information collectée en maints endroits du questionnaire, depuis les lieux de naissance, les lieux de résidence (à différents moments), les lieux de scolarisation, les lieux de séjour en dehors de la Polynésie, etc. Et cela pour les divers membres de l'entourage familial et avec des précisions diverses (tableau 1).

Tableau 1 Récapitulatif des différents lieux documentés dans *Feti'i e fenua*

La personne enquêtée	Son conjoint	Parents Parents du cjt	Fratrie Fratrie du cjt	Enfants de l'enquêté et/ou de son cjt	Proches
Naissance	Naissance	Naissance	Naissance	Naissance	Naissance
Résidence Mère		Résidence	Résidence	Résidence	Résidence
Résidence Conjoint					
Lieu Install Union	Lieu Install Union				
Ecole Primaire*	Ecole Primaire*			Ecole Primaire*	
Collège*	Collège*			Collège*	
Lycée*	Lycée*			Lycée*	
EcoleSup*	EcoleSup*			EcoleSup*	
Séjours hors PF*	Séjours hors PF*		Résidences d'enfants hors PF	Séjours hors PF*	
Bien(s) foncier(s)*	Bien(s) foncier(s)*				

Note : Les lieux assortis d'une * sont localisés plusieurs fois si la personne a fréquenté plusieurs établissements d'enseignements ou a séjourné plusieurs fois ou en différents lieux hors de Polynésie française.

À l'exception des lieux de séjour hors de la Polynésie française, la précision est effectuée en plusieurs temps. Un premier niveau distingue quatre entités : la Polynésie française (A), une autre collectivité ou département d'outre-mer (B), la France métropolitaine (C) et un autre pays (D).

1. Si la localisation relève de la Polynésie Française, alors des précisions sont demandées sur l'île, la commune associée et la commune (voir la figure 2a). Il y a 48 communes sur le territoire, 98

communes associées dans les 31 îles enquêtées. Les grandes îles peuvent comporter plusieurs communes, mais certaines communes regroupent plusieurs petites îles ;

Figure 2a Les localisations dans le DataModel (précisions administratives en Polynésie française)

Comas	Commune	Ile
Tehurui	Tumaraa	Raiatea
Tepoto Nord	Napuka	Tepoto Nord
Tevaitoa	Tumaraa	Raiatea
Tiarei	Hitiaa O Te Ra	Tahiti
Tikehau	Rangiroa	Tikehau
Toahotu	Taiaapu-Ouest	Tahiti
Tureia	Tureia	Tureia
Ua Huka	Ua Huka	Ua Huka
Uturoa	Uturoa	Raiatea
Vahitahi	Nukutavake	Vahitahi
Vaiaau	Tumaraa	Raiatea
Vairaatea	Nukutavake	Vairaatea
Vairao	Taiaapu-Ouest	Tahiti
Vaitoare	Tahaa	Tahaa
Vaiuru	Raivavae	Raivavae

Search: VAIU

116:116

Select Cancel

2. Si la localisation renvoie à des territoires ultra-marins, une précision est demandée sur ce territoire (figure 2b) ;

Figure 2b Les localisations dans le DataModel (identification des DROM)

Q06. Où résidait ta mère juste après ta naissance ?
Précisez la zone.

☐ 1. En Polynésie française

☒ 2. Dans une autre collectivité ou département d'outre-mer

☐ 3. En France métropolitaine

☐ 4. Dans un autre pays

Q06. Où résidait ta mère juste après ta naissance ?
Précisez le département ou la collectivité d'outre-mer

☐ 1. 971 - GUADELOUPE	☐ 10. 986 - WALLIS ET FUTUNA
☐ 2. 972 - MARTINIQUE	☐ 11. 988 - NOUVELLE-CALEDONIE
☐ 3. 973 - GUYANE	☐ 12. 989 - ILE DE CLIPPERTON
☐ 4. 974 - REUNION	
☐ 5. 975 - SAINT-PIERRE ET MIQUELON	
☐ 6. 976 - MAYOTTE	
☐ 7. 977 - SAINT-BARTHELEMY	
☐ 8. 978 - SAINT-MARTIN	
☐ 9. 984 - TERRES AUSTRALES ET ANTARCTIQUES FRANCAISES	

3. Lorsqu'il s'agit d'une localisation en France métropolitaine, une précision est demandée au niveau du département. Si elle n'est pas connue, un espace libre est proposé (figure 2c) afin que la

personne enquêtée puisse donner la localisation qu'elle connaît (la ville dans l'exemple de la figure 2c) ;

Figure 2c Les localisations dans le DataModel (identification dans l'Hexagone)

Q06. Où résidait ta mère juste après ta naissance ?
Précisez la zone.

☐ 1. En Polynésie française
☐ 2. Dans une autre collectivité ou département d'outre-mer
☒ 3. En France métropolitaine
☐ 4. Dans un autre pays

Enter a text of at most 30 characters

LibreDEPT	LibPays
LibDEPT	Pays
AnneeInstall 1964	DomCom
Zone 3 FR	DEPT ?
Ile	LibreDEPT BORDEAUX

4. Enfin, lorsque la localisation se situe à l'étranger, une nomenclature par pays est proposée, enrichie d'entités permettant de distinguer certains territoires éloignés de leurs pays (figure 2c), par exemple : Hawaï des Etats-Unis, (2). Une déclaration libre est possible lorsque la localisation est approximative.

Figure 2d Les localisations dans le DataModel (identification des régions/pays du monde-liste interactive)

D:\DM a finaliser\Referentiels\PAYS - Blaise 4.8 Database Browser

Nom	Code	CodeNum	CodeAlpha
GIBRALTAR	044	292	GIB
GRANDE-BRETAGNE	013	914	GBR
GRECE	009	300	GRC
GRENADE	473	308	GRD
GROENLAND	406	304	GRL
GUADELOUPE	458	312	GLP
GUAM	818	316	GUM
GUATEMALA	416	320	GTM
GUERNESEY	901	831	GUE
GUINEE EQUATORIALE	310	226	GNQ
GUINEE-BISSAO	257	624	GNB
GUYANA	488	328	GUY
GUYANE FRANÇAISE	496	254	GUF
HAITI	452	332	HTI
▶ HAWAII	999	999	HAW
HONDURAS	424	340	HND

Search:
HAWAII

99:292

Select Cancel

De plus, afin d'alléger la collecte des localisations des membres de la famille, il était d'abord demandé si la personne habitait « avec l'enquêté-e » ou « à côté », auquel cas l'information sur la localisation était dupliquée lors de l'apurement (section II.1).

I- 3.2 Les professions et activités

La situation au regard de l'activité, la profession et le statut ont systématiquement été collectés pour l'enquête et l'ensemble des membres de sa famille : conjoint, parents et beaux-parents, fratrie et belle-fratrie, enfants, soit près de 80 000 situations renseignées, selon des degrés de précision variables. Le tableau 2 récapitule les modalités de collecte, adaptées selon les liens de parenté.

Tableau 2 Récapitulatif des différentes variables de profession dans *Feti'i e fenua* selon la personne concernée

La personne enquêtée et son conjoint

14. Du point de vue du travail...

- ☐ Tu⁹ as un travail actuellement, quotidien ou pas ☐ tu cherches un travail
☐ Tu ne travailles pas ☐ Tu es retraité(e) ☐ Tu n'as jamais travaillé

15. Quel est ton travail ou le dernier travail que tu as exercé ?.....

16. Il s'agit d'un travail rémunéré comme....

1. ☐ employé.e du privé 2. ☐ employé.e du public 5. ☐ autre situation
3. ☐ employeur 4. ☐ indépendant.e, à votre compte

Les parents

- ☐ il/elle a un travail rémunéré ☐ il/elle ne travaille pas ☐ il/elle est retraité-e

Quel est ou était son (dernier) travail ?

Est (était-il)

1. ☐ employé.e du privé 2. ☐ employé.e du public 5. ☐ autre situation
3. ☐ employeur 4. ☐ indépendant.e, à votre compte

Les frères et sœurs

1. a un travail, quotidien ou pas 2. cherche un travail 3. ne travaille pas, au foyer 4. retraité(e)

Quel est ou était son (dernier) travail ?

Est (était-il)

1. ☐ employé.e du privé 2. ☐ employé.e du public 5. ☐ autre situation
3. ☐ employeur 4. ☐ indépendant.e, à votre compte

⁹ Le tutoiement est de rigueur en Polynésie Française, à l'exception de certains contextes particulièrement formels. L'ensemble du questionnaire est donc formulé à la deuxième personne du singulier.

Tableau 2 (suite) Récapitulatif des différentes variables de profession dans *Feti'i e fenua* selon la personne concernée

Les enfants			
1. est étudiant ou écolier	2. a un travail, quotidien ou pas	3. cherche un travail	4. ne travaille pas
5. autre			
Quel est ou était son (dernier) travail ?			
Est (était-il)			
1. ☐ employé.e du privé	2. ☐ employé.e du public	5. ☐ autre situation	
3. ☐ employeur	4. ☐ indépendant.e, à votre compte		

Selon les méthodes de codage expérimentées (SICORE¹⁰ ou méthode ISPF, nous y reviendrons), 15% à 17% des 79 876 libellés recueillis en clair s'avèrent incodables automatiquement et nécessitent donc d'être examinées individuellement et codées à la main. La difficulté de codification de ces libellés tient en partie à des spécificités polynésiennes, ou à l'imprécision des énoncés, collectés auprès d'un interlocuteur unique (proxy) amené à renseigner l'activité professionnelle d'un grand nombre de personnes.

De fait, les libellés non reconnus peuvent indiquer, en redondance avec la situation d'activité, que la personne n'a probablement jamais travaillé (« sans profession », « malade » par exemple) ou se référer à une profession, un type d'activité ou une appellation typiquement polynésien (« patente », « roulotte », « faa'apu »¹¹, « employé chez untel »...). Dans d'autres cas, le libellé mentionne une activité imprécise (« fonctionnaire », « ouvrier » ...) ou polysémique (« plongeur », « marin », ...) ou encore, une situation qui laisse présumer que la personne a travaillé mais sans que sa (dernière) profession n'ait été déclarée (« retraite », « cherche un emploi », « ne sait plus », ...). A cela, s'ajoute les erreurs de frappe et fantaisies orthographiques qui peuvent contrarier la reconnaissance de l'énoncé (« house keeping », « coprhacoulteur » ...).

I- 3.3 Les dates

Tout au long du questionnaire s'échelonnent les années de naissance et éventuellement de décès des différents membres de la famille ainsi que des repères temporels intermédiaires tels que :

- (1) l'âge ou l'année d'arrivée en Polynésie française pour les individus nés en dehors du territoire,
- (2) et l'âge ou l'année de départ pour des séjours de plus de 6 mois hors de Polynésie française.

De façon attendue, une bonne partie de ces dates sont imprécises (tableau 3) et cette imprécision est liée au type de lien familial. Les dates concernant les frères et sœurs du conjoint sont par exemple les plus fréquemment manquantes.

¹⁰ SICORE (Système Informatique de CODage des Réponses aux Enquêtes) est un système de codification automatique de libellés qui utilise un fichier de référence appelé fichier d'apprentissage (ou base de connaissance) lui servant d'exemples de chiffrage. Le libellé à chiffrer est ainsi comparé aux libellés contenus dans le fichier d'apprentissage et, lorsque le libellé est reconnu, le code associé est retenu.

¹¹ Le faa'apu se réfère à une activité d'autosubsistance et/ou de vente d'une petite production agricole domestique.

**Tableau 3 Fréquence des valeurs manquantes ou aberrantes
concernant les dates de naissance des membres de la famille**

	Nombre de valeurs manquantes	Nombre total d'observations	Pourcentage de valeurs manquantes
Mère	293	5141	5,70%
Père	481	5139	9,36%
Beau-Père	1055	3788	27,85%
Belle-Mère	866	3788	22,86%
Frères/sœurs d'Ego	5182	29257	17,71%
Frères/sœurs du cjt	5733	19452	29,47%
Enfant des deux	65	11363	0,57%
Enfant du conjoint	52	1935	2,69%
Enfant d'Ego	68	1302	5,22%

Source : données brutes *Feti'i e fenua* 2019-2020.

Concernant les dates de séjour en dehors de la Polynésie française ainsi que la date d'installation des personnes nées hors du territoire, les filtres et conditions appliquées par le DataModel ont garanti une qualité acceptable des dates et échéances déclarées.

I- 4. Différents types de variables brutes nécessitant des traitements distincts

La méthode de collecte CAPI (*Computer-Assisted Personal Interview*) permet de valider et faire des contrôles de cohérence automatiquement en temps réel au fur et à mesure de la saisie durant l'entretien. De plus, le DataModel (l'application informatique) offre la possibilité de présenter l'assistance nécessaire au chiffrement rapide des questions à nomenclatures via des menus déroulants et des fenêtres pop-up. De ce fait, les opérations d'apurement (tests de cohérence interne aux modules et externe entre modules) sont largement simplifiées.

Néanmoins, afin de faciliter les passations et ne pas alourdir l'interrogation en multipliant les vérifications bloquantes, on est amené à assouplir les critères de remplissage de l'information (par exemple des critères qui exigeraient qu'on insiste auprès des personnes interrogées pour qu'elles fournissent des dates même si la réponse est incertaine, ou encore en exigeant le département d'une résidence alors que la personne ne connaît que le nom de la localité). Cette souplesse nécessaire pour harmoniser la passation du questionnaire introduit du flou. Il subsiste ainsi des vérifications à entreprendre.

D'une manière générale, dans l'enquête *Feti'i e fenua*, les variables brutes peuvent être scindées en deux catégories nécessitant des traitements différents : des libellés en clair (remarques générales et associées aux variables de localisation mais surtout les activités professionnelles) ; et ensuite des variables numériques (dates, âges, effectifs).

Ces différents types de variable ne renvoient pas aux mêmes difficultés d'apurement.

II. Traitement des variables de localisation géographique

Chaque localisation est repérée à l'aide des caractéristiques détaillées dans le tableau 4, seule une variable est en saisie libre et demandera une codification spécifique de reconnaissance des lieux décrits avec une précision déclarative variable.

Tableau 4 Structure des variables de localisation dans le fichier brut (lieu de résidence)

ZoneResidence	PF = En Polynésie française OM = Dans une autre collectivité ou département d'outre-mer FR = En France métropolitaine ET = Dans un autre pays	<i>filtre</i>
IleResidence	Référentiel des îles de PF	Polynésie française
ArchipelResidence	Référentiel des archipels de PF	
ComasResidence	Référentiel des communes associées de PF	
CommuneResidence	Référentiel des communes de PF	
DomComResidence	Référentiel des DOM COM	OM
DeptResidence	Référentiel des départements de France	FR
PaysResidence	Référentiel des pays du monde	ET
DeptLibreResidence	<i>Saisie libre</i> (région, ville, zone, etc)	FR
AutreLieuResidence	<i>Saisie libre</i>	<i>Remarque</i>

Deux opérations caractérisent l'apurement des variables de localisation :

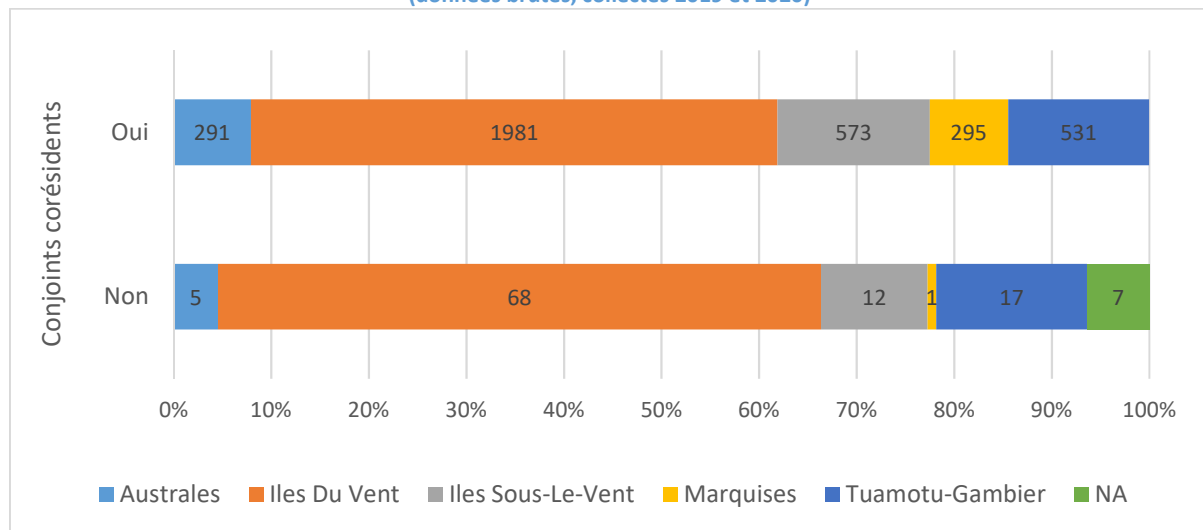
1. la première concerne uniquement les lieux de résidence, il s'agit de transférer les caractéristiques de la résidence de l'enquêté aux membres de la famille qui réside dans le même logement ou à proximité ;
2. la seconde consiste à coder les lieux décrits à partir des indications en saisie libre.

II- 1. Lieux de résidence des personnes corésidentes ou résidant à proximité

La première étape est le repérage des membres de la famille qui résident avec la personne enquêtée ou à proximité auxquels il faut attribuer les caractéristiques du lieu de résidence de la personne enquêtée dans les variables qui à l'état brut sont manquantes.

Selon le degré de parenté, la fréquence (faite sur les données brutes) de ces réponses varie (figures 3, 4, 5, 6).

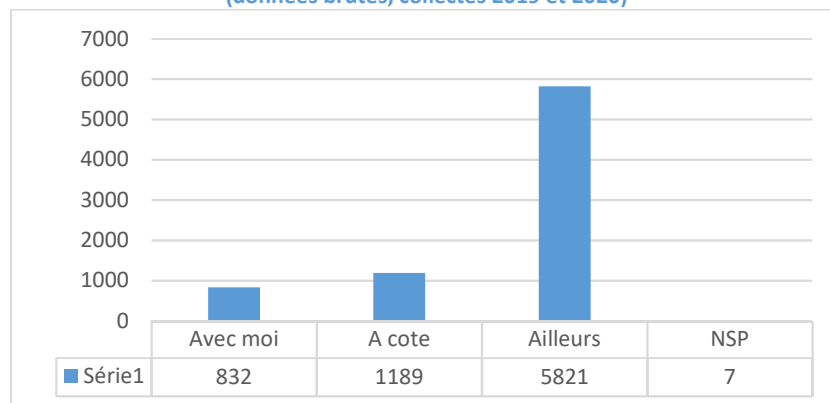
Figure 3 Répartition de la zone de résidence du conjoint en fonction du statut de résidence
(données brutes, collectes 2019 et 2020)



Note : Parmi les conjoints non corésidents 3 résident en Outre-Mer, 3 en Métropole et 2 à l'étranger
Source : données brutes Feti'i e fenua 2019-2020.

Un certain nombre de conjoints ne résident pas ensemble, souvent du fait des conditions de travail (travail en mer, conjoint en études hors de la Polynésie, emploi sur une île différente, etc.), c'est le cas de 4% des couples enquêtés (figure 3).

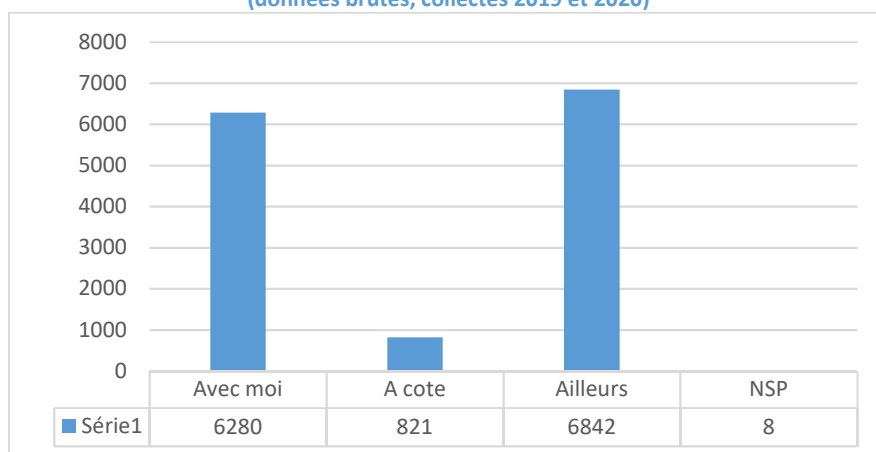
Figure 4 Corésidence des parents vivants (7 849) avec l'enquêté-e
(données brutes, collectes 2019 et 2020)



Source : données brutes Feti'i e fenua 2019-2020.

Plus d'un quart des parents (et beaux-parents) des enquêtés résident à proximité ou avec l'enquêté. On est en présence d'une importante corésidence intergénérationnelle. Les précisions de localisation doivent donc être reportées pour ces personnes.

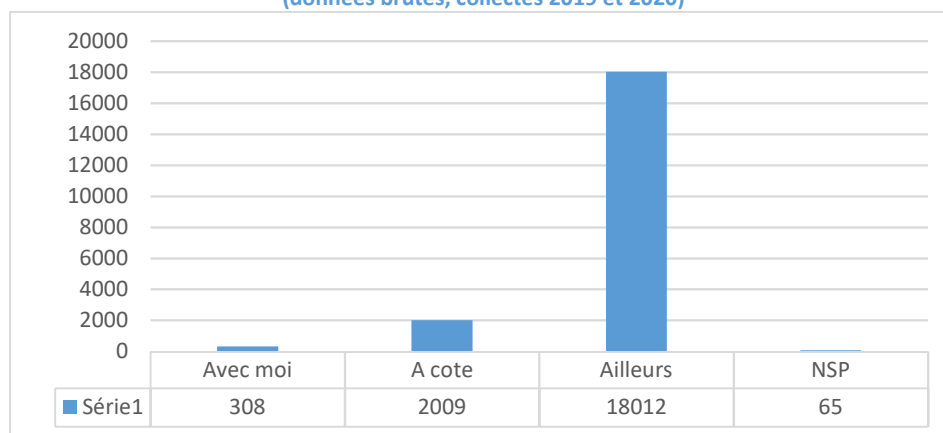
Figure 5 Corésidence des enfants vivants (13 951) avec l'enquêté-e
(données brutes, collectes 2019 et 2020)



Source : données brutes *Feti'i e fenua* 2019-2020.

Les enfants des personnes enquêtées ont des âges qui s'échelonnent de 1 à plus de 30 ans. La moitié d'entre eux résident avec l'enquêté ou à proximité. Les précisions de localisation doivent donc être reportées pour ces personnes.

Figure 6 Corésidence avec les frères et sœurs (20 394) de l'enquêté-e
(données brutes, collectes 2019 et 2020)



Source : données brutes *Feti'i e fenua* 2019-2020.

Quant aux frères et sœurs des enquêtés, les précisions de localisation doivent être reportées plus rarement car seuls 11% d'entre eux résident avec l'enquêté-e ou à proximité.

II- 2. Codification des localisations en saisie libre

La seconde étape consiste à coder le texte rempli dans la variable **DeptLibreResidence** et dans les remarques **AutreLieuRésidence** qui sont les 2 variables en saisie libre du tableau 4).

La première variable **DeptLibreResidence** a été renseignée lorsque la personne enquêtée ne connaît pas le numéro du département d'une localisation en Métropole.

700 modalités y figurent et la plupart d'entre elles ont été recodées automatiquement :

- **607 modalités** sont relatives à un ou des départements (plus ou moins précisément), d'autres pourraient **correspondre à plusieurs villes** (Fontenay par exemple) et d'autres encore correspondent à une **province historique** et sont donc moins précis que le département mais davantage que la région administrative actuelle.
- **27 modalités** ne font pas référence à la France Métropolitaine, mais à l'**étranger**, ou aux territoires des **Com ou Dom** ou encore aux **Communes** de la Polynésie française,
- **55 modalités** dont il est impossible d'extraire une quelconque information de localisation et que l'on code en **NA** (*non applicable*).

Concernant la variable **AutreLieuRésidence** qui recueille des remarques, 133 ont été recodées dans les variables existantes (30 remarques qui ne concernent pas les lieux).

Cette codification a amené :

1. à préciser le nom du département (variable **DeptResidence**)
2. ainsi qu'à la création de la variable **RegionResidence** lorsque la précision départementale n'a pu être déduite.

La variable **RegionResidence** contient les **13 modalités** (Auvergne- Rhône- Alpes, Bourgogne-Franche Comte, Bretagne, Centre-Val de Loire, Corse, Grand Est, Hauts de France, Ile de France, Normandie, Nouvelle Aquitaine, Occitanie, Pays de la Loire, Provence-Alpes-Côte d'Azur) qui font référence à la Métropole. Cette variable est ensuite renseignée et ajoutée pour tous les individus ayant un département de résidence.

Une codification similaire a été appliquée à l'ensemble des variables en saisie libre concernant le repérage d'une localisation.

III. La codification des professions et activités

Dans l'enquête *Feti'i e fenua*, 79 876 libellés de profession sont renseignés : il s'agit des professions des enquêtés, de leur conjoint, de leurs parents / beaux-parents, de la fratrie / belle-fratrie et leurs enfants.

- 5071 « Ego »
- 3718 « Conjoint »
- 11817 « Enfants »
- 16750 « Parents »
- 25722 « Fratrie »
- 16828 « Belle-Fratrie ».

Lors de la collecte, le recueil est fait avec trois questions présentées dans le tableau 2 sur lesquelles la post-codification va s'appuyer. Les variables correspondantes (Annexe A5) décrivent successivement la situation de la personne, variable **Travail** (en activité, au chômage, inactive, retraitée), la profession en clair (**DernierTravail**) puis le **StatutTravailleur** (salarié du public, du privé, indépendant).

Deux méthodes de post-codification ont été utilisées et comparées :

- une méthode de *matching* avec les professions décrites, et leur distribution, dans le recensement polynésien de 2017. Cette méthode est basée sur la distance de Jaro-Winkler proposé par les collègues de l'ISPF. Elle s'appuie sur les professions codées du recensement de la population et les variables **DernierTravail** (le libellé en clair) et **StatutTravailleur** scindé en **Statut** (salarié, indépendant) et **PublicPrivé** (secteur public ou privé),

et

- une méthode de codage automatique des professions via le logiciel SICORE, faite à l'Ined ; elle s'appuie également sur les variables **DernierTravail** (le libellé en clair) et **StatutTravailleur**.

III- 1. *Matching* de Jaro-Winkler

La distance de Jaro-Winkler mesure la dissimilarité entre deux chaînes de caractères en se fondant sur le nombre de transpositions de caractères nécessaires pour passer d'une chaîne à l'autre (Jaro, 1989 ; Winkler, 1999). Le résultat est normalisé de façon à avoir une mesure entre 0, dénotant l'égalité des chaînes comparées, et 1 qui représente l'absence de similarité. Cette mesure est particulièrement adaptée au traitement de chaînes courtes comme des noms et donc bien adaptée à la comparaison des libellés de profession telle que saisies dans **DernierTravail**.

Puis on effectue un tirage avec probabilité d'inclusion en fonction des variables annexes (**Statut** et **PublicPrivé**) quand on est dans un triplet non unique, c'est-à-dire lorsque plusieurs codifications correspondent au même triplet [**DernierTravail**, **Statut**, **PublicPrivé**] dans les données du recensement.

Il s'agit donc d'un appariement avec les professions en saisie libre du recensement de la population de 2017 (statut + profession en saisie libre pour être précis) puis d'un tirage aléatoire entre les différentes codifications possibles lorsque l'appariement donne lieu à plusieurs solutions. Les probabilités de tirage sont proportionnelles à la présence de chaque code dans les données du recensement.

Un des inconvénients majeurs de cette méthode est le fait que, même si l'on tient compte du statut (Indépendant / salarié et Public / privé) dans le *matching* initial, le tirage écrase complètement ce dernier.

On peut faire varier la précision du *matching* Jaro-Wrinkler grâce à un paramètre fixant la distance maximum à laquelle on peut apparier. Plus l'appariement est précis (donc la distance petite), plus le nombre d'observations non-codées sera important. La courbe ci-dessous (Figure 7) représente le nombre de professions codées par l'algorithme selon que l'on fait varier le paramètre de distance de 0,025 à 0,2.

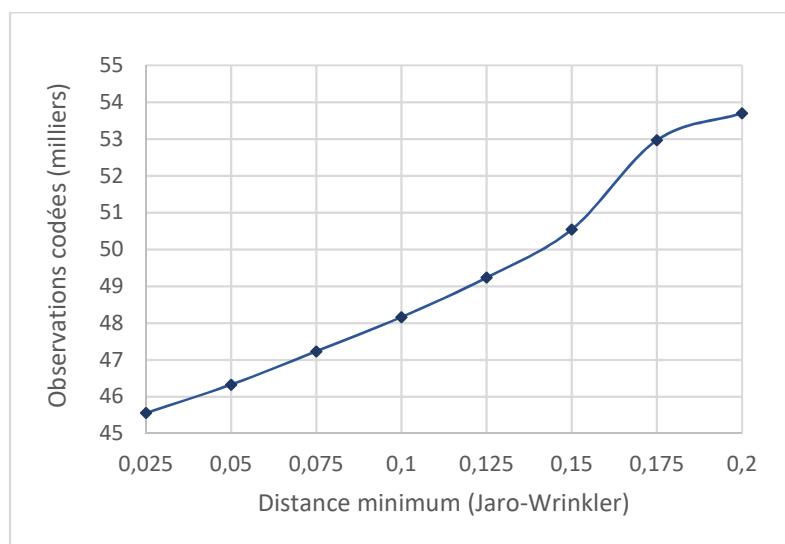


Figure 7 Nombre d'observations codées selon la distance choisie, ensemble des professions (tous modules confondus)

Source : libellés de professions renseignées dans les données brutes *Feti'i e fenua* 2019-2020.

Le nombre total de professions saisies, tous modules confondus, est de 79 873, parmi lesquelles 20 695 individus sont considérés non concernés : il s'agit des individus inactifs, étudiants, etc. Plus précisément la liste des modalités exclues rassemble :

- i. Les inactifs divers « AUCUN », « PAS DE TRAVAIL », « AU FOYER », « COTOREP », « EN PRISON », « S'OCCUPE DE SES PARENTS », etc. : recodés en 82
- ii. Les étudiants : recodés en 84
- iii. Les non concernés (trop jeunes ou décédés sans avoir jamais travaillé) : recodés en 0
- iv. Les « ne sait pas » (et équivalent) : recodés en 99.

Il reste donc 59 096 observations à coder. Le choix du paramètre de distance maximale repose sur deux critères : on souhaite coder un maximum de professions – nombre qui augmente avec la distance (Figure 7), tout en respectant un critère de qualité des codages qui nécessite quant à lui de limiter cette distance.

L'examen de la qualité des codages fournis par cet algorithme aux distances maximales de 0,08 et 0,1 et 0,15 ont permis de choisir la distance de 0,1 qui permet de coder un peu plus de 80% des observations.

III- 2. SICORE

Pour mettre en œuvre le programme SICORE un certain nombre d'opérations préalables doivent être menées

1. La suppression d'observations

Les individus pour lesquels le libellé (variable **DernierTravail**) ne correspondait pas à une profession ont été exclus de la base et n'ont pas fait l'objet d'un codage via SICORE. Parmi ces libellés, il y a notamment : au foyer (avec toutes ces déclinaisons possibles), écolier / collégien / lycéen / étudiant, décédé, etc.

Par cette opération 20 780 observations ont été écartées.

2. Le nettoyage du libellé de la profession

Sur les 59 096 observations restantes, un nettoyage du libellé de la profession (variable **DernierTravail**) a été fait, telles que la suppression des accents, la correction des fautes d'orthographe, etc.

3. La création de variables annexes

Grâce aux données collectées, deux variables SICORE ont pu être créées **Statut** qui distingue les employés des indépendants et **Pub** qui spécifie le secteur privé ou public de l'activité à partir de la variable **StatutTravailleur**.

A l'issue du recours à SICORE, on obtient les résultats présentés dans le tableau 5:

Tableau 5 – Distribution des retours de la procédure SICORE appliquée aux professions de l'enquête *Feti'i e fenua*)

code_retour	Signification	Fréquence	Pourcentage
C_C	Echec -Non reconnaissance du libellé	10 494	17.76
C_R	Echec partiel – Eventuel codage pouvant être considéré comme douteux	2 803	4.74
RC*	Réussite sous réserve de qualité – Réussite mais une (ou plusieurs) variable annexe nécessaire était à valeur manquante	45 799	77.50
TOTAL	-	59 096	100.0

Parmi les RC* (c'est-à-dire les réussites sous réserve de qualité), on ne compte en réalité que 44 506 observations avec un codage PCS attribué, puisque 1 293 observations se sont vues affectées un code rebut « T-HORS ». SICORE permet donc de coder un peu plus de 75% des observations. Par ailleurs il fournit un codage pouvant être considéré comme « douteux » pour un peu moins de 5% d'entre elles (code retour C_R).

Ce codage automatique des professions échoue donc à coder environ 20% des observations (code retour C_C et observations codées HORS). On ne compte enfin aucun cas de réussite « totale », avec ou sans utilisation des variables annexes disponibles (code_retour=(CCS,RCS)).

III- 3. Comparaison

D'un point de vue quantitatif, on observe peu de différences dans les performances des différentes méthodologies : la procédure par *matching* Jaro-Winkler permet un codage automatique d'environ 80% des professions (avec une précision choisie de 0,1) tandis que la procédure par SICORE permet de coder 75% des professions (avec un code retour RC : « sous réserve de qualité »). L'examen des différentes catégories d'échecs (tableau 6 ci-dessous) corroborent des volumes de « déchets » très comparables et va donc dans le même sens.

Tableau 6 – Comparaison des retours des deux procédures de codification appliquées aux professions de l'enquête *Feti'i e fenua*

	"HORS"	vides	Non concernés ?	Libellés
Codif Jaro-Winkler	1 006	11 100	20695 (codé 82, femme au foyer, nan... ou 84 : étudiant)	79 873
	12 106			
Codif SICORE	2 803 (échec partiel, douteux)	11 787 (échec)	20780 (PAS DE PROF : femme au foyer, nan, retraité...)	79 876
	13 890			

SICORE ne code pas plus mais code-t-il mieux ?

En premier lieu, SICORE évite le problème – précédemment signalé et relativement fréquent – d'incohérence entre la profession, son statut déclaré dans l'enquête et la codification codée par *matching* Jaro-Winkler. Pour rappel, bien que le statut (Indépendant / Salarié et Public / Privé) soit pris en compte dans le *matching* Jaro-Winkler initial, le tirage, qu'il soit aléatoire ou proportionnel à la distribution observée dans le recensement de 2017, écrase ensuite complètement cette caractéristique. Ce faisant, il introduit non seulement des discordances entre le déclaratif (profession et statut) et la codification – agriculteur indépendant (11) codé en salarié agricole (69) par exemple – mais aussi de l'hétérogénéité entre des situations professionnelles strictement identiques – des agriculteurs se déclarant indépendants (11) étant alternativement, selon le tirage, codés en indépendants (11) ou en salariés (69).

L'examen plus approfondi des codifications des professions de 500 répondants (Ego), respectivement codées par *matching* Jaro-Winkler et par SICORE, tend à confirmer la meilleure qualité de codification par SICORE. Sur 500 professions codées, 317 sont codées de la même façon selon les deux procédures, soit 63,4%. Mais lorsque des différences existent, la codification SICORE est généralement plus correcte. En définitive, 73,8% des professions de l'échantillon examinées sont correctement codées par SICORE contre 64,8% par *matching* Jaro-Winkler.

L'extrait des comparaisons présenté dans le tableau 7 rend compte du diagnostic de cette comparaison des codifications. Plus performante d'un point de vue qualitatif, la codification SICORE a été adoptée.

Les échecs de codification restants (libellés non reconnus, codifications douteuses...) en partie liés à des activités propres au contexte Polynésien représentent 22,5% des professions à coder. Elles font l'objet d'une codification manuelle *ad hoc*.

Tableau 7 Comparaison de la codification des professions de l'enquête *Feti'i e fenua* selon les deux méthodes (extrait)

DernierTravail	StatutTravailleur	Pub	matching Jaro- Winkler	Sicore	diagnostic
DIRECTEUR BANQUE	Salarié	Privé		37	ok Sicore
MERE AU FOYER	*	*	82		ok Jaro-Winkler
MANAGER	Salarié	Privé	46		possible Jaro-Winkler
FEMME DE MENAGE	Indépendant	Privé	68	56	ok Sicore
TECHNICIEN DE DIFFUSION	Salarié	Privé	46		ok Jaro-Winkler
CHEF COMMERCIAL	Salarié	Privé		46	ok Sicore
GERANT	Indépendant	Privé	21	22	ok Sicore
GERANT BUREAU D ETUDE	Indépendant	Privé	38		possible Jaro-Winkler
ENTREPRENEUR	Indépendant	Privé		22	ok Sicore
CONSEILERE PEDAGOGIQUE	Salarié	Public		42	ok Sicore
AUCUN	*	*	82		ok Jaro-Winkler
SURVEILLANT CENTRE PENNITENTIER	Salarié	Public		53	ok Sicore
AIDE SECRETAIRE	Salarié	Public	52		ok Jaro-Winkler
RESPONSABLE DE LIBRAIRIE	Salarié	Privé	38		trop élevé > 46
DOUANIER	Salarié	Public	45	52	pas de grade ok Sicore
OPTICIENNE	Indépendant	Privé		43	ok Sicore
CHEF D'EQUIPE	Salarié	Privé	12	62	ok Sicore
DIRECTRICE COMMERCIALE	Salarié	Privé	46	37	ok Sicore

Note : Rappel des grandes catégories de la codification des PCS 10 Agriculteurs exploitants ; 20 Artisans, commerçants et chef d'entreprise ; 30 Cadres ; 40 Professions intermédiaires ; 50 Employés ; 60 Ouvriers ; 82 inactifs.

IV. Apurement et imputation des dates

Cette partie du travail concerne l'apurement et l'imputation des âges/dates de naissance et décès de l'ensemble des membres de la famille. Pour rappel, la personne enquêtée peut alternativement répondre, selon la façon dont elle se remémore le plus facilement l'événement, que *la personne est née/décédée telle année ou à tel âge*. En outre, la saisie des dates de naissance et de décès était obligatoire pour chacune des personnes décrites. Il n'y a donc pas de valeur manquante dans les données brutes.

Les personnes interrogées l'étant sur la base d'un tirage par âge (âgées de 40 à 59 ans l'année de l'enquête), aucune imputation n'est à faire sur leur âge. Simplement, nous avons dû recoder la date de naissance d'un individu du module Conjoint à partir d'un individu similaire (NumFA=030271).

L'imputation concerne donc les dates de naissance et de décès des modules Enfants, Parents et Fratrie.

IV- 1. Qualification des données manquantes

Avant de choisir une méthode d'imputation, il convient d'examiner l'ampleur de la non-réponse concernant les dates demandées selon le lien de parenté à la personne enquêtée. Nous procédons donc initialement à cette évaluation.

Les répartitions des valeurs manquantes concernant les années de naissance et de décès par sous-module sont présentées dans le tableau 8 ci-dessous.

Puis, en vue de l'imputation de ces valeurs manquantes, nous avons procédé à une analyse des dates manquantes et de celles renseignées pour chacun des sous-modules des modules Parents, Enfants et Fratrie.

Tableau 8 Fréquence des valeurs manquantes des dates de naissance et décès des modules Enfants, Parents et Fratrie

	Années de naissance manquantes		Années de décès manquantes	
	Nombre	%	Nombre	%
Mère	293	5,70%	247	10,27%
Père	481	9,36%	299	9,80%
Beau-Père	1055	27,85%	409	18,92%
Belle-Mère	866	22,86%	294	17,30%
Frères/sœurs d'Ego	5182	17,71%	452	24,29%
Frères/sœurs du cjt	5733	29,47%	624	20,04%
Enfant des deux	65	0,57%	5	2,25%
Enfant du conjoint	52	2,69%	4	8,70%
Enfant d'Ego	68	5,22%	1	3,57%

Source : données Feti'i e fenua 2019-2020, après recodage des dates manquantes

Dates de naissance : Module Parents

Les valeurs manquantes constituent ici moins de 10% des observations pour les parents d'Ego et entre 20% et 30% des dates de naissance des parents des conjoints (tableau 8). En outre, les dates de naissance sont manquantes simultanément (tableau 1 en annexe A2) pour les deux parents dans un peu moins de la moitié des cas pour les parents d'Ego et dans une majorité de cas pour les parents de son conjoint.

Les histogrammes des figures 8a et 8b représentent les fréquences des dates de naissance déclarées des parents et beaux-parents (une fois les valeurs manquantes écartées). Ces graphiques ont permis de vérifier la régularité des distributions des dates de naissance des parents d'Ego et de son conjoint avant de procéder à l'imputation. De façon convaincante (comme attendu), les variables semblent être plus concentrées pour le module Ego qui est strictement âgé de 40 à 59 ans alors que l'amplitude d'âge de son Conjoint n'est pas limitée. On remarque également qu'aucune anomalie importante n'apparaît.

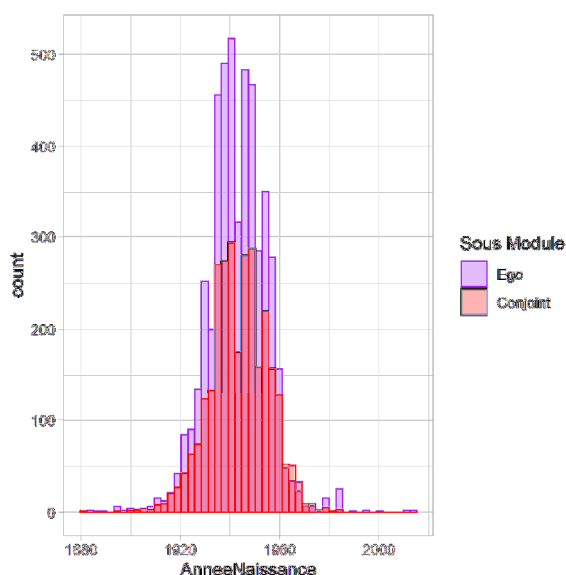


Figure 8a Distribution de l'année de naissance déclarée de la mère d'Ego (94%) et de son conjoint (77%)

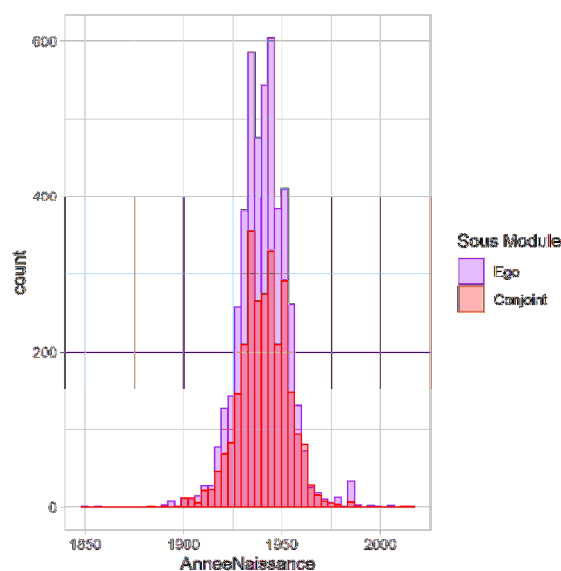


Figure 8b Distribution de l'année de naissance déclarée du père d'Ego (90%) et de son conjoint (72%)

Source : données *Feti'i e fenua* 2019-2020, après recodage des dates manquantes

Dates de naissance : Module Enfants

Une très grande majorité des enfants ont leurs dates de naissance renseignées, et ce quels que soient leurs parents (Ego, son conjoint ou bien les deux), comme le montre le tableau 8. En outre, rares sont les fratries ayant toutes les dates de naissance manquantes (cf. tableau 2 en Annexe A2).

La distribution obtenue est présentée dans la figure 1 de l'annexe A4 (le label « E-C » correspond aux enfants de la personne enquêtée (Ego) et de son conjoint « C » du conjoint seulement et « E » d'Ego seulement). Comme pour le module Parents elle permet de vérifier qu'aucune valeur inattendue ne présente de mode après l'apurement.

Enfin, il est intéressant de remarquer que plus de 88% des fratries du module enfant sont ordonnées.

Dates de naissance : Module Fratrie des enquêtés et de leur conjoint

Il y a, dans les fratries de ce module, bien plus de valeurs manquantes que dans celles du module Enfants. Il y en a même un peu plus que dans le module Parents. D'après le tableau 8, il y a davantage de dates de naissance manquantes en proportion dans le module Fratrie – un peu plus de 17% pour la fratrie d'Ego et un peu moins de 30% pour la fratrie de son conjoint.

La distribution des dates de naissance est représentée dans la figure 2 de l'annexe A4. Dans ce cas, on remarque une queue de distribution étendue à droite mais aucun mode ne permet de déterminer d'autres valeurs à recoder comme manquantes. Ces répartitions permettent donc ici aussi de vérifier la qualité des données après apurement.

Enfin dans ce module, ce ne sont que 57% des fratries qui sont décrites dans l'ordre.

Dates de décès : Module Parents

L'apurement, permet d'obtenir les distributions représentées dans les figures 9a et 9b. Ici encore, malgré des queues de distributions étendues vers la gauche, on ne remarque aucun mode qui suggérerait une valeur à recoder comme manquante. Par ailleurs, le nombre de décès par an augmente avec le temps jusqu'à l'année de l'enquête, ce qui est cohérent.

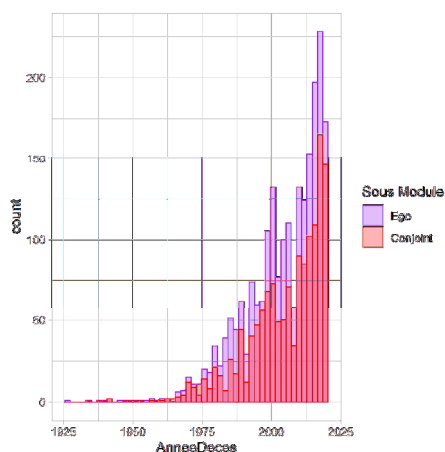


Figure 9a Distribution de l'année de décès déclarée de la mère d'Ego (90%) et de son conjoint (83%)

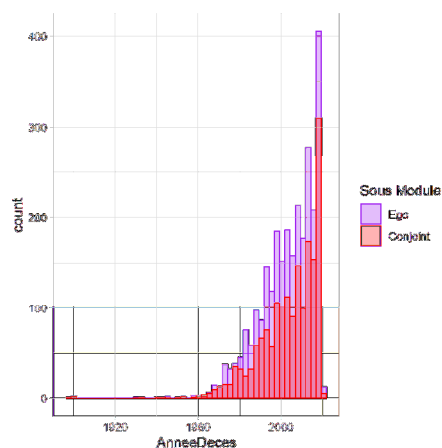


Figure 9b Distribution de l'année de décès déclarée du père d'Ego (90%) et de son conjoint (81%)

Source : données Feti'i e fenua 2019-2020, après recodage des dates manquantes

Par ailleurs, si les dates de décès sont manquantes dans des proportions comparables aux dates de naissance (tableau 8), on remarque qu'années de naissance et de décès sont souvent manquantes simultanément (tableau 4 en Annexe A2).

Dates de décès : Module Enfants

La distribution de ces dates de décès est représentée par l'histogramme de la figure 3 de l'annexe A4. Elle est irrégulière du fait du faible nombre d'individus décédés dans ce module étant donné leur âge. Pour les mêmes raisons, le nombre de dates de décès manquantes est très faible dans ce module (tableau 8).

Dates de décès : Module Fratrie

Après apurement, on aboutit à la distribution représentée dans la figure 10 pour les dates de décès de la fratrie d'Ego et de son conjoint (toujours en fonction du sous-module). Là encore le nombre de décès des frères et sœurs d'Ego et de son conjoint augmentent avec le temps pour atteindre leur maximum l'année de l'enquête. Notons que la distribution est plus irrégulière que pour le module Parents, ce qui était attendu, étant donné la distribution des dates de naissance et la part de valeurs manquantes dans ce module.

Ici aussi, les pourcentages de valeurs manquantes parmi les dates de décès sont similaires à ceux observés pour les années de naissance (tableau 8) et nombreux sont les individus du module Fratrie pour lesquels on n'a ni les dates de naissance ni les dates de décès (tableau 6 en Annexe A2).

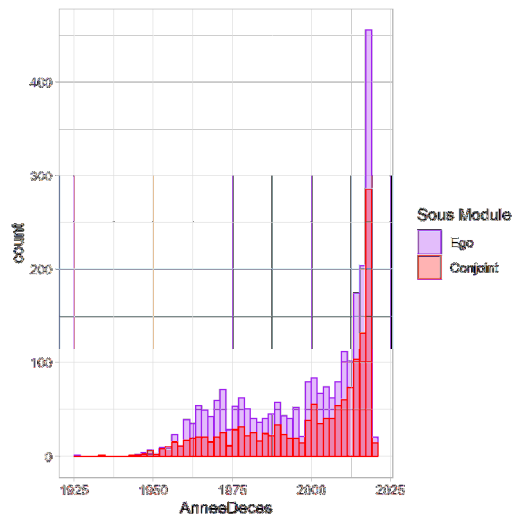


Figure 10 Distribution des années de décès déclarées des frères et sœurs d'Ego (76%) et de son conjoint (80%)

Source : données Feti'i e fenua 2019-2020, après recodage des dates manquantes

IV- 2. Imputation des dates manquantes

Éléments de méthodologie : *proximity mean matching* et imputation multiple

La deuxième partie du travail sur les dates de naissance et de décès consiste à imputer les valeurs manquantes. En principe, l'imputation revient à résoudre le problème des données manquantes, non pas en les remplaçant par la meilleure prédiction possible mais par une valeur plausible et exploitable (Little et Rubin 2002).

Il est possible de détailler cette démarche grâce aux postulats suivants (Little 1988) :

1. Fonder l'imputation sur une distribution prédite des données manquantes, étant donné les valeurs observées ;
2. Tenir compte du contexte : les déterminants des variables à imputer et leur probabilité d'être manquantes ;
3. Conserver les interactions entre les variables ;
4. Éviter d'extrapoler au-delà du domaine de distribution de la variable ;
5. Conserver une part d'aléatoire dans les prédictions (c'est-à-dire du bruit) ;
6. Prendre en compte l'imputation dans les estimations sur les données imputées (erreurs, variance, etc.)

La méthode que nous avons ici choisi d'appliquer pour l'imputation des dates de naissance et de décès des ascendants, des descendants et des collatéraux d'Ego et de son conjoint est appelée *proximity mean matching* adapté au cas multivarié. Nous l'associons à une imputation multiple dont le processus est détaillé dans l'encadré 1.

Nous allons maintenant présenter ces méthodes ainsi que leur intérêt pour l'imputation des dates de naissance et de décès des parents, des fratries et des enfants d'Ego et de son conjoint avant de décrire leur implémentation.

Encadré 1 : Le processus d'imputation

Le procédé qui a permis d'imputer des dates de naissance et de décès des modules Parents, Enfants et Fratrie peut être résumé comme suit :

- 1) Prédiction des variables à imputer à l'aide d'un modèle paramétrique – ici un modèle de régression logistique.
- 2) Appariement de chacune des valeurs manquantes avec ses k plus proches voisins parmi les valeurs renseignées, à partir des prédictions obtenues lors de l'étape précédente.
- 3) Tirage aléatoire parmi les valeurs renseignées pour ces k voisins. L'observation tirée sera la valeur imputée.
- 4) Itération de ce processus n fois (ici 5) : il s'agit de l'imputation multiple.
- 5) Calcul de la moyenne de ces n imputations. On sélectionne alors l'imputation la plus proche de cette moyenne de sorte à lisser les imputations (c'est-à-dire à réduire le bruit).

VI- 2.1 Le *proximity mean matching*

Le *proximity mean matching* peut être considéré comme une variante des méthodes d'imputation dites *hot deck* qui consistent à remplacer les données manquantes par des valeurs « proches », déjà présentes dans les données observées (Andridge et Little, 2010). Il s'agit d'apparier les données manquantes avec un échantillon de k valeurs observées grâce à une prédiction de cette valeur. On obtient ainsi un ensemble de valeurs voisines, parmi lesquelles la valeur imputée est ensuite tirée aléatoirement.

L'appariement des données manquantes avec un échantillon de valeurs observées repose sur une prédiction paramétrique de la variable à imputer à partir des données. Ici, nous avons choisi un modèle de régression logistique.

Du fait du tirage aléatoire d'une valeur parmi les k voisins ainsi définis, le *proximity mean matching* fait partie des méthodes *hot deck* dites aléatoires où l'on détermine un ensemble de voisins au sein duquel l'imputation sera tirée. À l'inverse les méthodes déterministes sélectionnent un seul « donneur » qui constituera l'imputation.

Cette méthode présente donc de nombreux avantages au vu des principes que nous avons cités plus haut :

- Étant donné que l'imputation est toujours une valeur observée dans le jeu de données d'origine, le *proximity mean matching* permet l'imputation des valeurs plausibles (postulat n°4) ;
- Par ailleurs, étant donné la manière dont est construite la métrique permettant l'appariement, l'imputation repose également sur prédiction de la variable à imputer (postulats n°1 et 2) ;
- En outre, du fait du tirage effectué pour choisir la valeur à imputer, il s'agit là d'une méthode *hot deck* dite aléatoire, par opposition aux méthodes déterministes comme les plus proches voisins (postulat n°5) ;

- Enfin, elle permet d'appliquer des méthodes dites de *chained equations*¹² propres à la correction de la non-réponse multivariée car permettant de prendre en compte les interactions entre les variables (postulat n°3).

2.2 Imputation Multiple

On associe l'imputation par *proximity mean matching* à une imputation multiple qui consiste simplement à effectuer plusieurs fois (itérer) l'imputation, de sorte à obtenir plusieurs versions de la base de données imputée – qui diffèrent du fait de la part d'aléatoire dans la méthode d'imputation.

Cette méthode a été développée par Donald B. Rubin dans les années 1970, pour rendre compte du processus d'imputation dans les estimations sur les données imputées, notamment dans les calculs de variance (cité par van Buuren, 2018). Elle permet normalement d'obtenir plusieurs jeux de données imputés sur lesquels sont effectuées les estimations, permettant ainsi de tenir compte de la part d'aléatoire introduite par l'imputation dans les résultats finaux (postulat n°6). Ici, le but étant d'obtenir un seul jeu de données final, en vue de la constitution d'un fichier production recherche (FPR), conserver ces multiples imputations n'est pas envisageable.

Néanmoins, on peut relever un autre intérêt du processus d'imputation multiple. Il permet en effet de lisser les résultats en réduisant le bruit introduit par la part d'aléatoire dans les imputations générées, à l'image des techniques de *bootstrap*¹³. Ainsi, on impute *in fine* la valeur, parmi les multiples imputations qui est la plus plausible, c'est-à-dire la plus proche de la moyenne des imputations produites.

IV- 3. Implémentation

La méthode semble bien fonctionner, et est facile à implémenter à l'aide du package *mice* sous R (van Buuren et Groothuis-Oudshoorn, 2011). On effectue ainsi les imputations des données manquantes à partir des trois plus proches voisins en générant 5 imputations parmi lesquelles on sélectionne la plus proche de la moyenne des différentes imputations.

Nous présentons maintenant les deux méthodes qui nous ont permis d'effectuer ces imputations. La première a l'avantage de conserver les relations d'âge au sein des familles mais ne permet pas d'imputer toutes les dates de naissance et de décès. La seconde, *a priori* un peu moins précise, permet donc l'imputation des valeurs restantes.

Méthode 1 – Imputation par famille

Afin de prendre en compte au mieux les corrélations entre les âges des frères et sœurs au sein des fratries, on cherche à imputer les années de naissance des frères et sœurs ainsi que de leurs parents simultanément. On effectuera ainsi les imputations des modules Parents et Fratrie d'une part et du module Enfants d'autre part. Pour cela on travaille à partir de bases de données mises en forme où

¹² Il s'agit de méthodes itératives permettant d'estimer les distributions conditionnelles des variables que l'on cherche à imputer.

¹³ Méthode statistique de ré-échantillonnage et de tirage aléatoire pour approcher par simulation la distribution inconnue d'un estimateur (Efron, 1979). Nommée en s'inspirant des péripéties de la vie du Baron de *Münchhausen* qui se sortit des sables mouvants en tirant sur ses cheveux (ou sur les lacets de ses bottes – *bootstrap*-).

chaque ligne correspond à une famille. L'imputation par fratrie nécessite par ailleurs de constituer plusieurs bases de données autant que de tailles de fratrie.

Nous décrirons ici le processus de mise en forme et d'imputation des données à partir de l'exemple de familles aux fratries de 4 personnes.

Imputation par famille : mise en forme de la base, exemple des fratries de 4 personnes

Cette première méthode consiste donc à imputer en même temps toutes les dates de naissance manquantes en utilisant les informations disponibles sur toute la famille. Pour ce faire, on constitue une base des ascendants et collatéraux :

- Ego ou son Conjoint :
 - Année de Naissance,
 - Indicatrice du fait d'être un enfant donné à *fa'a'amu*¹⁴ ;
- Les trois autres membres de la fratrie :
 - Années de naissance et de décès,
 - Indicatrice du fait d'être un enfant donné à *fa'a'amu*,
 - Indicatrice du fait d'être en vie actuellement ;
- Les parents :
 - Années de naissance et de décès,
 - Indicatrice du fait d'être en vie actuellement.

Dans le cas du module Enfants, on utilise quasiment les mêmes informations sur les fratries d'enfants et leurs parents – c'est-à-dire Ego et son conjoint ; ainsi que la variable qui indique de qui est chaque enfant (Ego, son conjoint ou les deux).

Une fois ces bases de données constituées, on peut imputer les données manquantes à l'aide du package *mice* de R. On commence par représenter les données manquantes (figure 11) : comme l'indiquent les résultats exploratoires présentés en première partie on est ici dans le cas de données manquantes multivariées. Il est donc nécessaire de prendre en compte les relations entre les variables qui peuvent être manquantes simultanément à l'aide des équations en chaîne qui effectuent les estimations à partir des distributions conditionnelles des variables prises en compte dans l'imputation.

¹⁴ De façon succincte, précisons qu'un enfant donné à *fa'a'amu* est un enfant confié à d'autres parents.

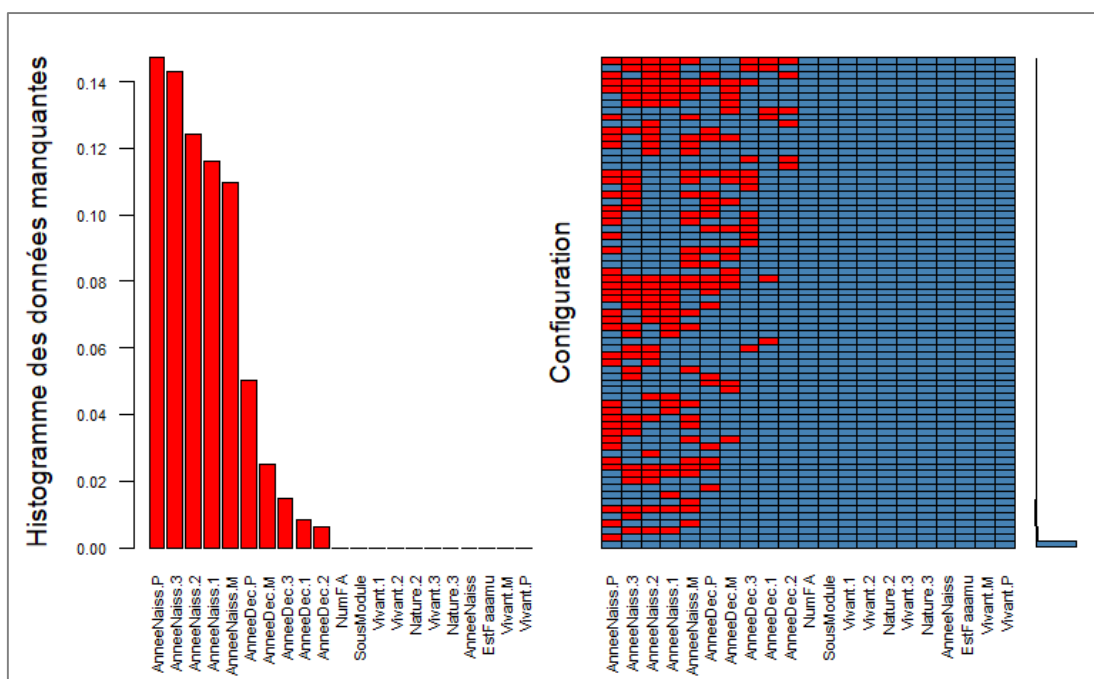


Figure 11 Représentation des données à imputer mises en forme, pour les fratries de 4 individus (Ego ou son conjoint et leurs 4 frères et sœurs) – Graphique effectué l'aide du package VIM sous le logiciel R
Source : données Feti'i e fenua 2019-2020, après recodage des dates manquantes

Lecture figure 11 :

Le diagramme en barres à gauche (figure 11) permet de quantifier les données manquantes pour chacune des variables sélectionnées. Par exemple la variable **AnnéeNaiss.M** (année de naissance de la mère d'Ego ou de son/sa conjoint) présente un peu plus de 14% de valeurs manquantes. On remarque que seules les années de naissance et de décès présentent des données manquantes.

Le graphique de droite permet quant à lui d'étudier chacune des configurations de données manquantes, c'est-à-dire, le fait pour chacune des variables décrites plus haut d'être manquantes simultanément. Les cases en rouge représentent les valeurs manquantes et celles en bleu, les valeurs renseignées. Ainsi, la dernière ligne du graphique correspond à la configuration où aucune des données n'est manquante. L'avant-dernière, la configuration où seule l'année de naissance du père n'est pas renseignée.

Enfin, l'étroit diagramme en barres de droite quantifie la présence de chacune de ces configurations dans les données. Dans l'exemple, le cas où aucune de ces variables n'est manquante semble être la configuration modale de notre jeu de données, ce qui est d'ailleurs cohérent avec les pourcentages déjà examinés de valeurs manquantes.

On applique ensuite un modèle de correction de la non-réponse multivariée par *proximity mean matching* et imputation multiple qui permet d'obtenir des imputations de bonne qualité.

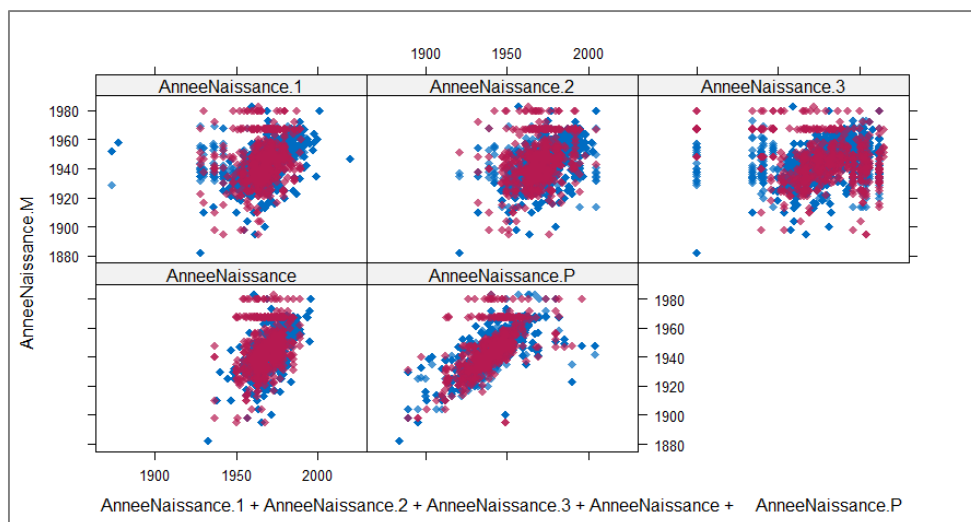


Figure 12 Année de naissance de la mère en fonction des années de naissance du père et de la fratrie, valeurs imputées (en rose) et valeurs observées (en bleu)

Source : données Feti'i e fenua 2019-2020, après imputation des dates manquantes

Lecture figure 12 :

Les cinq nuages de points de la figure 12 représentent la distribution des années de naissance des trois frères et sœurs (**AnneeNaissance.1**, **AnneeNaissance.2** et **AnneeNaissance.3**), celle d'Ego ou symétriquement de son conjoint (**AnneeNaissance**) ainsi que du père (**AnneeNaissance.P**) en fonction de l'année de Naissance de la mère (**AnneeNaissance.M**). Les points bleus correspondent aux cas où l'année de naissance de la mère était renseignée et ceux en rose à ceux où cette variable a été imputée. Ces graphiques permettent d'avoir un aperçu de la qualité des imputations en comparant la distribution des imputations à celle des valeurs renseignées. Ici, les deux nuages de points se superposent ce qui indique que les imputations sont vraisemblables : elles demeurent bien dans le domaine de distribution des variables observées et les relations entre les différentes variables du modèle (antériorité des dates de naissance selon le rang des enfants, etc.) sont conservées par la méthode choisie. Les figures 1 à 3 en annexe A4 font de même pour les autres dates imputées.

Cependant les bases de données ainsi constituées causent des problèmes de colinéarité pour des fratries trop nombreuses, ce qui empêche l'imputation. On ne peut donc pas imputer les dates de naissance et de décès des fratries de plus de 13 frères et sœurs pour le module Enfants ainsi que les fratries de plus de 10 frères et sœurs pour le module Fratrie (et donc de leurs parents également).

On aboutit à plus des trois quarts des observations manquantes imputées pour le module Fratrie, près de 70% pour le module Parents et plus de 95% pour le module Enfants (tableau 9).

Tableau 9 Bilan de la première imputation des dates de naissances pour chaque module de l'enquête Feti'i e fenua, 2020.

Module	Fratrie	Parents	Enfants
Obs. imputées	2 061	7 600	179
Obs. restantes	631	3 292	6
Total manquantes	2 692	10 892	185
Total observations	17 858	48 670	14 599

Méthode 2 – Imputation pour chaque individu

Il reste désormais à traiter les dates de la ligne « Obs. restantes » du tableau 9. Celles-ci vont être imputées à partir d'une sélection plus large de variables et en gardant le format initial des données, c'est-à-dire une ligne par individu. Cette méthode conserve dans une moindre mesure les relations d'âge au sein des familles mais a l'avantage de fonctionner sur les observations que nous n'avons pas pu imputer avec la première méthode. On ajoute donc aux prédicteurs de la première méthode la zone et l'archipel de naissance des individus ainsi que leur situation vis-à-vis du travail.

Une fois cette imputation effectuée, on construit une variable catégorielle indiquant, pour chacune des observations si elle a été imputée et par quelle méthode l'imputation a été effectuée le cas échéant. Ainsi, les modules Fratrie, Parents et Enfants contiennent deux variables **Naiss_Imp** et **Dec_Imp** qui portent respectivement sur les années de naissance et de décès et qui valent 0 lorsque la variable est observée, 1 si elle a été imputée avec la première méthode par famille et 2 si elle l'a été avec la deuxième méthode par individu.

Conclusion

Ce document de travail présente les opérations statistiques post-collecte relevant principalement de l'apurement et de l'imputation. Ces deux opérations essentielles de la production de données d'enquêtes quantitatives, relativement peu connues et documentées dans la littérature de sciences sociales, conditionnent la production du Fichier de Production et de Recherche (FPR), mis à disposition de la communauté scientifique pour l'analyse et la valorisation des résultats de l'enquête.

L'enquête *Feti'i e fenua* « Famille, territoire et relations intergénérationnelles en Polynésie française » (Ined-Ispf 2020) est la première enquête Famille collectée sur le territoire de la Polynésie française. Inspirée de l'enquête *Famille et logements* associée au recensement réalisée en France métropolitaine, elle a été collectée dans 31 îles du territoire auprès de 5 139 répondants avec un remarquable taux de réponse de 86%. Recueillant auprès de personnes âgées de 40 à 59 ans des informations géographiques, socio-économiques et relationnelles concernant les membres de la famille sur 3 générations, elle permettra de reconstituer, les déplacements de chacun, les relations entretenues entre les différentes personnes, la dispersion de la famille dans et au-delà du territoire. La description de l'espace résidentiel de la famille est en effet un outil essentiel à la compréhension des dynamiques familiales aussi bien en termes de mobilités de ses membres que de solidarités entre eux.

Nous avons choisi de présenter trois types d'opérations de post-collecte en les illustrant concrètement à partir de trois exemples : celui de la codification des lieux et des professions et enfin l'imputation de la non-réponse concernant les dates (naissance ou décès) des membres de la famille.

Dans le cas de la codification des localisations se pose la question de la précision des informations. Typiquement une information est connue « il vit en Bretagne » « il étudie à Tulsa » « elle est née dans le bateau », néanmoins elle n'a pas été saisie dans les termes de précision attendue. Il s'agit alors de reconnaître l'information, d'en extraire la meilleure précision, d'en proposer une codification qui puisse être analysée ensuite avec les informations plus précisément récoltées. C'est ainsi que de

nouvelles variables peuvent être créées à cet effet comme la variable REGION plus précise que le pays (ou la Métropole) et moins que le département.

Dans le cas de la codification des professions (près de 80 000), alors que la collecte reste limitée (dans ses détails) par rapport à l'ensemble des informations nécessaires à une codification en PCS raffinée, nous comparons la codification par SICORE (développée et utilisée à l'Insee) à celle que l'on peut obtenir par une méthode : Jaro-Wrinkler de comparaison de chaînes de caractères. Quelle que soit la méthode, un peu moins de 14 000 professions restent à coder à la main, mais SICORE reste très performant qualitativement.

Enfin, une grande partie de l'effort de préparation du Fichier de production et de recherche (FPR) a été consacré à l'imputation des dates manquantes. Il a fallu dans un premier temps analyser la nature de la non-réponse qui suit en effet des régularités : elle dépend par exemple de la proximité du lien. Ainsi un quart des dates de naissance des parents du conjoint sont manquantes versus 6% de celle de la mère de l'enquêté. D'autre part, l'imputation d'une date peut être « cadrée » par l'échantillon de l'ensemble des dates connues de la famille. Aussi, au lieu de se contenter d'une imputation *hot deck* courante, la méthode adoptée ici, dite d'imputation multiple, permet un ajustement avec une touche de variation aléatoire.

Au-delà de cette présentation des opérations d'apurement et d'imputation de l'enquête *Feti'i e fenua*, nous avons donc montré dans ce document les questions que pose l'homogénéisation des informations recueillies dans une collecte quantitative. Les comparaisons effectuées, les choix retenus illustrent ici la façon dont on adapte les méthodes disponibles à la qualité et la nature des données. Sont également rendues accessibles et documentées les opérations souvent invisibles qui conduisent des données brutes aux données largement accessibles à l'analyse.

Bibliographie

- Andridge, R., Little R. (2010). "A Review of Hot Deck Imputation for Survey Nonresponse", *International Statistical Review*, 78, 40-64.
- Buuren, S. van (2018) *Flexible Imputation of Missing Data, Second Edition (2nd ed.)*. CRC Press. Chapman & Hall/CRC Interdisciplinary Statistics.
- Buuren, S. van, Groothuis-Oudshoorn K. (2011). "Mice: Multivariate Imputation by Chained Equations in R". *Journal of Statistical Software* 45, n° 3.
- Caron, N. (2005). *La correction de la non-réponse par repondération et par imputation*, Document de travail de la Direction des Statistiques Démographiques et Sociales de l'Insee, n°M0502. http://www.insee.fr/fr/publications-et-services/docs_doc_travail/m0502.pdf
- Chauvet, G., Deville, J.C., Haziza, D. (2011). "On Balanced Random Imputation in Surveys", *Biometrika*, 98, 459-471.
- Deville J.-C., Särndal C.-E., Sautory O. (1993). "Generalized Raking Procedures in Survey Sampling", *Journal of the American Statistical Association*.
- Efron Bradley (1979). "Bootstrap Methods: Another Look at the Jackknife", *Annals of Statistics*, vol. 7, n° 1, p. 1-26.
- Ford B. L. (1983). "An Overview of Hot-Deck Procedures, Incomplete Data" in *Sample Surveys*, volume 2, Academic Press.
- Haziza, D. et Beaumont, J.-F. (2007). "On the Construction of Imputation Classes in Surveys", *International Statistical Review*, 75, 25-43.
- Jaro Matthew A. (1989). "Advances in Record Linking Methodology as Applied to the 1985 Census of Tampa Florida." In *Journal of the American Statistical Society* 84 (406): 414–20.
- KISH, L. (1965). *Survey Sampling*. New York: Wiley.
- KISH, L. and Frankel, M. R. (1974). "Inference from Complex Samples (with discussion)", *Journal of the Royal Statistical Society*. (B series), 36, 1-37.
- Little R.J.A. (1986). "Survey Nonresponse Adjustments for Estimates of Means", *International Statistical Review*.
- Little, R. J. A., et Rubin D. B. (2002) *Statistical Analysis with Missing Data Second Edition*. John Wiley & Sons, Inc.
- Winkler William E. (1990). "String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage." In: *Proceedings of the Section on Survey Research Methods, American Statistical Association*.
- Winkler William E. (1999). "The State of Record Linkage and Current Research Problems." In: *Statistics of Income Division*, Internal Revenue Service Publication R99/04.

A1 - Détermination de l'échantillon des répondants

La constitution de l'échantillon final rassemble les entretiens des répondants éligibles :

les femmes et hommes âgés de 40 à 59 ans

Exceptionnellement, si la personne désignée ne pouvait répondre à l'enquête (indisponible, absente le temps de l'enquête...), son **CONJOINT** – et exclusivement son conjoint – pouvait répondre à sa place, du fait de la « symétrie » familiale dans la collecte des informations, il a été toléré que le conjoint disponible de la personne tiré au sort par la méthode aléatoire Kish, souvent une femme, soit interviewée à sa place, cela s'est produit dans 824 cas. La variable **Present** du module EGO permet d'identifier ces questionnaires.

Il s'agit donc des individus pour lesquels- **STATUT** (dans la base FicheAdresse) = ERV ou RAF
Le nombre de répondants = **5140**

Néanmoins :

deux enquêtés pour lesquels le nombre d'individus éligibles dans le ménage (Somme de la variable Kishable) est égale à 0.

Il s'agit de NumFA=020346 et NumFA=012599. Ces deux personnes sont respectivement âgées de 37 ans (année de naissance 1982) et 38 ans (année de naissance 1981). Ils ont donc été retirés.

On retient donc les individus pour lesquels :

STATUT (dans la base FicheAdresse) = ERV ou RAF et Nombre d'individus "Kishable" > 0
(dans la base MENAGE)

Le nombre de répondants => **5138**

MAIS on a repêché un individu (**NumFA= 012826**) dont le STATUT = IAJ (impossible à joindre), car il a répondu intégralement et est dans la tranche d'âge requise (45 ans). Nous avons donc estimé que le statut était erroné.

L'échantillon rassemble donc 5139 répondants

A2 – Ages manquants selon le lien de parenté

1. Apurement

Les dates de naissance et de décès ne comportent au départ aucune donnée manquante puisque leur saisie était obligatoire dans le DataModel. La première étape du travail sur les dates a donc consisté à repérer les valeurs censées indiquer que la date est inconnue (par exemple : 1850) pour chacun des modules puis à les recoder en données manquantes.

Leur repérage a reposé sur la vraisemblance de ces valeurs, sur les remarques qui pouvaient les accompagner et sur l'observation des données (repérage de modes peu vraisemblables au vu de la distribution des dates de naissance et de décès).

Dates de naissance

Le repérage de ces dates nous a permis de sélectionner, pour chacun des modules, les valeurs à recoder comme manquantes :

- Module parents : nous avons choisi de supprimer certaines valeurs peu vraisemblables : 1850, 1860, 2019, 2020 ; ainsi que d'autres également peu vraisemblables et associées à des remarques indiquant que la date n'est pas connue ;
- Module enfants : nous avons retiré les dates suivantes 1850, 1860, 1898 ;
- Module fratrie : nous avons recodé en valeurs manquantes les dates de naissance correspondant aux années 1850 et 1860 ainsi que les valeurs peu vraisemblables quand elles étaient associées à des remarques indiquant que l'enquêté(e) ne connaissait pas la date de naissance de son frère ou de sa sœur.

Dates de décès

L'apurement des dates de décès des trois modules (parents, enfants, fratrie) a été effectué de la même manière que pour les dates de naissance.

- Module Parents : on a ici choisi de retirer, en plus des valeurs dont les remarques révélaient que l'enquêté ne connaissait pas la date, les dates de décès suivantes : 1850, 1860 et 9999
- Module Enfants : de même que pour les dates de naissance, nous avons retiré quelques dates de décès des enfants décédés à partir des remarques et des valeurs aberrantes (ici 1850, 1860, 1898 et 1900)
- Module Fratrie : les valeurs que nous avons considérées comme aberrantes dans ce cas étaient 1850, 1860 et 9999.

Une fois l'apurement effectué et la non réponse quantifiée (tableau 3 section I- 3), nous avons procédé à quelques statistiques descriptives afin d'observer les distributions.

Tableau 1 Dates de naissance manquantes simultanément ou non au sein des couples de parents, répartition par sous-module

	Date de naissance de la mère d'Ego				Date de naissance de la belle-mère d'Ego		
Date de Naissance du Père d'Ego		Renseigné	Manquant	Date de Naissance du beau-Père d'Ego		Renseigné	Manquant
	Renseigné	4576	84		Renseigné	2558	175
	Manquant	272	209		Manquant	364	691

Tableau 2 Valeurs manquantes simultanément au sein des fratries du module Enfants, répartition par sous-module

Enfants de	Toutes	Certaines	Aucune	Total
Les deux	7	36	3818	3861
Conjoint	11	19	836	866
Ego	18	12	586	616

Tableau 3 Valeurs manquantes simultanément au sein des fratries du module Fratrie, répartition par sous-module

Fratrie de	Toutes	Certaines	Aucune	Total
Conjoint	571	653	2252	3476
Ego	373	856	3714	4943

Tableau 4 Dates de naissance et de décès manquantes simultanément ou non, répartition par sous-module

Sous-Module		Décès Renseigné	Décès Manquant
Mère	Naiss. Renseignée	2211	0
	Naiss. Manquante	0	194
Père	Naiss. Renseignée	2645	181
	Naiss. Manquante	106	118
Belle-mère	Naiss. Renseignée	934	0
	Naiss. Manquante	0	412
Beau-Père	Naiss. Renseignée	1072	90
	Naiss. Manquante	246	254

Tableau 5 Valeurs manquantes pour les dates de décès du module Enfants, répartition par sous-module

Enfant de		Décès Renseigné	Décès. Manquant
Les deux	Naiss. Renseignée	26	0
	Naiss. Manquante	1	1
Conjoint	Naiss. Renseignée	38	3
	Naiss. Manquante	4	1
Ego	Naiss. Renseignée	211	1
	Naiss. Manquante	6	4

Tableau 6 Années de naissance et de décès manquantes simultanément ou non, module Fratrie, répartition par sous-module

Enfant de		Décès Renseigné	Décès Manquant
Conjoint	Naiss. Renseignée	1173	140
	Naiss. Manquante	236	312
Ego	Naiss. Renseignée	2157	203
	Naiss. Manquante	332	421

A3 : Années de naissance et de décès déclarées

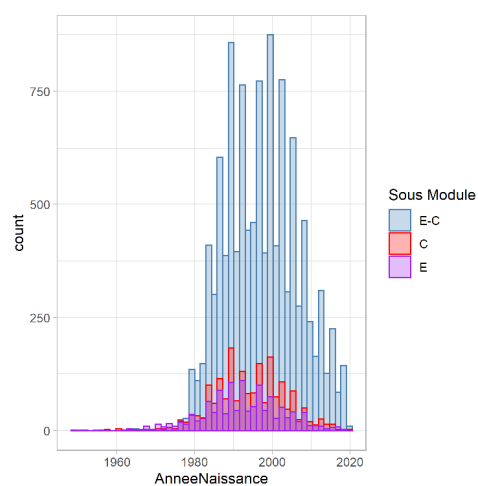


Figure 1 Distribution des années de naissance déclarées des enfants d'Ego (95%), de son conjoint (97%) et des deux (99%)

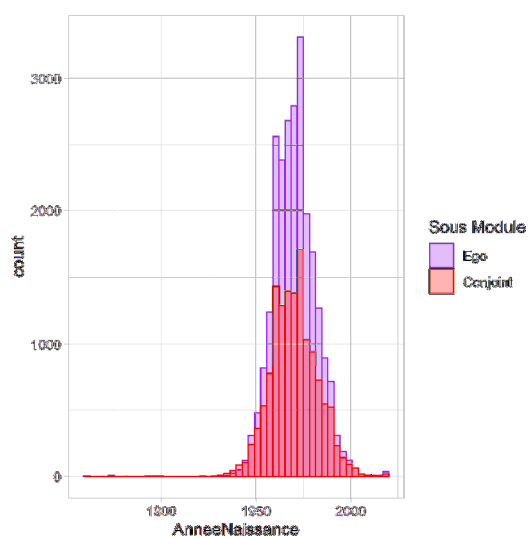


Figure 2 Distribution des années de naissance déclarées des frères et sœurs d'Ego (82%) et de son conjoint (70%)

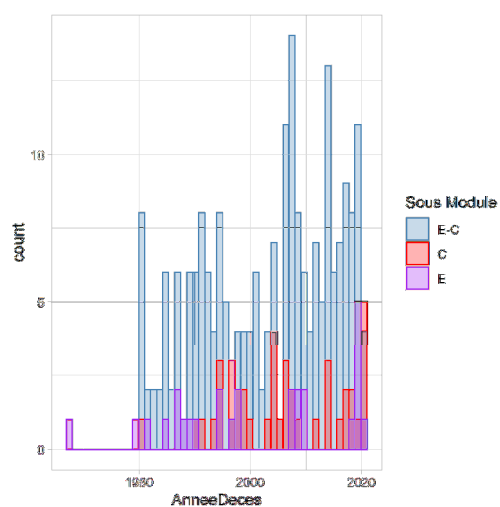


Figure 3 Distribution des années de décès déclarées des enfants d'Ego (96%), de son conjoint (91%) et des deux (98%)

A4 - Qualité des imputations des années de naissance du père et des fratries (fratries de 4 personnes d'Ego ou de son conjoint)

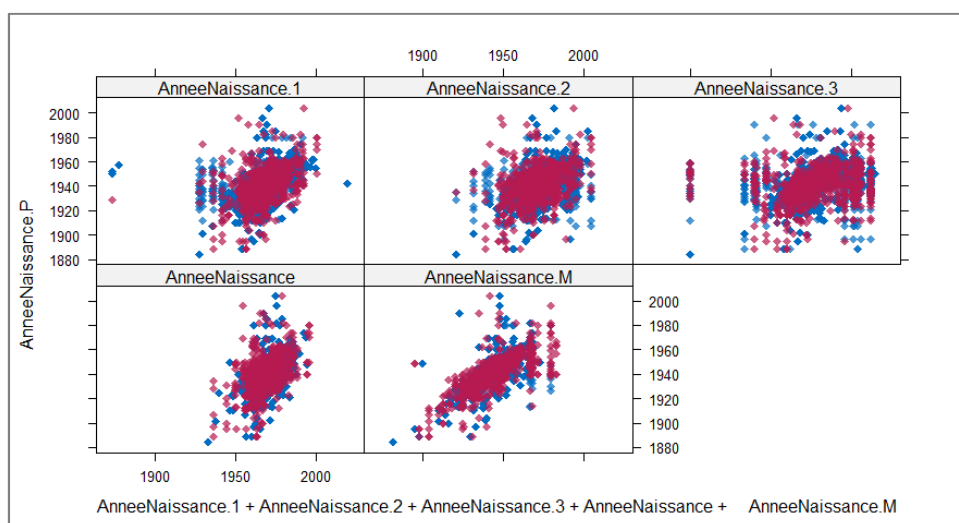


Figure 1 Année de naissance du père en fonction des années de naissance de la mère et de la fratrie, valeurs imputées (en rose) et valeurs observées (en bleu)

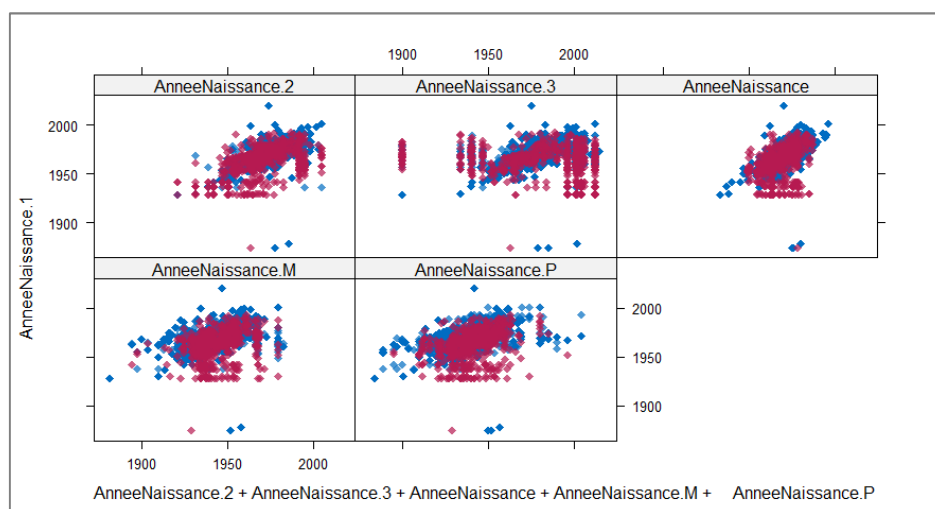


Figure 2 Année de naissance du premier membre de la fratrie en fonction des années de naissance des parents et du reste de la fratrie, valeurs imputées (en rose) et valeurs observées (en bleu)

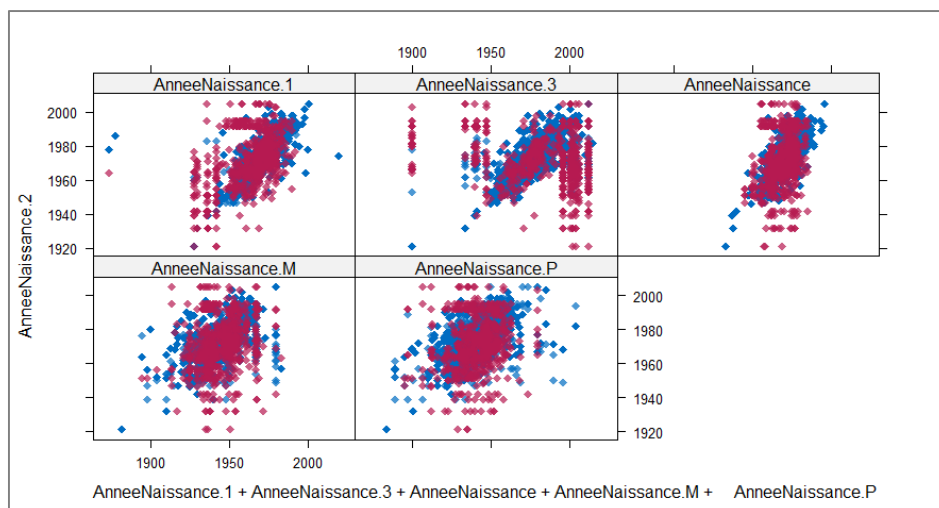


Figure 3 Année de naissance du deuxième membre de la fratrie en fonction des années de naissance des parents et du reste de la fratrie, valeurs imputées (en rose) et valeurs observées (en bleu)

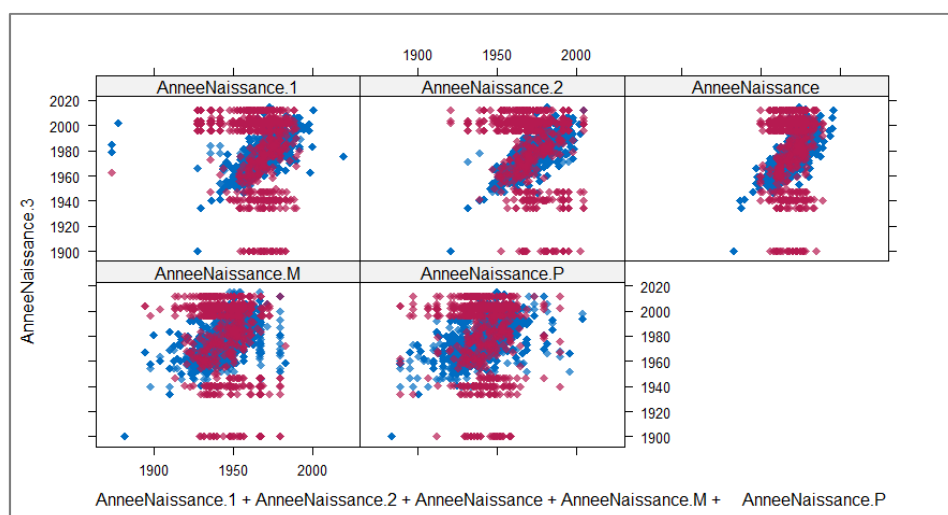


Figure 4 Année de naissance du troisième membre de la fratrie en fonction des années de naissance des parents et du reste de la fratrie, valeurs imputées (en rose) et valeurs observées (en bleu)

A5 – Variables de professions

Identification des variables de recueil (du DataModel) des professions pour les membres de la famille
(enquête *Feti'i e fenua*)

Module	Variables		
	<i>Travail</i>	<i>DernierTravail</i>	<i>StatutTravailleur</i>
Ego	Q14	Q15	Q16
Conjoint	Q14C	Q15C	Q16C
Parents	Q50/Q50C/Q60/Q60C	Q51/Q51C/Q61/Q61C	Q52/Q52C/Q62/Q62C
Fratie	Q73/Q73C	Q74/Q74C	Q75/Q75C
Enfants	Q91	Q92	Q93

A6 – Répertoire des abréviations

Bootstrap : méthode statistique de ré-échantillonnage et de tirage aléatoire pour approcher par simulation la distribution inconnue d'un estimateur. Nommée en s'inspirant des péripéties de la vie du Baron de *Münchhausen* qui se sorti des sables mouvants en tirant sur ses cheveux (ou sur les lacets de ses bottes –*bootstrap*-)

CAPI : *Computer-Assisted Personal Interview*

DataModel : Programme informatisé de saisie de l'enquête implanté sur les ordinateurs portables des enquêteurs

Feti'i e fenua : liens et territoire/patrimoine

FPR : Fichier de Production et de Recherche

Ined : Institut National d'Etudes Démographiques

ISPF : Institut de la Statistique de Polynésie française

MICE : Multiple Imputation by Chained Equations

SICORE : Système Informatique de COdage des Réponses aux Enquêtes

TCM : Tableau de Composition du Ménage

A7 – Questionnaire de l'enquête *Feti'i e fenua*

La famille de Bertrand

L'entretien fictif de Bertrand est un questionnaire « rempli » des informations d'une famille inventée.

Utilisé lors des formations des enquêtrices, il permet de reconstituer l'enchaînement du questionnement et de présenter le déroulement de l'entretien.

Etude de la famille et de son territoire en Polynésie

Feti'i et fenua

Intro

Nous préparons une étude sur la famille en Polynésie, les endroits où chacun vit et les occupations des différents membres de la famille...

>>>Module Personne éligible <<<

Quel est ton prénom ?

RICHIE

Tu te considères comme...

☒

un homme

☐

une femme

☐

Mahu

☐

Rae rae

En quelle année es-tu né(e) ou quel est ton âge ?

24 A

« Pour commencer et afin de savoir avec qui faire l'enquête, peux-tu me lister, s'il y en a, les personnes qui ont entre 40 et 59 ans parmi celles qui habitent dans ce logement »

habiter ici, c'est-à-dire que c'est leur logement habituel, même s'ils sont temporairement absents

Quel est son prénom ?

mon père BERTRAND

XXX se considère comme...

☒

un homme

☐

une femme

☐

Mahu

☐

Rae rae

En quelle année est né XXX ou quel est son âge ?

57 A

etc...

RECAPITULATIF

- RICHIE, homme 24 ans

- BERTRAND, homme, 57 ans

TIRAGE AU SORT

Listes des personnes éligibles

- BERTRAND, homme, 57 ans

ELOI est-il bien l'interlocuteur qui répondra aux questions de l'enquête pour lui-même et le reste de la famille ?

1. ☒ Oui, c'est bien BERTRAND qui va répondre lui-même

2. ☐ Non, c'est la conjointe de BERTRAND qui va répondre à sa place

>>>Module Ego<<<

4. Où es-tu né.e ?

☐ en Polynésie française → Dans quelle île, commune associée, archipel ?

Ile Commune associée → Q6

☐ dans une autre collectivité ou département d'outre-mer → Lequel ?.....

☒ en France métropolitaine → Dans quel département ? à Toulon.....

☐ dans une autre pays → Lequel ?.....

5. Tu es né en dehors de la PF, quand, en quelle année t'es-tu installé.e en Polynésie française la dernière fois ? 1980

6. Où résidait ta mère juste après ta naissance ?

1 ☐ en Polynésie française

2 ☐ dans une autre collectivité ou département d'outre-mer → Lequel ?.....

3 ☒ en France métropolitaine → Dans quel département ? ajouter la possibilité de taper le lieu en clair Toulon

4 ☐ dans une autre pays → Lequel ?.....

7. As-tu été Fa'a'amu ?

☐ oui ☒ non

8. Combien as-tu eu de frères/demi-frères ou adoptés (fa'a'amu),

y compris si certains sont décédés ? (Si NSP, Indiquez ceux connus de l'enquête ou à défaut)

01 F/ demi-F / F fa'a'amu

Combien as-tu eu de sœurs/demi-sœurs ou adoptées (fa'a'amu),

y compris si certains sont décédés ? (Si NSP, Indiquez ceux connus de l'enquête ou à défaut)

01 S/ demi-S / S fa'a'amu

Le chiffre donné par
l'enquête est-il
précis ?

9. As-tu eu des enfants ou adopté (fa'a'amu), y compris si certains sont décédés ? ☐ Non

☒ Oui → 10. Combien ? 01 fils nés 01 filles nées

 fils fa'a'amu filles fa'a'amu

11. Quel est ton niveau d'étude ?

☐ Tu n'es pas allé à l'école → Q13

☐ as appris à lire et compter/instruction ménagère/CJA

☐ Ecole primaire

☐ Collège (de la 6^e à la 3^e)

☐ BEP, CAP ou équivalent

☒ Lycée (2^{nde}, 1^{ère}, terminale)

☐ Etudes supérieures

12. Les localisations sont précisées en Commune Archipel / Département Pays pour l'étranger

plusieurs lieux successifs peuvent être saisis pour chaque niveau scolaire

Où as-tu fréquenté l'école primaire ?

Ile Commune associée Toulon Archipel

Ecole primaire ensuite ?

Où as-tu fréquenté le collège ?
Ile Commune associée <u>à Toulon</u> Archipel
le collège ensuite ? <u>au Harve</u>
Où as-tu fréquenté le lycée ?
Ile Commune associée <u>à Brest</u> Archipel
le lycée ensuite ?

13. As-tu séjourné en dehors de la Polynésie française pour une (des) période(s) de plus de 6 mois ?

☐ Oui ☒ Non → 14.

Décrivons ce séjour hors du territoire :

Localisation

En quelle année ou à quel âge a commencé le séjour

En quelle année ou à quel âge s'est terminé ce séjour

Quel était l'objectif du séjour (plusieurs réponses possibles) ?

☐ Etudier ☐ Travailler ☐ Service/Engagement militaire ☐ Autre

Avec qui es-tu parti (plusieurs réponses possibles si le statut a évolué durant le séjour) ?

☐ Seul ☐ Avec les parents ☐ En couple ☐ Avec les enfants

14. Du point de vue du travail...

☐ Tu as un travail actuellement, quotidien ou pas

☐ tu cherches un travail

☐ Tu ne travailles pas

☒ Tu es retraité(e)

☐ Tu n'as jamais travaillé → 17.

15. Quel est ton travail ou le dernier travail que tu as exercé ? Cuisinier dans l'armée

16. Il s'agit d'un travail rémunéré comme....

1. ☐ employé.e du privé

2. ☒ employé.e du public

5. ☐ autre situation (au foyer)

3. ☐ employeur

4. ☐ indépendant.e, à votre compte

17. Par ailleurs quel(s) type(s) d'activité exerces-tu habituellement ? (plusieurs réponses possible)

Par ailleurs, est-ce que tu fais du fa'apu pour vendre, tu vas à la pêche pour vendre ?	☒
Par ailleurs, est-ce que tu fais du fa'apu ou tu vas à la pêche pour ta propre consommation ?	☒
Par ailleurs, est-ce que tu fais et tu vends de l'artisanat ou tu fais de la danse, des spectacles ?	☒
Par ailleurs, est-ce que tu vends des services : réparation, repassage, mécanique, couture ?	☒
Par ailleurs, est-ce que tu as une activité familiale non-rémunérée (s'occuper d'une vieille personne, des enfants ?	☒
Par ailleurs, est-ce que tu as un travail bénévole dans une association sportive ou religieuse ?	☒

18. Touches-tu une pension au titre d'un handicap ?

☒ Oui

☐ Non

19. Es-tu actuellement en couple (si pas de conjoint dans le ménage) ?

1. ☒ Oui, avec une personne qui vit dans le même logement

2. ☐ Oui, avec une personne qui vit dans un autre logement

3. ☐ Non

20. Du point de vue légal, es-tu...

☒ marié.e (ou séparé.e non divorcé.e)

☐ pacsé.e

☐ veuf.ve

☐ divorcé.e

☐ célibataire
(jamais marié.e)

si Q19 = 3.Non

19a As-tu une relation amoureuse

☐ Oui ➡ Q21.

si Q19 = 2.Oui, avec une personne qui vit dans un autre logement

Si vous êtes en couple actuellement avec une personne qui vit la plupart du temps dans un autre logement

21. Où habite ton conjoint actuellement ?

22. Vous voyez votre conjoint... (de visu)

- ☐ Tous les jours ☐ toutes les semaines ☐ tous les 15 jours ☐ tous les mois
☐ 2 ou 3 fois par an ☐ 1 fois par an ☐ moins souvent

si Q19 = Non & Q19a = Non l'enquêté n'est pas en couple et n'a pas de relation amoureuse

➡ Module Logement

23. Vous êtes en couple,

Où vous êtes-vous installés à votre mariage / mise en couple ?

- ☐ chez toi ☐ chez tes parents ☐ Dans la famille de ton conjoint
☐ nous ne nous sommes pas installés ensemble.

✗ Autre : à la caserne

Localisation en Guadeloupe

>>>>Module conjoint<<<<

1.c Quel est le prénom de ton conjoint ? (surnom éventuellement) Josie

2.c Se considère t'elle/il comme ☐ un homme ✗ une femme ☐ un mahu ☐ un rae rae ☐ autrement

3.c En quelle année est né.e ton conjoint ? | 1 | 9 | 5 | 9 | Si NSP : Quel est son âge ? | | | |

4.c Où est-il/elle né.e ? en Guadeloupe

5.c En quelle année s'est-il/elle installé.e en Polynésie française la dernière fois ? | 1 | 9 | 8 | 0 |

➡ 6.c. Du point de vue légal, ton conjoint est-il...

- ✗ marié.e ☐ pacsé.e ☐ veuf.ve ☐ divorcé.e ☐ célibataire

7.c Etait-il Fa'a'amu ? ☐ oui ✗ non

8.c Combien ton conjoint a-t-il/elle eu de frères /demi-frères ou adoptés y compris si certains sont décédés ?

(Si NSP, Indiquez ceux connus de l'enquête)

| | 2 | F/ demi-F / F fa'a'amu

Combien ton conjoint a-t-il/elle eu de sœurs /demi-sœurs ou adoptées y compris si certains sont décédés ? (Si NSP, Indiquez ceux connus de l'enquête)

| | 2 | S /demi- S / S fa'a'amu

9.c Votre conjoint a-t-il eu des enfants ou adopté (fa'a'amu) au cours d'une précédente union, y compris si certains sont décédés ? ✗ Non

- ☐ Oui ➡ 10.c Combien ? | | | | fils nés / fils fa'a'amu
| | | | filles nés / filles fa'a'amu

11.c Quel est son niveau d'étude ?

- ☐ Il/elle n'est pas allé à l'école → **Q15.c**
☐ as appris à lire et compter/instruction ménagère/CJA
☐ Ecole primaire
 ☐ Collège (de la 6^e à la 3^e)
☒ BEP, CAP ou équivalent
 ☐ Lycée (2nde, 1^{ère}, terminale)
 ☐ Etudes supérieures

12.c Les localisations sont précisées en Commune Archipel / Département Pays pour l'étranger

plusieurs lieux successifs peuvent être saisis pour chaque niveau scolaire

Où a-t-il/elle fréquenté l'école primaire ? <i>à St Rose à la Guadeloupe</i>	Ile	Commune associée	Archipel
Où a-t-il/elle fréquenté le collège ? <i>à St Rose</i>	Ile	Commune associée	Archipel
Où a-t-il/elle fréquenté le lycée ? <i>en Guadeloupe</i>	Ile	Commune associée	Archipel
Où a-t-il/elle fait ses études supérieures ?	Ile	Commune associée	Archipel

13.c A-t-il/elle séjourné en dehors de la Polynésie française pour une (des) période(s) de plus de 6 mois ?

- ☐ Oui
 ☒ Non → **14.c**

Décrivons ce séjour hors du territoire :

Localisation

En quelle année ou à quel âge a commencé s'est terminé le séjour

Quel était l'objectif du séjour ? ☐ Etudier ☐ Travailler ☐ Service/Engagement militaire ☐ Autre

Avec qui est-il-elle partie ? ☐ Seul ☐ Avec les parents ☐ En couple ☐ Avec les enfants

14.c Du point de vue du travail...

- ☐ Votre conjoint a un travail actuellement, quotidien ou pas
 ☐ il-elle cherche un travail
☐ il-elle ne travaille pas
 ☒ il-elle est retraité(e)
 ☐ il-elle n'a jamais travaillé → **17.c**

15.c Quel est son travail ou le dernier travail qu'il/elle a exercé ?

serveuse restauration

16.c Il s'agit d'un travail rémunéré comme....

1. ☒ employé.e du privé
 2. ☐ employé.e du public
 5. ☐ autre situation (au foyer)
 3. ☐ employeur
 4. ☐ indépendant.e, à votre compte

17.c Et par ailleurs quel (s) type(s) d'activité exerce-t-il/elle habituellement ? (plusieurs réponses possible)

Par ailleurs, est-ce que _____ fait du fa'apu pour vendre, tu vas à la pêche pour vendre ?	<i>✓</i>
Par ailleurs, est-ce que _____ fait du fa'apu ou va à la pêche pour sa propre consommation ?	<i>✓</i>
Par ailleurs, est-ce que _____ fait et vend de l'artisanat ou fait de la danse, des spectacles ?	<i>✓</i>
Par ailleurs, est-ce que _____ vend des services : réparation, repassage, mécanique, couture ?	<i>Cuisine</i>
Par ailleurs, est-ce que _____ a une activité familiale non-rémunérée (s'occuper d'une vieille personne, des enfants ?)	<i>✓</i>
Par ailleurs, est-ce que _____ a un travail bénévole dans une association sportive ou religieuse ?	<i>✓</i>

18.c Touche-il/elle une pension au titre d'un handicap ?

- ☐ Oui
 ☒ Non

>>>>>Module logement<<<<<

Nous allons maintenant parler du logement/fare où tu habites actuellement...

26. Il s'agit...

- ☐ Fare ATR, MTR, FEI OPH...
- ☒ Autre type de **maison individuelle**
- ☐ Immeuble collectif de l'OPH
- ☐ Autre immeuble collectif
- ☐ Hôtel, pension de famille
- ☐ Habitation mobile : bateau, véhicule
- ☐ Construction provisoire

26a. Ce logement a été construit par...

- ☒ Une entreprise
- ☐ Toi-même
- ☐ Ton conjoint
- ☐ Un membre de ta famille ou de celle de ton conjoint
- ☐ Autre

27. Concernant ce logement où tu habites, es-tu ou ton conjoint...

- ☐ Locataire(s) d'un logement social (OPH, OTHS...) → Q29
- ☐ Locataire(s) d'un logement privé → Q29
- ☐ Propriétaire(s) (accédant) d'un fare (MTR, ATR, FEI, OPH...)
- ☒ Propriétaire(s) de votre logement sans indivision
- ☐ Propriétaire de votre logement en indivision **avec** une/plusieurs pers. de **ta** famille
- ☐ Propriétaire de votre logement en indivision **avec** une/plusieurs pers. de la famille **de ton conjoint**
- ☐ Logé gratuitement → **Q27c tu es logé gratuitement par une/plusieurs pers. de**
- ta** famille ☐ de la famille **de ton conjoint** ☐ la mairie ou le gouvernement ☐ autre → Q29

Q27b qui est le propriétaire...

- ☐ toi ☐ ton conjoint
- ☒ vous deux

28. A qui appartient le terrain où se trouve ce logement...

- ☐ il appartient à la commune ou à l'Eglise
- ☒ il t'appartient en bien propre (à toi ou ton conjoint)
- ☐ il appartient à **ta** famille (indivision)
- ☐ il appartient à la famille **de ton conjoint** (indivision)
- ☐ Autre préciser.....

29. Ce logement que vous habitez...

- ☒ toi ou un parent l'avez rénové et transformé, agrandi
- ☐ tu l'occupes sans transformations majeures

32. Votre logement actuel est-il équipé...

- | | | | |
|--|---|---|---------------------------------------|
| D'un toit anticyclonique | ☐ Oui | ☒ Non | ☐ toit encordé |
| De pilotis (Surélevé) | ☐ Oui | ☒ Non | |
| D'un garage, remise | ☒ Oui | ☐ Non | |
| De bâtiments agricoles, professionnels | ☐ Oui | ☒ Non | |
| D'une terrasse, balcon | ☒ Oui | ☐ Non | |

- Il y a...un séjour / salon ☒ Oui à l'intérieur ☐ Oui à l'extérieur *fare pote'e* ☒ Non
- des chambres ☒ Oui nb 4 ☐ Non
- une salle de bain ☒ Oui à l'intérieur ☐ Oui à l'extérieur ☒ Non
- une cuisine ☒ Oui à l'intérieur ☐ Oui à l'extérieur *fare tutu* ☒ Non

33. A proximité, y a-t-il d'autres fare où vivent d'autres membres de la famille ?

- ☐ Non ☒ Oui

peux-tu me préciser qui réside dans chaque fare: *ex. la famille de ma cousine Mary*

ma fille et son mari habitent à côté

34. En dehors de ton lieu d'habitation actuel, es-tu, toi ou ta famille, propriétaire, seul ou en indivision, d'un ou plusieurs terrain(s) ou logement(s) en Polynésie ou ailleurs ?

☐ Oui

☒ Non → [Filtre](#)

Pour chaque bien : Peux-tu me préciser...

Où se trouve ce bien ?

De quel type de bien s'agit-il ?

- ☐ d'un terrain cultivé ☐ d'une habitation occupée
☐ d'un terrain non-cultivé ☐ d'une habitation non-occupée

En es-tu propriétaire ?

- ☐ en bien propre ☐ en indivision

Comment as-tu accédé à la propriété ?

tu l'as acheté / il fait partie de ton héritage

De quel côté vient cet héritage

du côté de ton père / du côté de ta mère/des deux côtés

Filtre :

pas en couple



page suivante

34c. Et ton conjoint, ou sa famille, est-il propriétaire, seul ou en indivision, d'un ou plusieurs terrain(s) ou logement(s) en Polynésie ou ailleurs ? *Plusieurs réponses possibles*

☒ Oui

☐ Non → [Filtre](#)

Pour chaque bien : Peux-tu me préciser...

Où se trouve ce bien ? *en Guadeloupe*

De quel type de bien s'agit-il ?

- ☐ d'un terrain cultivé ☒ d'une habitation occupée
☒ d'un terrain non-cultivé ☐ d'une habitation non-occupée

En est-il propriétaire ?

- ☐ en bien propre ☒ en indivision

Comment a-t'il accédé à la propriété ?

elle l'a acheté / il fait partie de son héritage

De quel côté vient cet héritage

du côté de son père / du côté de sa mère/des deux côtés

Parlons de votre famille... >>> Module père d'EGO <<<

Nous allons maintenant parler de ton père, privilège celui qui t'a élevé.

44. Qui a joué pour toi le rôle de père ou de figure paternelle	☒ ton père (biologique) ☐ l'homme qui t'a élevé → Quel est votre lien (de parenté) ? ☐ tu n'as pas eu de figure paternelle.....	
45. (Pour se repérer) Quel est son prénom ?	MARCEL	
47. Est-il toujours en vie ?	Non <i>Il est décédé</i> → En quelle année ? 2005	
46. Où est-il né ? Polynésie française : île, commune ? <i>si NSP, archipel</i> Ailleurs en France : département ? Autre pays : lequel ? En quelle année est-il né ? son âge	en Bretagne à Quimper 1932 <i>si décédé → Q51</i>	
48b. Où habite-t-il actuellement ?	☐ Avec moi ☐ A côté ☐ Ailleurs	
48c. Localisation Polynésie française : île, commune ? <i>si NSP, archipel</i> Ailleurs en France : département ? Autre pays : lequel ?		
48d. et avec qui habite-t-il dans ce logement ?	☐ seul ☐ avec son conjoint avec un ou plusieurs ☐ de ses frères et sœurs ☐ de ses enfants ☐ de ses petits enfants ☐ avec d'autres personnes	
49. Etes-vous en contact (face à face, téléphone, internet...)...	☐ tous les jours ☐ toutes les semaines ☐ tous les 15 jours ☐ tous les mois ☐ 2 ou 3 fois par an ☐ 1 fois par an ☐ moins souvent ☐ jamais	
50. Du point de vue du travail...	☐ il a un travail rémunéré ☐ il ne travaille pas ☐ il est retraité	
52b. Par ailleurs quel(s) type(s) d'activités exerce-t-il actuellement ? <i>(plusieurs réponses possible)</i>	☐ il fait un faapu pour vendre, va à la pêche pour vendre ☐ il fait un faapu ou va à la pêche pour sa propre consommation ☐ il fait et vend de l'artisanat ou fait de la danse, spectacles ☐ il propose des services (réparation, mécanique) ☐ il exerce une activité familiale non-rémunérée (auprès pers. âgées...) ☐ il exerce un travail bénévole dans association sportive, religieuse	
51. Quel est ou était son (dernier) travail ?	Enseigne de vaisseau (armée)	
52. Est (ou était)-il...	☐ employé du privé ☒ employé du public ☐ employeur ☐ indépendant, à son compte ☐ autre situation (au foyer)	
Filtre : Le père est décédé → Page suivante (questions sur la mère)		
53. dans la vie courante, a-t-il besoin, a-t-il recours régulièrement à son entourage pour ... ? <i>(plusieurs réponses possibles)</i>	Dans la vie courante pour : - faire les courses - faire le ménage - faire ses repas - faire sa toilette - autre chose :	☐ a recours à qq de son entourage Qui ? ☐ est autonome, se débrouille seul ☐ aurait besoin mais n'a personne pour l'aider
54. Concernant sa santé, depuis le 1^{er} janvier 2019 a-t-il ...	☐ une pension Cotorep ☐ été hospitalisé (y compris une intervention sans séjour) ? combien de fois ☐ fait l'objet d'une évacuation sanitaire ? combien de fois ☐ bénéficié d'une aide à domicile ou de soins infirmiers	

décédé



Et votre mère...

>>> Module mère d'EGO <<<

54. Qui a joué pour toi le rôle de mère ou de figure maternelle	☒ votre mère (biologique) ☐ la femme qui vous a élevé → Quel est votre lien (de parenté) ?	
55. (Pour se repérer) Quel est son prénom ?	 MARTINE	
57. Est-elle toujours en vie ?	Oui Elle est décédée → En quelle année ? _ _ _	
56. Où est-elle née ? Polynésie française : ile, commune ? <i>si NSP, archipel</i> Ailleurs en France : département ? Autre pays : lequel ? En quelle année est-elle née ? son âge	 AUDIERNE en Finistère 1 9 4 0 si décédée → Q61	
58b. Où habite-t-elle actuellement ?	☐ Avec moi ☐ A côté ☒ ailleurs	
58c. Localisation Polynésie française : ile, commune ? <i>si NSP, archipel</i> Ailleurs en France : département ? Autre pays : lequel ?	 à BREST	
58d. Et avec qui habite-t-elle dans ce logement ?	☒ seule ☐ avec son conjoint ☐ avec un ou plusieurs ☐ de ses frères et sœurs ☐ de ses enfants ☐ de ses petits enfants ☐ avec d'autres personnes	
59. êtes-vous en contact (face à face, téléphone, internet...)...	☐ tous les jours ☐ toutes les semaines ☐ tous les 15 jours ☒ tous les mois ☐ 2 ou 3 fois par an ☐ 1 fois par an ☐ moins souvent ☐ jamais	
60. Du point de vue du travail...	☐ elle a un travail rémunéré ☐ ne travaille pas ☒ elle est retraitée	
62b. Par ailleurs quel(s) type(s) d'activités exerce-t-elle actuellement ? (plusieurs réponses possible)	☐ elle fait un faapu pour vendre, va à la pêche pour vendre ☐ elle fait un faapu ou va à la pêche pour sa propre consommation ☐ elle fait et vend de l'artisanat ou fait de la danse, spectacles ☐ elle propose des services (repassage, couture) ☐ elle exerce activité familiale non-rémunérée (auprès pers. âgées...) ☐ elle exerce un travail bénévole dans association sportive, religieuse rien	
61. Quel est ou était son (dernier) travail ?	 au foyer	
62. Est (ou était)-elle...	☐ employée du privé ☐ employée du public ☐ employeur ☐ indépendante, à son compte ☒ autre situation (au foyer)	
Filtre : La mère est décédée → Aller page suivante		
63. dans la vie courante, a-t-elle besoin, a-t-elle recours régulièrement à son entourage pour ... ? (plusieurs réponses possible)	Dans la vie courante pour : - faire les courses - faire le ménage - faire ses repas - faire sa toilette - autre chose : au maison de retraite ☒ a recours à qq de son entourage → Autre Qui ? ☐ est autonome, se débrouille seule ☐ aurait besoin mais n'a personne pour l'aider	
64. Concernant sa santé, depuis le 1^{er} janvier 2019 a-t-elle ...	☐ une pension Cotorep non ☐ été hospitalisée (y compris une intervention sans séjour) ? combien de fois _ _ non ☐ fait l'objet d'une évacuation sanitaire ? non combien de fois _ _ ☒ bénéficié d'une aide à domicile ou de soins infirmiers	

décédée



Filtre : si le répondant n'est pas en couple actuellement

→ aller page 12

Parlons maintenant des parents de votre conjoint.e

>>> Module père du conjoint <<<

44c. Qui a joué pour ton cjt le rôle de père ou de figure paternelle	☐ son père (biologique) ☐ l'homme qui l'a élevé → Quel est leur lien (de parenté) ? ☒ il n'a pas eu de figure paternelle.....	
45c. (Pour se repérer) Quel est son prénom ?		
47c. Est-il toujours en vie ?	Il est décédé → En quelle année ? _ _ _ _	
46c. Où est-il né ? Polynésie française : ile, commune ? <i>si NSP, archipel</i> Ailleurs en France : département ? Autre pays : lequel ? En quelle année est-il né ?	_ _ _ _ <i>si décédé → Q51c</i>	
48Cb. Où habite-t-il actuellement ?	☐ Avec moi ☐ A côté ☐ ailleurs	
48c. Localisation Polynésie française : ile, commune ? <i>si NSP, archipel</i> Ailleurs en France : département ? Autre pays : lequel ?	 ☐ seul ☐ avec son conjoint avec un ou plusieurs	
48Cd. et avec qui habite-t-il dans ce logement ?	☐ de ses frères et sœurs ☐ de ses enfants ☐ de ses petits enfants ☐ avec d'autres personnes	
49c. Etes-vous en contact (face à face, téléphone, internet...)...	☐ tous les jours ☐ toutes les semaines ☐ tous les 15 jours ☐ tous les mois ☐ 2 ou 3 fois par an ☐ 1 fois par an ☐ moins souvent ☐ jamais	
50c. Du point de vue du travail...	☐ il a un travail rémunéré ☐ il ne travaille pas ☐ il est retraité	
52cb. Par ailleurs quel(s) type(s) d'activités exerce-t-il actuellement ? <i>(plusieurs réponses possible)</i>	☐ il fait un faapu pour vendre, va à la pêche pour vendre ☐ il fait un faapu ou va à la pêche pour sa propre consommation ☐ il fait et vend de l'artisanat ou fait de la danse, spectacles ☐ il propose des services (réparation, mécanique) ☐ il exerce une activité familiale non-rémunérée (auprès pers. âgées...) ☐ il exerce un travail bénévole dans association sportive, religieuse	
51c. Quel est ou était son (dernier) travail ?		
52c. Est (ou était)-il...	☐ employé du privé ☐ employé du public ☐ employeur ☐ indépendant, à son compte ☐ autre situation (au foyer)	
Filtre : Le père du conjoint est décédé → Page suivante (belle-mère)		
53c. dans la vie courante, a-t-il besoin, a-t-il recours régulièrement à son entourage pour ... ? <i>(plusieurs réponses possible)</i>	Dans la vie courante pour : - faire les courses - faire le ménage - faire ses repas - faire sa toilette - autre chose :	☐ a recours à qq de son entourage Qui ? ☐ est autonome, se débrouille seule ☐ aurait besoin mais n'a personne pour l'aider
54c. Concernant sa santé, depuis le 1 ^{er} janvier 2019, a-t-il...	☐ une pension Cotorep ☐ été hospitalisé (y compris une intervention sans séjour) ? combien de fois _ _ ☐ fait l'objet d'une évacuation sanitaire ? combien de fois _ _ ☐ bénéficié d'une aide à domicile ou de soins infirmiers	

décédé



Parlons maintenant de la mère de votre conjoint.e

>>> Module mère du conjoint <<<

54c. Qui a joué pour ton cjt le rôle de mère ou de figure maternelle	☐ sa mère (biologique) ☒ la femme qui l'a élevé → Quel est votre lien (de parenté) ? <i>sa grand-mère maternelle</i>	
55c. (Pour se repérer) Quel est son prénom ?	<i>CHARLENE</i>	
57c. Est-elle toujours en vie ? <i>non</i>	<i>Elle est décédée</i> → En quelle année ? <i>1990</i> → Q61c	
56c. Où est-elle née ? Polynésie française : ile, commune ? <i>si NSP, archipel</i> Ailleurs en France : département ? Autre pays : lequel ? En quelle année est-elle née ?	<i>à la Martinique</i> <i>NSP</i> <i>si décédée → Q61c</i>	
58Cb. Où habite-t-elle actuellement ?	☐ Avec moi ☐ A côté ☐ ailleurs	
58Cc. Localisation Polynésie française : ile, commune ? <i>si NSP, archipel</i> Ailleurs en France : département ? Autre pays : lequel ?	☐ seule ☐ avec son conjoint avec un ou plusieurs	
58b. Et avec qui habite-t-elle dans ce logement ?	☐ de ses frères et sœurs ☐ de ses enfants ☐ de ses petits enfants ☐ avec d'autres personnes	
59c. Êtes-vous en contact (face à face, téléphone, internet...)...	☐ tous les jours ☐ toutes les semaines ☐ tous les 15 jours ☐ tous les mois ☐ 2 ou 3 fois par an ☐ 1 fois par an ☐ moins souvent ☐ jamais	
60c. Du point de vue du travail...	☐ elle a un travail rémunéré ☐ ne travaille pas ☐ elle est retraitée	
62Cb. Par ailleurs quel(s) type(s) d'activités exerce-t-il actuellement ? <i>(plusieurs réponses possible)</i>	☐ elle fait un faapu pour vendre, va à la pêche pour vendre ☐ elle fait un faapu ou va à la pêche pour sa propre consommation ☐ elle fait et vend de l'artisanat ou fait de la danse, spectacles ☐ elle propose des services (repassage, couture) ☐ elle exerce activité familiale non-rémunérée (auprès pers. âgées...) ☐ elle exerce un travail bénévole dans association sportive, religieuse	
61c. Quel est ou était son (dernier) travail ?	<i>Ouvrière dans la canne</i>	
62c. Est (ou était)-elle...	☒ employée du privé ☐ employée du public ☐ employeur ☐ indépendante, à son compte ☐ autre situation (au foyer)	
FILTRE : La mère du conjoint est décédée → Aller page suivante		
63c. dans la vie courante, a-t-elle besoin, a-t-elle recours régulièrement à son entourage pour ... ? <i>(plusieurs réponses possible)</i>	Dans la vie courante pour : - faire les courses - faire le ménage - faire ses repas - faire sa toilette - autre chose :	☐ a recours à qq de son entourage Qui ? ☐ est autonome, se débrouille seule ☐ aurait besoin mais n'a personne pour l'aider
64c. Concernant sa santé, depuis le 1 ^{er} janvier 2019 a-t-elle ...	☐ une pension Cotorep ☐ été hospitalisée (y compris une intervention sans séjour) ? combien de fois ☐ fait l'objet d'une évacuation sanitaire ? combien de fois ☐ bénéficié d'une aide à domicile ou de soins infirmiers	

décédée



Nous allons maintenant parler de tes frères et sœurs

>>> Module frères et sœurs <<<

Nous allons parler du 1^{er} frère / de la 1^{ère} sœur

☒ Tu es d'accord pour en parler

☐ Tu ne connais pas ce frère / cette sœur

☐ Tu ne souhaites pas parler de ce frère / cette sœur

(Pour se repérer) Quel est son prénom ?	HELENE	..RONAN..
66. Préciser le lien de parenté entre toi et XXX...	Biologique/demi-frère/sœur/fa'a'amu	frère biolo
67. XXX se considère-t-il comme...?	☐ M ☒ F ☐ Mahu ☐ Rae rae ☐ A	M
70. XXX est-il toujours en vie ? Si il/elle est décédé.e : En quelle année ?	1 9 6 4 <i>Di'edee</i>	Toujours en vie
68. En quelle année est-il/elle né.e ? (âge)	1 9 5 4	1 9 5 7
69. Où est-il/elle né.e ? Polynésie française : ile, commune ? Ailleurs en France : département ? Autre pays : lequel ?	Quimper si DCD → Q74	Quimper
71. Où habite-t-il/elle actuellement ?	☐ Avec moi ☐ A côté ☐ Ailleurs	à l'orient
71a. Localisation Polynésie française : ile, commune ? Ailleurs en France : département ? Autre pays : lequel ?	☐ seul(e) ☐ avec son conjoint avec un ou plusieurs	
71b. Et avec qui habite-t-il/elle dans ce logement ?	☐ de ses propres frères et sœurs ☐ de ses enfants ☐ un ou ses deux parents ☐ avec d'autres personnes	avec sa femme
72. A quelle fréquence es-tu en contact avec XXX (face à face, téléphone, internet...) ?	1. tous les jours 2. toutes les semaines 3. tous les 15 jours 4. tous les mois 5. deux ou 3 fois/an 6. une fois par an 7. moins souvent 8. jamais	Noter le n° de la réponse : une fois par an
73. Du point de vue du travail, actuellement, il/elle...	1. a un travail, quotidien ou pas 2. cherche un travail 3. ne travaille pas, au foyer 4. retraité(e)	retraité
74. Quel est ou était son (dernier) travail ?	sans objet <i>De la marine</i>	de la marine
75. Il s'agit d'un travail rémunéré comme...	1. employé.e du privé 2. employé.e du public 3. employeur 4. indépendant, à son compte 5. autre situation (au foyer)	2
75b. Bénéficie-t-il/elle d'une pension	COTOREP ☐ Oui ☐ Non	non
76. Combien a-t-il/elle eu d'enfants ? (tous y compris fa'a'amu)	 si 0 → F/S suivant	0 3 si 0 → F/S suivant
77. Certains de ces enfants vivent-ils en dehors de la Polynésie française ?	☐ Oui 0 0 ☐ Non → F/S suivant	☒ Oui 0 3 ☐ Non → F/S suivant
78. Où vivent-ils ? Ailleurs en France : Indiquer où Autre pays : Indiquer le ou les noms de pays	1..... 2..... 3..... 4.....	1. Nantes 2. NSP 3. l'orient 4.....

écédé.e



Nous allons maintenant parler >>> Module frères et sœurs du conjoint <<<

je ne les connais pas

Procéder colonne par colonne	Bô-Frère/sœur 1	Bô-Frère/sœur 2	
(Pour se repérer) Quel est son prénom ?	Frère 1.....	Frère 2.....	
66c. Préciser le lien de parenté entre ton conjoint(e) et XXX...	Biologique/demi-frère/sœur/fa'a'amu	Bio	Shirley Bio
67c. XXX se considère-t-il comme...	☒ M ☐ F ☐ Mahu ☐ Rae rae ☐ A	F	F
70c. XXX est-il toujours en vie ? Si il/elle est décédé.e : En quelle année ?	oui	oui	oui
68c. En quelle année est-il/elle né.e ? (âge)		1961	20 A Guadeloupe
69c. Où est-il/elle né.e ? Polynésie française : ile, commune ? Ailleurs en France : département ? Autre pays : lequel ?	Guadeloupe	Guadeloupe	Guadeloupe
71c. Où habite-t-il/elle actuellement ?	☐ Avec moi ☐ A côté ☒ Ailleurs	ailleurs	ailleurs
71Ca. Localisation Polynésie française : ile, commune ? Ailleurs en France : département ? Autre pays : lequel ?	Guadeloupe	en France	NSP
71Cb. Et avec qui habite-t-il/elle dans ce logement ?	☐ seul(e) ☐ avec son conjoint NSP avec un ou plusieurs ☐ de ses propres frères et sœurs ☐ de ses enfants ☐ un ou ses deux parents ☐ avec d'autres personnes		
72c. A quelle fréquence ton conjoint est en contact avec XXX (face à face, téléphone, internet...) ?	1. tous les jours 2. toutes les semaines 3. tous les 15 jours 4. tous les mois 5. deux ou 3 fois/an 6. une fois par an 7. moins souvent 8. jamais	Noter le n° de la réponse :	
73c. Du point de vue du travail, actuellement, il/elle...	1. a un travail, quotidien ou pas 2. cherche un travail 3. ne travaille pas, au foyer 4. retraité(e)		③
74c. Quel est ou était son (dernier) travail ?	Coureur	infirmière	infirmière
75c. Il s'agit d'un travail rémunéré comme...	1. employé.e du privé ① 2. employé.e du public 3. employeur 4. indépendant, à son compte ⑤ autre situation (au foyer)	②	⑤
75Cb. Bénéficie-t-il/elle d'une pension	COTOREP ☐ Oui ☐ Non nm	nm	nm
76c. Combien a-t-il/elle eu d'enfants ? (tous y compris fa'a'amu)	② si 0 → F/S suivant	NSP si 0 → F/S suivant	0
77c. Certains de ces enfants vivent-ils en dehors de la Polynésie française ?	☐ Oui ☐ Non → F/S suivant	☐ Oui ☐ Non → F/S suivant	
78c. Où vivent-ils ? Ailleurs en France : Indiquer où Autre pays : Indiquer le ou les noms de pays	1..... 2..... 3..... 4.....	1..... 2..... 3..... 4.....	

Nous allons maintenant parler de tes enfants

>>> Module Enfants d'Ego et de son conjoint <<<

Procéder colonne par colonne	Enfant 1	Enfant 2
(Pour se repérer) Quel est son prénom ?	... <u>AREUE</u> ...	... <u>RICHIE</u> ...
77. XX se considère-t-elle/il comme	☐ Garçon ☒ Fille ☐ Mahu ☐ Rae rae ☐ Autre	<u>garçon</u>
77b XX est-il, est-elle	☒ ton enfant biologique ☐ un enfant accueilli à faa'amu ☐ un de tes enfant confié (à faa'amu)	<u>bio.</u>
78. En quelle année est-il/elle né.e ?	<u>1990</u> ou Age <u> </u> <u> </u> <u> </u>	<u>24 A</u>
79. Où est-il/elle né.e ? Polynésie française : île, commune ? si NSP, archipel Ailleurs en France : département ? Autre pays : lequel ?	<u>TAHITI</u> <u>NSP</u>	<u>TAHITI</u> <u>NSP</u>
81. XXX est-il/elle toujours en vie ? Si il/elle est décédé.e : En quelle année ?	☐ DCD <u> </u> <u> </u> <u> </u> <u> </u> <u> </u> <u> </u> → Enfant suivant	<u> </u>
82. Où habite-t-il/elle actuellement ?	☐ Avec moi → Q86 ☒ A côté → Q83 ☐ ailleurs	<u>avec moi</u>
82. Localisation Polynésie française : île, commune. en France. Autre pays : lequel ?		
83. Avec qui XXX habite dans ce logement ?	☐ seul(e) ☒ avec son conjoint et avec un ou plusieurs ☐ de ses propres frères et sœurs ☒ de ses enfants ☐ un ou ses deux parents ☐ avec d'autres personnes	
84. Dans quel type de logement XXX habite t-il/elle ?	1. hébergé chez une personne de la famille 2. dans son propre logement (privé) 3. dans un logement à vous ou 'opu'he 4. en pensionnat 5. en cité universitaire / foyer 6. en caserne 7. autre (précisez.....)	<u> </u>
85. A quelle fréquence es-tu en contact avec XXX (face à face, téléphone, internet...) ?	1. tous les jours 2. toutes les semaines 3. tous les 15 jours 4. tous les mois 5. deux ou 3 fois/an 6. une fois par an 7. moins souvent 8. jamais	<u> </u>
86. Cet enfant dépend-t-il de ses parents financièrement ?	☐ Oui, régulièrement ☒ Oui, occasionnellement ☐ Non, jamais	<u>Oui régulièrement</u>
87. Cet enfant t' aide-t-il financièrement ?	☐ Oui, régulièrement ☐ Oui, occasionnellement ☒ Non, jamais	<u>Non</u>
88. Dirais-tu que tu es proche affectivement de XXX ?	1. Oui, plutôt ou très proche 2. Non, pas vraiment	Noter le n° de la réponse : <u>Oui</u> <u> </u>

Q77c Cet enfant est

- ☐ 1. de toi et ton conjoint actuel
☐ 2. de ton conjoint uniquement
☐ 3. de toi uniquement

Suite	Enfant 1	Enfant 2
89. Quel est son niveau d'étude ?	1. Il/elle n'a jamais été à l'école 2. a appris à lire et compter/instruction ménagère/CJA 3. Ecole primaire 4. Collège (6 ^{ème} – 3 ^{ème}) 5. Niveau BEP, CAP ou équivalent 6. Lycée (2 ^{nde} , 1 ^{ère} , terminale) 7. Enseignement supérieur	Noter le n° de la réponse : __ CAP cuisine
90. Où a-t-il fréquenté... l'école primaire le collège le lycée Où se sont déroulées ses études supérieures ?	P. rae Localisations des différents établissements faa'a faa'a	P. rae faa'a faa'a
91. Du point de vue du travail, actuellement... XXX est XXX a-t-il/elle toutefois déjà travaillé dans le passé ?	1. est étudiant ou écolier 2. a un travail, quotidien ou pas 3. cherche un travail 4. ne travaille pas 5. autre, précisez	Noter le n° de la réponse : __ cherche du travail
92. Quel est ou était son (dernier) travail.	Recepti n wiste	Cuisinier de cuisine
93. Il s'agit d'un travail rémunéré comme...	1. employé.e du privé 2. employé.e du public 3. employeur 4 indépendant, son compte 5. autre situation (au foyer)	Noter le n° de la réponse : 1
94. A-t-il/elle déjà séjourné en dehors de la Polynésie française au moins 6 mois ?	Oui Non	oui
95. Où s'est déroulé ce (ces) séjour(s) hors du territoire ?	Localisation(s)	Nouvelle Calédonie
96. XXX touche-t-il/elle une pension au titre d'un handicap ?	Oui Non	non

>>> Module Proches <<<

(Pour se repérer) Quel est son prénom ?	TETU A O R A	
98. Quel est le lien entre toi et XXX...	voisin	
99. XXX se considère-t-il comme...?	☒ M ☐ F ☐ Mahu ☐ Rae rae ☐ A	
100. En quelle année est-il/elle né.e ? (âge)	1 _ _ _ _ 62A	1 _ _ _ _
111. Où est-il/elle né.e ? Polynésie française : île, commune ? Ailleurs en France : département ? Autre pays : lequel ?	BORA BORA	
112. Où habite-t-il/elle ?	☐ Avec moi → Q115 ☒ A côté → Q114 ☐ Ailleurs	
113. Localisation Polynésie française : île, commune ? Ailleurs en France : département ? Autre pays : lequel ?		
114. A quelle fréquence es-tu en contact avec XXX (face à face, téléphone, internet...) ?	1. tous les jours 2. toutes les semaines 3. tous les 15 jours 4. tous les mois 5. deux ou 3 fois/an 6. une fois par an 7. moins souvent 8. jamais	Noter le n° de la réponse : 1 _
115. En quoi XXX est-il, est-elle proche ?	on faisait du sport maintenant on va pêcher ensemble	

Les habitants du logement

BERTRAND, 57 ans, née en 1963 à BREST, arrivé en 1980, français

JOSIE, épouse de BERTRAND, 61 ans, née en Guadeloupe, arrivée en 1980, française

RICHIE, leur fils, né en 1996 à Pirae, français

MIA, fille de RICHIE, 2 ans née à Faa'a, vit là en garde alternée, française



1 _ _ h 1 _ _

FIN : Nous avons terminé ! Merci beaucoup de votre participation. Souhaitez-vous donner votre avis sur ce questionnaire ?

.....

Leïla Fardeau, Éva Lelièvre et l'équipe ATOLLS (Archipels, Territoires et MObilités FamiLiaLes), *L'enquête Feti'i e fenua (Enquête Famille, territoire et relations intergénérationnelles en Polynésie française) : Apurement et imputation des données*, Paris, Ined, Document de travail, 262

