

Refonte des pondérations de l'enquête suivi de la demande touristique

Documents de travail

N° 2022-12 — Juillet 2022



Résumé

Le suivi de la demande touristique (SDT)¹ permet d'estimer, à partir de leurs réponses à l'enquête, le nombre de nuits passées par les résidents de France métropolitaine en dehors de leur domicile, que ce soit en France (métropole ou DOM) ou à l'étranger.

Ce document de travail explicite les méthodes de redressement utilisées pour améliorer les estimations issues de SDT, en proposant une stratification ex-post des données collectées, ainsi qu'une correction de la non-réponse. Cette refonte a conduit à rehausser les niveaux des estimations par rapport aux statistiques publiées antérieurement par la DGE, mais ne modifie pas les analyses en matière d'évolutions.

Les cinq premières parties décrivent les différentes méthodes retenues pour calculer les cinq systèmes de poids utilisés dans l'enquête. La sixième partie synthétise les différents travaux. Enfin la dernière partie révisé les résultats et mesure l'impact des redressements.

Summary

The tourist demand survey (SDT) makes it possible to estimate the number of nights spent by residents of metropolitan France in all accommodation in metropolitan France, whether commercial or non-commercial, in the overseas departments and abroad, based on their declaration. Insee took over responsibility for the SDT survey in 2020 from the Direction Générale des Entreprises (DGE). It is carried out by an external service provider.

This working paper explains the adjustment methods used to improve the estimates from SDT, proposing an ex-post stratification of the data collected, as well as a correction for non-response. This recasting has led to an increase in the levels of the estimates compared with the statistics previously published by the DGE, but does not modify the analyses in terms of changes.

The first five parts describe the different methods used to calculate the five weight systems used in the survey. The sixth part summarises the different works. Finally, the last part revises the results and measures the impact of adjustments.

1 L'Insee a pris la responsabilité de l'enquête Suivi de la Demande Touristique (SDT) en 2020, à la suite de la direction générale des entreprises (DGE). Elle est réalisée par un prestataire extérieur pour le compte de l'Insee.

Table des matières

Introduction.....	3
1. Calcul des pondérations mensuelles des individus.....	4
1.1. Post-stratification des échantillons.....	4
1.2. Correction de la non-réponse totale.....	5
1.3. Calage sur marges.....	5
2. Calcul des pondérations annuelles des individus.....	7
Bilan des étapes de calcul des poids individus mensuels et annuels :.....	8
3. Calcul des pondérations des voyages décrits.....	9
3.1. L'échantillonnage des voyages décrits.....	9
3.2. Calage sur le nombre de voyages et de nuitées déclarés.....	10
4. Calcul des pondérations des dépenses.....	12
4.1. Post-stratification des voyages en groupes de dépenses homogènes.....	12
4.2. Pseudo-poids initial et correction de la non-réponse.....	14
4.3. Calage sur le nombre de voyages et de nuitées par type d'hébergement.....	15
4.4. Le cas particulier de l'année 2020.....	16
5. Calcul des pondérations des séjours.....	17
6. Synthèse des étapes de calculs.....	18
7. Bilan.....	19

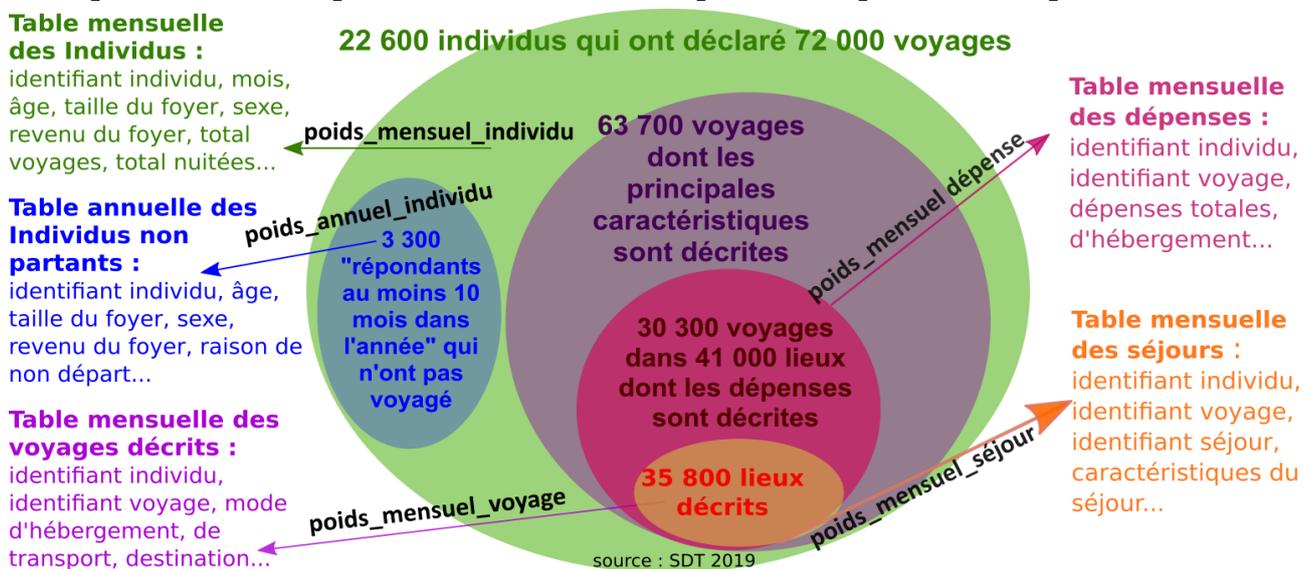
Introduction

L'enquête Suivi de la Demande Touristique (SDT) répond au volet « demande » du règlement européen de 2011 sur les statistiques du tourisme. Elle alimente aussi les comptes satellites du tourisme et la ligne tourisme de la balance des paiements établie par la Banque de France. Elle est la seule source permettant de disposer d'information sur les déplacements marchands et non marchands des résidents français. L'enquête SDT, sous la responsabilité de la Direction Générale des Entreprises jusqu'en 2019, est réalisée par un prestataire pour l'Insee depuis le 1^{er} janvier 2020. L'expertise du processus et des données passés a révélé l'existence de biais, et a conduit à un investissement méthodologique de l'Insee pour les réduire et améliorer la qualité.

L'échantillon SDT sous-représente les catégories de population les plus enclines à voyager, notamment les jeunes et les personnes aisées. Les méthodes de redressement mises en œuvre par le passé présentent plusieurs écueils, certains pouvant être corrigés *a posteriori* : l'absence de poids initial, liée au caractère non aléatoire du panel, peut en effet être partiellement corrigée par une stratification ex post des données collectées en fonction des variables les plus corrélées avec la thématique étudiée. Par ailleurs, le calage sur marge peut être amélioré par une correction préalable de la non-réponse. Cette note présente comment ces deux étapes essentielles de construction des pondérations ont été mises en œuvre, quand c'était possible, pour produire les différents systèmes de poids de l'enquête SDT.

L'enquête mensuelle SDT permet de produire 5 tables de résultats statistiques, chacune de ces tables correspondant à un échantillonnage différent et donc à une variable de pondération spécifique. La suite de la note s'attache à décrire, pour chaque système de poids, la méthode retenue pour calculer la nouvelle variable de pondération.

Principales caractéristiques des 5 tables annuelles produites à partir de l'enquête SDT :



1. Calcul des pondérations mensuelles des individus

1.1. Post-stratification des échantillons

Les travaux menés sur la base des données SDT 2019 ont permis de déterminer les caractéristiques individuelles les plus corrélées à la thématique des voyages touristiques, parmi les variables présentes dans l'enquête, et pour lesquelles on dispose, par ailleurs, d'une estimation fiable. Le revenu, l'âge du répondant, la taille du foyer et dans une moindre mesure le sexe, sont les caractéristiques les plus corrélées à la probabilité de réaliser un voyage touristique pour un mois donné. En particulier, l'âge du répondant a fait l'objet d'un travail spécifique pour explorer finement le lien entre l'âge et la probabilité de voyager. Deux tranches d'âges ont été retenues comme étant les plus enclines à réaliser des voyages professionnels, i.e les 35-44 ans et les 45-54 ans. Les travaux ont conduit aussi à revoir le découpage des âges élevés. La probabilité de voyager chute en effet significativement à partir de 70 ans, d'où l'intérêt de distinguer cette dernière classe d'âge des autres. Compte-tenu de la difficulté d'accorder convenablement les concepts de revenu (voir 1.3 infra), cette dernière variable n'a pu être retenue à cette étape, mais intervient néanmoins au moment du calage sur marge. Ainsi, l'échantillon des panélistes interrogés a été stratifié en croisant chaque mois, l'âge du répondant (6 classes), la taille du foyer (3 classes) et le sexe du répondant, comme suit :

Tranches d'âge du répondant		Taille du foyer		Sexe du répondant	
15-24 ans					
25-34 ans	X	1 personne	X	Homme	
35-44 ans		2 personnes		Femme	
45-54 ans		3 personnes et plus			
55-69 ans					
70 ans et plus					

Lorsque leurs effectifs étaient trop faibles ou parce qu'elles avaient peu de sens, certaines strates ont été regroupées avec d'autres. Les foyers des personnes âgées de 70 ans et plus sont rarement composés de plus de 2 personnes : dans ce cas, la taille du foyer a été ramenée à deux modalités (1 personne ou 2 personnes et plus). Les jeunes âgés de 15 à 24 ans sont peu nombreux dans l'échantillon, pour eux la taille du foyer n'a pu être prise en compte. La grille utilisée pour la stratification des pondérations des individus comprend finalement 30 strates.

Ainsi, chaque mois, pour chaque individu k appartenant à la strate h , on obtient les pseudos-poids initiaux suivants :

$$\mathbf{Poids_initial}_k = N_h / n_h$$

avec

$Poids_initial_k$: pseudo-poids initial de l'individu k appartenant à la strate h

N_h : taille de la strate h estimée par le recensement de la population sur le champ de la population métropolitaine, âgées de 15 ans et plus et vivant en logement ordinaire.

n_h : taille de l'échantillon initial de la méthode précédente pour la strate h

1.2. Correction de la non-réponse totale

Testée sur tous les mois de l'année 2019, la non-réponse à l'enquête SDT dépend significativement de l'âge, de la taille du foyer et du panel d'origine de l'enquêté (la collecte s'appuie en effet sur la mobilisation conjointe de plusieurs panels), cette dernière variable permettant de distinguer également le mode d'interrogation (internet, papier). Tenant compte de la post-stratification établie précédemment, la non-réponse à l'enquête a donc été redressée au niveau de l'âge, de la taille du foyer, du sexe et du panel d'origine.

Ainsi, le poids corrigé de la non-réponse de l'individu k appartenant à la strate h et au panel j :

$$\mathbf{Poids_CNR}_k = \mathbf{Poids_initial}_k / \mathbf{P}_{h,j}$$

avec

$\mathbf{Poids_CNR}_k$: le pseudo-poids initial individu corrigé de la non-réponse

$\mathbf{Poids_initial}_k$: le pseudo-poids initial de l'individu k

$\mathbf{P}_{h,j}$: probabilité de réponse des individus appartenant à la strate h et au panel j

1.3. Calage sur marges

Une dernière étape de calage sur marges a permis de tenir compte du revenu du foyer dans le calcul des nouveaux poids².

Tout comme la variable d'âge, la variable de revenu du foyer a fait l'objet d'un travail spécifique de regroupement en classes. L'analyse des corrélations entre l'appartenance à une des 13 tranches de revenus où les enquêtés étaient appelés à se placer et la probabilité de réaliser différents types de voyages (professionnel, à l'étranger, etc.) a conduit à retenir 6 classes de revenus. Cela a permis notamment d'isoler les classes de revenus les plus élevées, qui sont également celles les plus enclines à voyager et les plus sous-représentées dans l'échantillon. La variable ainsi obtenue est très fortement corrélée aux variables d'intérêt de l'enquête, mais ne trouve pas d'équivalent naturel dans les sources dont dispose l'Insee. En particulier, il s'agit d'un revenu déclaré par l'enquêté, alors que les statistiques sur les revenus habituellement produites par l'Insee s'appuient sur l'exploitation de sources administratives, fiscales et sociales. La somme des revenus perçus et des prestations sociales tirée de l'enquête revenus fiscaux et sociaux (ERFS) a néanmoins été utilisée pour approcher au mieux « la somme des ressources mensuelles nettes du foyer avant prélèvement à la source » qui est collectée dans SDT. L'écart entre les deux estimations étant impossible à estimer, le choix a été fait de procéder pour cette variable de revenu à un calage pénalisé, plutôt que de manière exacte, afin d'approcher sans trop de contrainte la structure des revenus tirée de l'ERFS.

L'étape de calage sur marge fait aussi intervenir, de manière exacte cette fois, la répartition de la population par région de résidence. Enfin, la répartition par âge et sexe a été aussi introduite dans le calage sur marge, pour maintenir le découpage en strates retenues lors de la 1^{re} étape du calcul des poids initiaux. Enfin, la marge totale utilisée pour le calage correspond à la population résidente de France métropolitaine, en logement ordinaire, tirée du dernier recensement de la population disponible au moment des calculs, ici 2017. Le choix de se limiter aux logements ordinaires a pour objectif de se rapprocher au mieux du champ de l'enquête, qui, a priori, exclut les populations vivant en communauté.

² À noter que les individus qui n'ont pas déclaré de revenu sont considérés comme des non-répondants. On en comptabilise 207 en 2019.

Cette dernière étape permet de déterminer un poids mensuel pour chaque individu de l'échantillon, noté par la suite *poids_mensuel_individu*.

Percentiles de la variable *poids_mensuel_individu* en milliers, calculés selon les mois de l'année 2019 :

mois	0%	1%	10%	25%	50%	75%	90%	99%	100%
janvier	0	0,9	1,3	1,9	2,6	3,8	4,9	8,0	16,7
février	0	0,9	1,3	1,9	2,7	3,8	5,0	8,2	19,4
mars	0	0,9	1,4	2,0	2,7	3,9	5,2	8,4	18,9
avril	0	0,9	1,3	1,9	2,7	3,9	5,1	8,5	19,6
mai	0	1,0	1,4	1,9	2,7	3,9	5,1	8,7	20,0
juin	0	1,0	1,4	1,9	2,8	4,0	5,5	9,6	18,4
juillet	0	0,9	1,4	2,0	2,8	4,1	5,5	9,7	20,6
août	0	1,0	1,4	1,9	2,8	4,0	5,5	10,4	23,0
septembre	0	1,0	1,4	2,0	2,7	3,9	5,3	9,8	23,7
octobre	0	0,9	1,4	1,9	2,6	3,9	5,3	10,0	22,6
novembre	0	1,0	1,4	1,9	2,7	4,0	5,5	10,5	21,9
décembre	0	1,0	1,4	1,9	2,7	4,0	5,6	11,1	21,8

2. Calcul des pondérations annuelles des individus

Un questionnaire annuel est adressé en janvier N+1 aux individus du panel qui n'ont déclaré aucun voyage pour motif personnel au cours de l'année N. Ce module rétrospectif sur l'année écoulée interroge les panélistes sur les raisons de non-départ. Pour s'assurer qu'il s'agit bien de personnes n'ayant pas voyagé au cours de l'année, il est adressé uniquement aux « bons répondants », c'est-à-dire à ceux qui ont systématiquement répondu n'avoir fait aucun voyage pour motif personnel et ce à au moins 10 enquêtes mensuelles de l'année N.

Le calcul des pondérations est donc réalisé pour un sous-échantillon constitué uniquement des « bons répondants ». Le champ de calcul des poids annuels est donc réduit à tous les panélistes ayant répondu au moins à 10 enquêtes mensuelles au cours de l'année, qu'ils aient voyagé ou non. Tous les autres ont un poids annuel égal à zéro et ne sont pas interrogés.

Les caractéristiques individuelles les mieux corrélées à la probabilité de voyager déterminées lors des modélisations mensuelles, restent significatives pour l'analyse annuelle. Les mêmes strates sont donc utilisées pour calculer les poids initiaux du module annuel, c'est-à-dire en croisant l'âge, la taille du foyer et le sexe. Les caractéristiques des individus, notamment l'âge évoluant au cours de l'année, ce sont les dernières caractéristiques connues qui sont retenues pour associer une strate à un individu. Les données livrées à l'Insee ne font pas apparaître de non-réponse dans le module annuel. Les raisons du non-départ interrogé lors de ce module sont imputées quand l'individu ne répond pas, ce qui revient à faire l'hypothèse que les non-répondants ont les mêmes comportements (raisons de non-départ) que les répondants. Il n'est donc pas possible de corriger la non-réponse pour établir ce poids annuel.

Par ailleurs, en n'interrogeant que les « répondants au moins 10 mois dans l'année » qui n'ont pas voyagé, l'hypothèse qui est faite indirectement est que ces « bons répondants » ont la même probabilité de voyager que le reste de l'échantillon qui n'est pas interrogé. Il est néanmoins possible que la non-réponse soit plus fréquente chez les panélistes qui ne voyagent jamais. Ce dernier point ne pouvant être mesuré, les étapes de calcul des pondérations décrites par la suite ne permettent pas de corriger cet éventuel biais d'échantillonnage.

Les pseudo-poids initiaux du module annuel s'écrivent pour chaque individu k appartenant à la strate h :

Si l'individu k a répondu à au moins 10 enquêtes mensuelles : $Poids_initial_k = N_h / R_h$,

Sinon : $Poids_initial_k = 0$

avec

$Poids_initial_k$: pseudo-poids annuel initial de l'individu k appartenant à la strate h

N_h : taille de la strate h estimée par le recensement de la population sur le champ de la population métropolitaine, âgées de 15 ans et plus et vivant en logement ordinaire.

R_h : Nombre d'individus ayant répondu à au moins 10 enquêtes mensuelles et appartenant à la strate h pour la dernière enquête mensuelle renseignée

Comme pour les poids individus mensuels, les pseudo-poids initiaux font l'objet d'un calage sur marge exact pour la structure de la population par tranche d'âge, par sexe et par région de résidence et d'un calage pénalisé sur le revenu.

Distribution des poids annuels 2019 des individus non partant en milliers :

année	0%	1%	10%	25%	50%	75%	90%	99%	100%
2019	0	1,1	1,5	1,8	2,8	3,8	5,7	12,0	25,9

Bilan des étapes de calcul des poids individus mensuels et annuels :

Type de pondération	Strates retenues pour le calcul des poids initiaux	Redressement de la non-réponse	Marges	Sources
Mensuelle pour les Individus	Mois Tranche d'âge Taille du foyer Sexe	Tranche d'âge Taille du foyer Sexe Panel d'origine	<u>Calage exact :</u> Région de résidence Tranche d'âge Sexe <u>Calage pénalisé :</u> Revenus	RP* 2017 en logements ordinaires ERFS** 2018
Annuelle pour les Individus « bons répondants »	Tranche d'âge Taille du foyer Sexe	Pas d'information sur la non-réponse	<u>Calage exact :</u> Région de résidence Tranche d'âge Sexe <u>Calage pénalisé :</u> Revenus	RP* 2017 en logements ordinaires ERFS** 2018

*RP : Recensement de la Population

**ERFS : Enquête Revenus Fiscaux et Sociaux

3. Calcul des pondérations des voyages décrits

À l'occasion de chaque interrogation mensuelle, les panélistes SDT déclarent un nombre total de voyages et de nuitées réalisés au cours du mois selon le motif, personnel ou professionnel. Le nombre total de voyages et de nuitées réalisés par les résidents français selon le motif est estimé, par extrapolation avec le poids individus mensuel, comme suit :

$$\text{Total_voyages}_{m,r} = \sum_k (\text{poids_mensuel_individu}_k * \text{nb_voyages_déclarés}_{m,r})$$

$$\text{Total_nuitées}_{m,r} = \sum_k (\text{poids_mensuel_individu}_k * \text{nb_nuitées_déclarés}_{m,r})$$

avec :

m : le mois correspondant à la date de retour du voyage

r : la raison du voyage, personnelle ou professionnelle

Ces totaux mensuels serviront de marges pour le calcul des poids voyages.

3.1. L'échantillonnage des voyages décrits

Parmi les voyages déclarés, seuls certains sont décrits par les enquêtés. À l'occasion de cette étape de description, l'enquêté renseigne notamment la durée, le mode d'hébergement, le mode de transport ou encore la destination d'un ou de plusieurs voyages. Ce sont les panélistes eux-mêmes qui choisissent les voyages qu'ils souhaitent décrire, recevant la consigne de privilégier les voyages les plus coûteux et donc indirectement les plus longs. Il y a donc un choix non aléatoire des voyages choisis pour la phase de description.

Pour chaque voyage, un poids unique doit être estimé. Il dépendra notamment du poids attribué à l'individu lors de l'étape précédente et permettra d'extrapoler l'ensemble des caractéristiques du voyage, notamment le nombre de nuitées. Pour tenir compte à la fois du poids de l'individu et de l'étape de sélection des voyages à décrire, le pseudo-poids voyage « initial » correspond au poids mensuel de l'individu multiplié par un coefficient de redressement correspondant au nombre de voyages déclarés rapporté au nombre de voyages décrits :

$$\text{poids_mensuel_voyage} = \text{poids_mensuel_individu} * \text{nombre de voyage déclarés} / \text{nombre de voyages décrits}$$

Cette 1^{re} étape de calcul permet d'attribuer un 1^{er} poids à chacun des voyages d'un même individu. Ce poids permet également de pondérer le nombre de nuitées, c'est-à-dire la durée de chaque voyage, ce qui pose des difficultés. Dans de nombreux cas, en effet, le ratio « voyages déclarés / voyages décrits » est supérieur au ratio « nuitées déclarées / nuitées décrites », ce qui conduit, après extrapolation, à surestimer le nombre de nuitées :

- dans plusieurs cas, il y a plus de voyages déclarés que de voyages décrits, alors que le nombre de nuitées déclarées est égal au nombre de nuitées décrites. Pour ne pas surestimer le nombre de nuitées, le coefficient de redressement est fixé à 1 ;
- les panélistes décrivent les voyages les plus longs, ce qui conduit à surestimer le nombre de nuitées ; Pour les voyages décrits de plus de 25 nuits, le coefficient de redressement est fixé à 1 . En effet, dans la mesure où l'on observe les voyages qui se sont terminés un mois donné, il est peu probable qu'un enquêté ait terminé deux voyages de 25 nuits le même mois ;
- les individus qui voyagent beaucoup, les plus aisés par exemple, sont plutôt « rares » dans l'échantillon, ils ont donc des pondérations individus assez élevées. Par ailleurs, ils sont li-

mités dans le nombre de voyages pouvant être décrits (3 maximum). À une pondération initiale élevée, s'ajoute donc un coefficient de redressement qui l'est tout autant, produisant des poids parfois très élevés. Pour éviter une sur-pondération de ces individus, le nombre de voyages déclarés est plafonné à 10. Il est en effet peu probable qu'un enquêté ait fait 10 voyages distincts tous terminés le même mois avec entre chacun un retour au domicile. Enfin, le coefficient de redressement est plafonné à 5, ce qui signifie qu'un voyage décrit ne peut au maximum représenter que 5 voyages ayant des caractéristiques (hébergement, transport, destination) identique un mois donné.

Pour chaque individu, pour un mois et pour un motif donné, le coefficient de redressement et le pseudo-poids initial des voyages s'écrivent :

$$\text{coefficient de redressement} = \min(\min(10, \text{nombre de voyages déclarés}) / \text{nombre de voyages décrits}, 5)$$
$$\text{poids_mensuel_voyage} = \text{poids_mensuel_individu} * \text{coefficient de redressement}$$

3.2. Calage sur le nombre de voyages et de nuitées déclarés

Les choix méthodologiques précédents tendent à borner le coefficient de redressement, donc à sous-estimer le nombre de voyages réalisés. À cela s'ajoutent des individus qui ne décrivent aucun des voyages déclarés. Ces voyages ne participent donc pas à la pondération des voyages. En conséquence, le nombre total de voyages et de nuitées estimés avec le poids_initial voyage sous-estime significativement le nombre total de voyages et de nuitées déclarés :

- en 2019, le nombre de voyages atteint avec le poids_initial_voyage est de 20 % inférieur au total déclaré (193,8 millions de voyages contre 246,4 millions déclarés).
- l'écart sur les nuitées est plus faible, le comportement de réponse des enquêtés qui décrivent les voyages les plus longs ne pouvant être totalement corrigé par le coefficient de redressement modélisé précédemment. Ainsi, toujours sur l'année 2019, le nombre de nuitées extrapolées par le poids_initial_voyage est inférieure de 5 % au nombre total de nuitées déclarées (1 124,7 millions de nuitées estimées contre 1 195,2 déclarées).

Une étape de calage sur marge est donc nécessaire, pour réduire le décalage entre nombre de voyages et nombre de nuitées. Elle est réalisée chaque mois en distinguant le motif de déplacement, personnel ou professionnel et utilisant comme marge le nombre total de voyages et de nuitées déclarés.

Suite à cette étape, quelques voyages ont *in fine* des poids encore trop élevés. Les poids sont tronqués de sorte qu'un voyage décrit ne puisse en représenter plus de 40 000. Cette étape a un impact acceptable sur l'estimation totale, puisque le nombre total de voyages ainsi obtenus s'élève à 245,9 millions de voyages contre 246,4 déclarés (- 0,2 %). Les poids obtenus à l'issue de cette étape seront notés par la suite *poids_mensuel_voyage*.

Percentiles de la variable poids_mensuel_voyage en milliers, calculés selon les mois de l'année 2019 :

mois	0%	1%	10%	25%	50%	75%	90%	99%	100%
janvier	0,0	0,0	0,0	1,2	3,4	5,5	7,5	14,7	40,0
février	0,0	0,0	0,0	1,6	3,6	5,5	7,5	13,3	40,0
mars	0,0	0,0	0,0	0,0	3,6	5,8	8,0	13,8	40,0
avril	0,0	0,0	0,0	0,0	3,4	5,6	7,7	13,1	40,0
mai	0,0	0,0	0,0	0,0	3,4	5,6	7,9	14,4	40,0
juin	0,0	0,0	0,0	0,0	3,4	5,9	8,7	16,8	40,0
juillet	0,0	0,0	0,0	0,0	3,5	6,0	8,6	16,1	40,0
août	0,0	0,0	0,0	0,0	3,6	6,0	8,9	17,4	40,0
septembre	0,0	0,0	0,0	0,0	3,3	5,6	8,0	16,4	40,0
octobre	0,0	0,0	0,0	0,0	3,4	5,8	8,2	16,4	40,0
novembre	0,0	0,0	0,0	0,0	3,5	6,1	8,7	17,8	40,0
décembre	0,0	0,0	0,0	0,0	3,7	6,3	8,9	17,9	40,0

La distribution des poids des voyages montre davantage de poids proches de 0 que la distribution des poids des individus. Il s'agit dans la plupart des cas de voyages réalisés par des individus qui avaient déjà un poids parmi les plus faibles, et que le calage sur marge a encore diminué. On vérifie néanmoins que la répartition des voyages tenant compte des principales caractéristiques individuelles associées aux déplacements touristiques, revenus, âge, taille du foyer, est préservée.

motif	âge	extrapolation avec le poids_voyage		extrapolation avec le poids_individu	
		Nombre de voyages	%	Nombre de voyages	%
personnel	15-24	29 907	12%	30 906	13%
	25-34	33 832	14%	33 479	14%
	35-44	35 658	14%	33 296	14%
	45-54	37 587	15%	36 049	15%
	55-64	51 840	21%	54 898	22%
	70 et plus	25 445	10%	25 900	11%
	Total	214 268	87%	214 527	87%
professionnel	15-24	3 775	2%	4 686	2%
	25-34	6 989	3%	6 421	3%
	35-44	7 843	3%	7 377	3%
	45-54	8 449	3%	8 185	3%
	55-64	4 044	2%	4 364	2%
	70 et plus	556	0%	840	0%
	Total	31 657	13%	31 873	13%
Total	245 925	100%	246 400	100%	

4. Calcul des pondérations des dépenses

Pour récupérer l'information sur les dépenses, une nouvelle étape d'échantillonnage est mise en place sur la base cette fois des voyages décrits. Cet échantillonnage dépend du panel d'origine des répondants. Pour les répondants web, il varie avec les caractéristiques des voyages : ainsi, les voyages à l'étranger, les voyages longs, les voyages coûteux, font l'objet d'une réinterrogation web systématique sur les dépenses, tandis que pour les voyages en France métropolitaine, les voyages courts et en hébergement non payant, les taux de réinterrogation sont plus faibles. Quant aux répondants « papier », ils choisissent eux-mêmes le voyage pour lequel ils décrivent les dépenses avec la consigne de choisir le plus long voyage ou le plus coûteux.

4.1. Post-stratification des voyages en groupes de dépenses homogènes

L'enquête de réinterrogation n'est pas conduite comme un sondage à 2 degrés. Les poids des individus calculés initialement ne peuvent donc être réutilisés.

Pour tenir compte autant que possible de l'échantillonnage à probabilité inégale mis en œuvre selon le type de voyage, les pondérations des dépenses sont recalculées intégralement, à partir d'un pseudo-poids initial propre aux caractéristiques des voyages, sans tenir compte des caractéristiques des individus. Ce pseudo-poids initial est donc calculé à partir d'une stratification des voyages décrits, basée sur deux types de critères :

- stratification sur les variables qui expliquent le mieux la dépense par personne participant au voyage et par jour occasionnée par un voyage.
- stratification sur les variables qui participent à la sélection du voyage

Finalement, les variables les mieux corrélées au montant des dépenses et qui participent à la sélection des voyages réinterrogés sont :

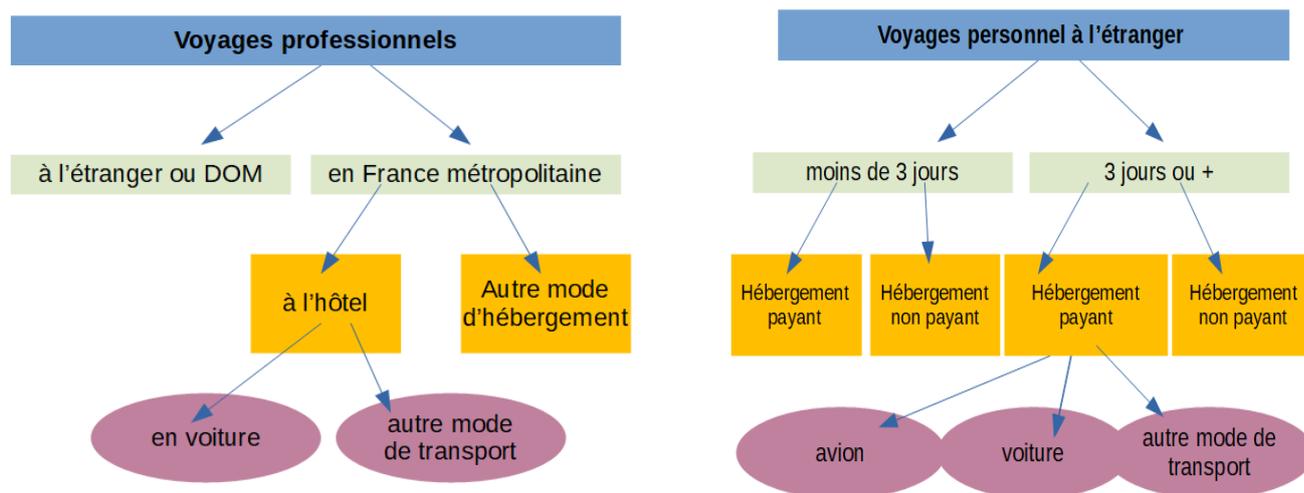
motif	durée	destination	hébergement	transport
personnel	moins de 3 jours	France métropolitaine	hôtel	voiture
	3 jours et plus		X autre payant	X avion
professionnel	X ou X	Étranger et DOM	non payant	autre
	moins de 4 jours			
	de 4 à 9 jours			
	10 jours ou plus			

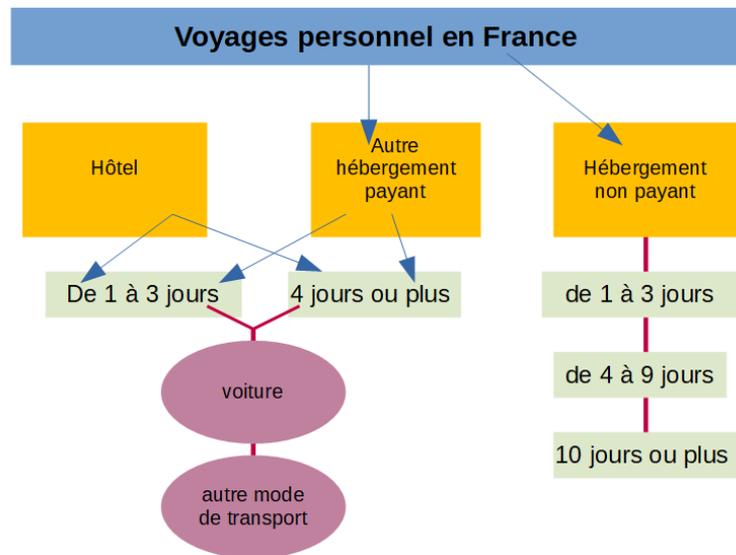
Une fois ces strates homogènes du point de vue du niveau des dépenses constituées, la taille de chaque strate a été observée, afin de s'assurer de la robustesse de la répartition. Les effectifs par strate ont été observés pour chaque mois de l'année, étant attendu que la stratification retenue in fine soit conservée pour chaque mois de l'année. En effet, il existe une forte saisonnalité en fonction du type de voyages observés. Par exemple, on observe très peu de voyages professionnels en août et très peu de voyages personnels en octobre. Finalement, compte-tenu des effectifs observés et tout en respectant des niveaux de dépenses homogènes, certaines strates ont pu être regroupées :

- En particulier, les voyages professionnels à l'étranger sont finalement regroupés en une seule strate : ils sont peu nombreux dans l'échantillon. Généralement plutôt court, à l'hôtel et avec des moyens de transport plus coûteux que la voiture, ils sont assez peu diversifiés ;

- Les voyages professionnels en France métropolitaine relèvent le plus souvent d'un hébergement payant, il n'est donc pas utile de distinguer les hébergements non payants des autres. La seule distinction de l'hôtel des autres modes d'hébergement permet d'obtenir des strates homogènes du point de vue de la dépense en hébergement ;
- Les dépenses des voyages personnels à l'étranger de courte durée présentent de fortes similarités dès lors qu'on distingue l'hébergement marchand, en général l'hôtel, du non marchand. À l'inverse, les voyages plus longs dans un hébergement payant sont plus nombreux et plus diversifiés. Pour ceux-ci, il est alors opportun de détailler les modes de transports, qui indirectement nous renseignent sur la destination et sur le mode d'hébergement plus détaillé (l'usage de la voiture pour les séjours en camping, l'avion pour les voyages plus coûteux et plus lointain, etc.) ;
- Les dépenses des voyages personnels en France métropolitaine sont très majoritaires dans l'échantillon, notamment les déplacements en hébergement non marchand. Leur taux de réinterrogation dépend de la durée. Ainsi les voyages de 3 jours ou moins sont très peu réinterrogés (taux inférieurs à 40 %), tandis que les voyages de plus de 10 jours le sont presque systématiquement (taux supérieurs à 80 %). La durée du voyage (court/moyen/long) en hébergement non marchand est donc pris en compte dans la stratification pour redresser le biais de sélection. Cela permet notamment de contrôler le nombre de nuitées final qui a tendance à être surestimé ;
- Les voyages personnels en France métropolitaine et en hébergement payant sont davantage réinterrogés, ils ont pu être stratifiés avec davantage de critères que la durée. Les séjours en hôtel ont systématiquement été isolés car cela reste le mode d'hébergement le plus coûteux. La voiture a été isolée des autres moyens de transport. Cela a permis d'isoler indirectement les moyens de transport plus coûteux comme le train et l'avion. Dans l'idéal les voyages en avion auraient pu être isolés, cependant ils restent assez rares certains mois de l'année.

Regroupement en 18 strates en fonction des effectifs mensuels :





4.2. Pseudo-poids initial et correction de la non-réponse

Pour calculer un pseudo-poids initial de dépense, il est nécessaire de passer par une étape de modélisation de la non-réponse pour distinguer l'échantillon des voyages interrogés, qui participent au calcul du pseudo-poids initial, de celui des voyages répondus, qui participe à la correction de la non-réponse.

L'information disponible sur la non-réponse pour la réinterrogation des voyages est fonction du mode de réponse. Pour les répondants web, la non-réponse n'existe pas, car la réinterrogation des dépenses a lieu dans la continuité de la description des voyages. Cette étape est transparente pour l'enquête. L'échantillon des voyages interrogés est donc le même que l'échantillon des voyages répondus pour les répondants internet.

Pour les répondants papier, la probabilité de répondre n'existe qu'au niveau de l'individu, mais pas au niveau du voyage. Pour se ramener à une information exploitable au niveau de chaque voyage, et distinguer les voyages répondus parmi les voyages interrogés, les échantillons ont été constitués comme suit :

- Soit l'individu a répondu à l'enquête de réinterrogation : le(s) voyage(s) qu'il a choisi de renseigner sont considérés comme faisant partie de l'échantillon interrogé et sont donc « répondus » ; les autres voyages qui n'ont pas été choisis par le répondant, sont considérés comme non-réinterrogés. Ils sortent de l'échantillon initial. Dans ce cas, il n'y a pas de non-réponse.
- Soit l'individu n'a pas répondu au questionnaire de réinterrogation : ne sachant pas quel(s) voyage(s) aurait(ent) été choisi(s), tous sont considérés comme devant être interrogés et font donc partie de l'échantillon initial, et tous sont considérés comme non-répondus.

Ainsi, chaque mois, pour la réinterrogation des dépenses du voyage v appartenant à la strate s , on obtient le pseudo-poids initial suivant :

$$\text{Poids_initial_réinterrogation}_v = N_s / n_s$$

avec

Poids_initial_réinterrogation_v : pseudo-poids initial des dépenses du voyage v appartenant à la strate s

N_s : taille de la strate s donnée par le poids_voyage_mensuel calculé précédemment

n_s : taille de l'échantillon des voyages interrogés comme défini précédemment pour la strate s

La correction de la non-réponse est réalisée au niveau des strates précédemment identifiées auxquelles s'ajoute le mode de réponse, papier ou web.

Ainsi, le poids pseudo-initial de la réinterrogation des dépenses du voyage v appartenant à la strate s et au mode de réponse j corrigé de la non-réponse s'écrit :

$$\mathbf{Poids_réinterrogation_CNR}_v = \mathbf{Poids_initial_réinterrogation}_v / p_{s,j}$$

avec

$Poids_réinterrogation_CNR}_v$: le pseudo-poids initial de la réinterrogation corrigé de la non-réponse

$Poids_initial_réinterrogation}_v$: le pseudo-poids initial du voyage réinterrogé v fonction de la strate s à laquelle il appartient

$p_{s,j}$: probabilité de répondre à la réinterrogation du voyage v appartenant à la strate s et au mode de réponse j , avec si $j=web$, $p_{s,j}=1$

La distribution des poids initiaux est très correcte, avec un poids maximum systématiquement inférieur à 18 000, quelle que soit la strate ou le mois observés. En revanche, on constate quelques poids extrêmes après correction de la non-réponse, induits par des taux de réponse des répondants papier particulièrement faibles. Cela concerne massivement les voyages courts, qui sont rarement choisis par les répondants papier, mais aussi les voyages pour motifs professionnels. Pourtant, il est nécessaire de rééquilibrer le poids de ce type de voyages pour réduire autant que possible le biais de sélection. Dans la suite, il paraît donc acceptable de conserver des poids plus élevés que dans les autres systèmes de pondération. Dans 5 strates, qui regroupent au total environ 80 répondants, les poids sont supérieurs à 100 000. Ils sont alors tronqués de sorte qu'un voyage décrit ne puisse en représenter plus de 100 000. Cette étape a un impact acceptable sur l'estimation totale, puisque le nombre total de voyages ainsi obtenus s'élève à 242,8 millions de voyages contre 245,8 avant l'étape de tronquage. Il y a un écart minime avec le total de voyages attendus de 245,9 millions correspondant à la somme de la pondération voyage (somme de poids_voyage) car une strate ne comporte aucun répondant : il s'agit des interrogés papier pour leurs voyages professionnels en France, en voiture et à l'hôtel au mois d'août.

Le biais de sélection des voyages réinterrogés a un impact sur le nombre de nuitées

La stratification des voyages tenant compte autant que possible de la durée du voyage ne permet pas de résorber complètement le biais induit par la sélection systématique des voyages les plus longs. Le nombre de nuitées calculé avec le pseudo-poids initial redressé de la non-réponse calculé précédemment atteint 1 252,534 millions de nuitées contre 1 194,307 déclarées par les individus interrogés (+ 5 %). Une étape de calage sur marge est donc nécessaire pour recalculer les nombres de voyages et de nuitées sur les marges déclarées.

4.3. Calage sur le nombre de voyages et de nuitées par type d'hébergement

L'étape de calage sur marge est nécessaire pour recalculer le nombre de nuitées sur les totaux déclarés. Autant que possible en fonction des effectifs, le calage tient compte des différents types d'hébergement, car ces données sont utilisées pour différentes opérations, notamment pour répartir les dé-

penses dans le cadre du Compte Satellite du Tourisme (CST). Finalement, l'étape de calage repose sur les critères suivants :

- le nombre de voyages et de nuitées pour motif personnel est recalé chaque mois sur les totaux déclarés par type d'hébergement en détaillant les modes d'hébergement payant et en regroupant l'hébergement non payant. Au final, 5 modalités sont conservées : hôtel, camping, gîte et location, autre hébergement payant et hébergement non payant ;
- le nombre de voyages et de nuitées pour motif professionnel est recalé chaque mois sur des marges différentes. D'une part, certains mois, les effectifs sont faibles, ce qui limite le nombre de marges possibles. D'autre part, les voyages à l'hôtel sont très majoritaires, tandis que les autres modes d'hébergement sont plus rares, marchand, comme non marchand. Pour les voyages professionnels, 2 modalités sont conservées : hôtel et autre hébergement (payant et non payant).

Les poids issus du calage sur marge atteignent parfois quelques valeurs élevées, notamment les voyages courts des mois d'été qui sont systématiquement non décrits au profit des voyages longs de la période estivale. Pour limiter les valeurs extrêmes, tout en conservant des poids permettant de redresser l'important biais de sélection des voyages, les poids sont tronqués à 130 000. Un seul voyage est concerné.

Les poids obtenus in fine seront notés par la suite *poids_mensuel_dépense*.

Percentiles de la variable poids_mensuel_dépense en milliers, calculés selon les mois de l'année 2019 :

mois	0%	1%	10%	25%	50%	75%	90%	99%	100%
janvier	2,1	3,3	4,2	4,6	6,0	7,4	8,7	31,3	31,3
février	1,9	3,1	4,1	5,2	6,3	8,1	10,2	27,0	27,0
mars	1,2	3,4	4,2	5,1	6,3	7,0	8,6	52,0	52,3
avril	1,4	3,2	3,9	5,0	8,1	11,2	18,2	22,4	22,4
mai	1,0	3,4	3,8	5,1	6,9	8,6	11,6	16,6	22,1
juin	0,1	3,1	4,2	5,3	7,8	9,1	11,9	18,3	18,3
juillet	1,7	2,9	3,5	4,0	4,8	6,8	11,0	111,7	125,0
août	1,2	3,0	4,2	4,5	5,4	8,6	10,0	82,6	106,4
septembre	0,6	2,8	4,0	4,4	5,1	9,0	11,3	123,1	130,0
octobre	1,2	3,2	4,1	4,8	6,7	10,0	14,0	65,1	65,5
novembre	1,3	2,6	3,4	4,5	5,9	6,6	7,5	11,9	13,2
décembre	0,4	2,9	4,3	4,4	5,3	6,8	14,0	71,5	71,5

4.4. Le cas particulier de l'année 2020

En 2020, le calcul des poids de réinterrogation n'a pas été corrigé de la non-réponse, car la non-réponse papier, habituellement élevée pour certaines strates, était encore plus massive, consécutive aux différentes périodes de confinement. La principale conséquence est que pour le mode d'interrogation papier, de nombreuses strates avaient un taux de réponse très faible, donc mécaniquement un poids corrigé de la non-réponse qui devenait très élevé, un seul voyage papier répondu en représentant un très grand nombre. Pour éviter de produire des poids extrêmes et faire peser sur quelques voyages, les dépenses de beaucoup d'autres, cette étape n'a pas été réalisée sur les mois de l'année 2020. Cela signifie que seules les caractéristiques des voyages (motif, durée, hébergement, transport, etc) ont participé au calcul de la pondération des dépenses avant le calage sur marge.

En 2020, l'étape de calage a elle aussi été adaptée. Pour les mois d'avril et de novembre 2020, mois entièrement confinés avec donc très peu de voyages recensés, le calage sur les nuitées personnelles a été réduit à deux marges, nuitées marchandes et nuitées non marchandes.

5. Calcul des pondérations des séjours

À l'occasion d'un voyage, un ou plusieurs lieux, encore appelés séjours, sont visités. Au moment de la réinterrogation, l'enquêté déclare le nombre total de lieux différents qu'il a visité lors de son voyage, qui sera noté par la suite *nombre_lieux_déclarés*. Une partie d'entre eux seulement fait l'objet d'une description plus approfondie, notée par la suite *nombre_lieux_décrits*.

Le poids attribué à ces différents lieux de séjour est directement déduit du poids de réinterrogation calculé précédemment comme suit :

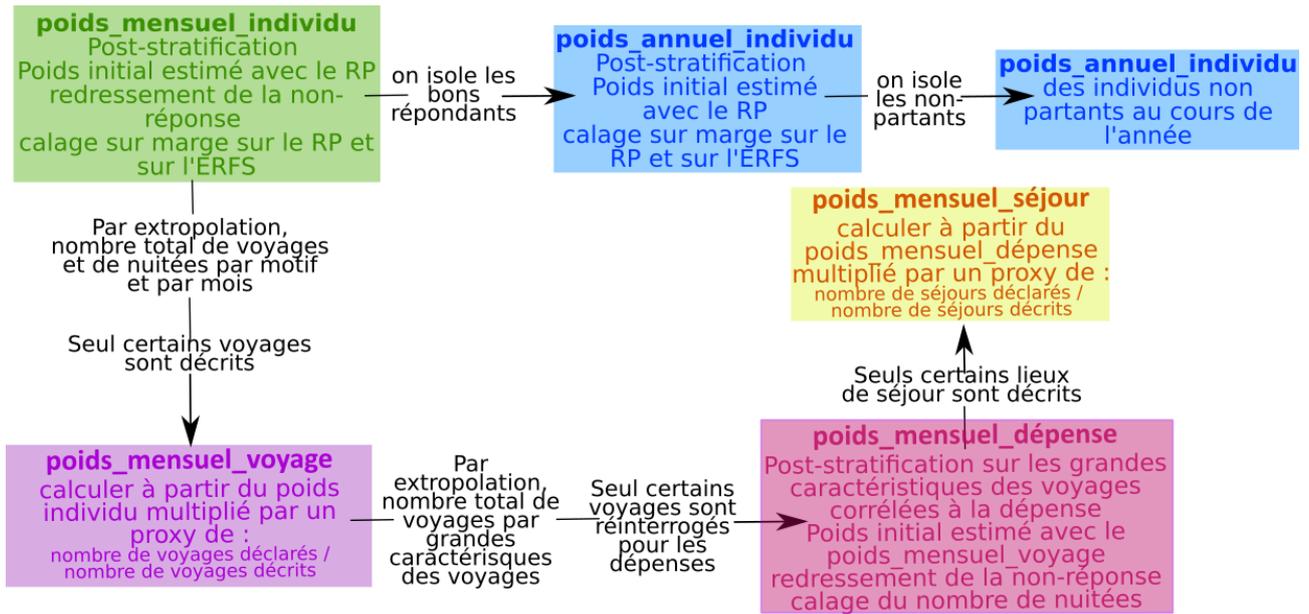
$$\text{poids_mensuel_séjour} = \text{poids_mensuel_réinterrogation} * \text{nombre_lieux_déclarés} / \text{nombre_lieux_décrits}$$

Comme pour les poids *poids_mensuel_dépense*, les poids sont bornés à 80 000, car il y a quelques valeurs extrêmes, notamment pour les voyageurs qui ont décrit un seul lieu de séjour pour 4 ou 5 déclarés.

Percentiles de la variable *poids_mensuel_séjour* en milliers, calculés selon les mois de l'année 2019 :

mois	0%	1%	10%	25%	50%	75%	90%	99%	100%
janvier	3,7	3,7	4,3	4,6	6,5	8,5	14,9	29,6	59,7
février	3,4	3,4	4,6	5,7	6,7	9,1	15,4	29,3	52,8
mars	3,3	3,9	4,4	5,8	6,1	9,2	16,7	33,4	80,0
avril	3,8	3,8	4,1	5,8	9,2	13,8	15,1	37,1	72,0
mai	3,5	3,9	4,1	5,6	7,1	9,5	14,2	28,3	60,9
juin	3,1	3,6	4,4	6,7	8,3	11,8	12,8	32,7	59,1
juillet	3,6	3,6	4,9	5,2	5,6	9,0	13,3	35,5	80,0
août	3,6	3,7	4,9	5,7	6,0	10,1	21,0	63,6	80,0
septembre	3,2	3,3	4,0	4,8	6,1	10,6	22,2	69,3	80,0
octobre	3,3	3,3	4,4	5,0	7,7	13,2	24,8	49,6	80,0
novembre	3,9	4,1	4,6	4,8	5,0	6,4	7,3	23,4	62,4
décembre	4,6	4,6	5,9	6,7	6,7	15,8	16,3	40,7	65,4

6. Synthèse des étapes de calculs



7. Bilan

La refonte des différents systèmes de pondération des données de l'enquête SDT a des effets significatifs sur les résultats. D'une manière générale, les niveaux des principales variables issues de l'enquête sont plus élevés qu'initialement, ce qui semble plus cohérent avec les autres sources d'information disponibles sur le sujet, notamment avec l'enquête de fréquentation hôtelière. Les résultats issus des pondérations de SDT avec la précédente méthode sous-estimaient significativement le nombre de voyages et de nuitées réalisés par les résidents français et en particulier les déplacements professionnels. Les dépenses réalisées au cours des voyages, examinées dans le cadre du compte satellite du tourisme, semblaient tout aussi sous-estimées par l'enquête.

Les comparaisons entre l'ancien et le nouveau système de pondération menées sur l'année 2019 montrent en effet une réestimation à la hausse du nombre de voyages réalisés de l'ordre de + 14 %, et notamment des voyages professionnels (+ 26 %). Les nuitées sont réévaluées globalement de 9 %, et les nuitées professionnelles de 20 %. La repondération des données individuelles a en effet permis de donner plus de poids aux grands voyageurs, qui réalisent davantage de voyages pour motif personnel mais aussi plus de voyages professionnels, mais moins longs. Les taux de départ progressent de l'ordre de 5 points, 4 points pour les voyages pour motif personnel et de l'ordre de 1 point pour les voyages pour motif professionnel.

Effet du changement de pondération des individus sur le nombre de voyages :

en milliers de voyages

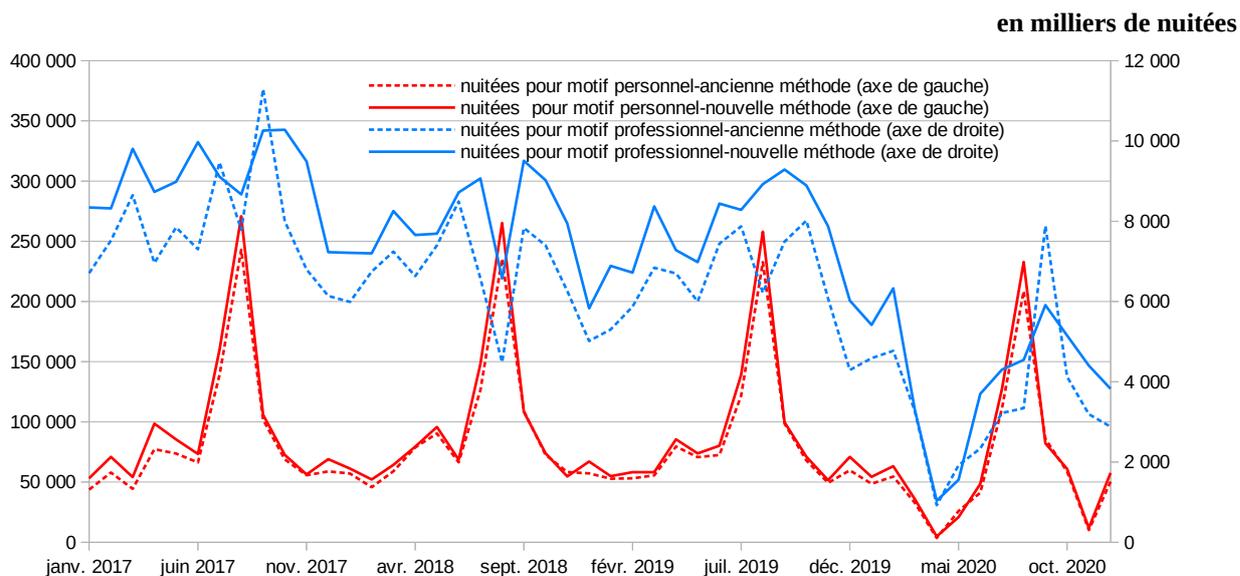
année	pondération individu ancienne méthode			pondération individu nouvelle méthode			Écart ancienne / nouvelle méthode		
	voyages personnels	voyages professionnels	total voyages	voyages personnels	voyages professionnels	total voyages	voyages personnels	voyages professionnels	total voyages
2018	193 346	26 265	219 611	214 003	31 967	245 971	+10,7%	+21,7%	+12,0%
2019	191 624	25 313	216 937	214 527	31 873	246 400	+12,0%	+25,9%	+13,6%
2020	130 068	13 646	143 714	148 248	15 690	163 939	+14,0%	+15,0%	+14,1%

Effet du changement de pondération des individus sur le nombre de nuitées :

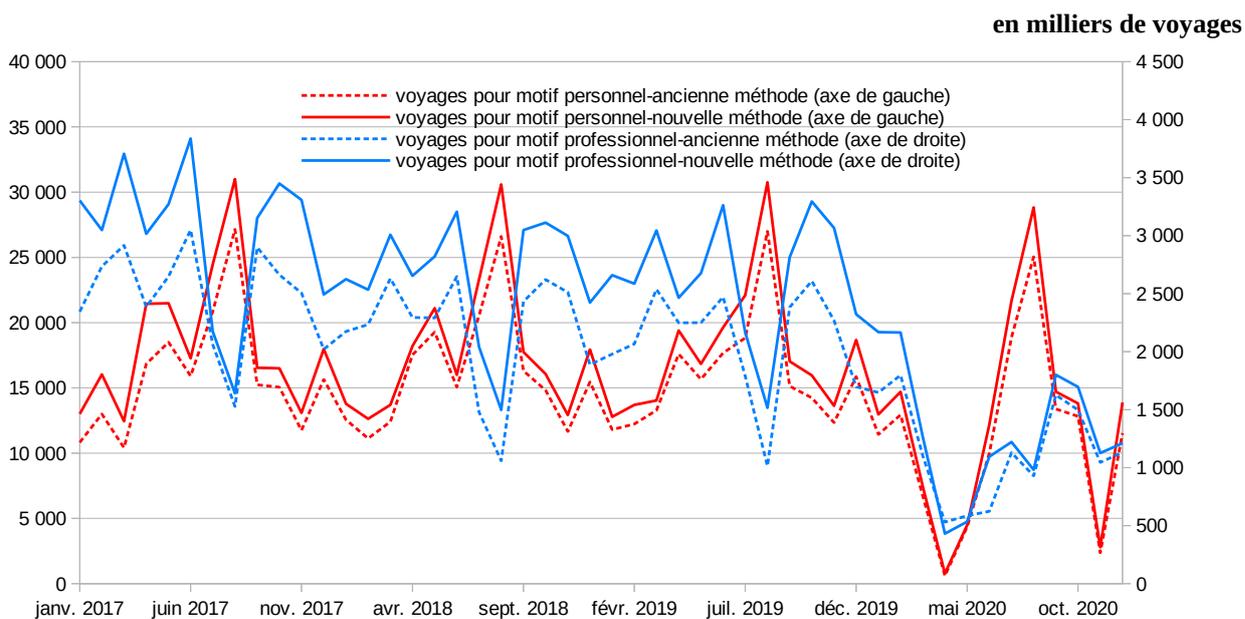
en milliers de nuitées

année	pondération individu ancienne méthode			pondération individu nouvelle méthode			Écart ancienne / nouvelle méthode		
	nuitées personnelles	nuitées professionnelles	total nuitées	nuitées personnelles	nuitées professionnelles	total nuitées	nuitées personnelles	nuitées professionnelles	total nuitées
2018	1 056 589	80 077	1 136 666	1 138 635	94 678	1 233 313	+7,8%	+18,2%	+8,5%
2019	1 014 454	78 117	1 092 572	1 101 222	93 950	1 195 172	+8,6%	+20,3%	+9,4%
2020	731 661	42 432	774 093	800 682	49 446	850 129	+9,4%	+16,5%	+9,8%

Évolution du nombre de nuitées mensuelles selon le motif et la pondération individu utilisée :



Évolution du nombre de voyages mensuels selon le motif et la pondération individu utilisée :



Évolution du taux de départ selon la pondération individu utilisée :

Taux de départ 2019 selon le mois, le motif et la pondération individu utilisée :

année 2019	pondération ancienne méthode			pondération nouvelle méthode			écart		
	Taux de départ	Taux de départ pour motif personnel	Taux de départ pour motif professionnel	Taux de départ	Taux de départ pour motif personnel	Taux de départ pour motif professionnel	Taux de départ	Taux de départ pour motif personnel	Taux de départ pour motif professionnel
janvier	20,8%	18,9%	3%	23,5%	20,7%	4%	+2,6 %	+1,8 %	+1,0 %
février	22,0%	20,0%	3%	25,3%	22,7%	4%	+3,3 %	+2,7 %	+0,9 %
mars	22,4%	20,0%	3%	25,0%	22,0%	4%	+2,6 %	+2,0 %	+0,8 %
avril	29,7%	27,7%	3%	33,8%	31,4%	3%	+4,1 %	+3,6 %	+0,7 %
mai	25,8%	23,9%	3%	29,1%	26,5%	4%	+3,3 %	+2,6 %	+0,9 %
juin	29,4%	27,2%	3%	32,8%	29,9%	5%	+3,3 %	+2,6 %	+1,3 %
juillet	31,3%	29,8%	2%	36,2%	34,4%	3%	+4,9 %	+4,6 %	+0,7 %
août	43,1%	42,2%	1%	48,4%	47,1%	2%	+5,3 %	+5,0 %	+0,8 %
septembre	25,7%	23,3%	3%	29,5%	26,5%	4%	+3,8 %	+3,2 %	+0,9 %
octobre	24,7%	22,4%	3%	28,7%	25,5%	5%	+4,0 %	+3,1 %	+1,3 %
novembre	21,2%	18,9%	3%	24,3%	21,3%	4%	+3,1 %	+2,4 %	+1,2 %
décembre	26,9%	25,6%	2%	31,7%	29,8%	3%	+4,7 %	+4,2 %	+1,1 %

La structure des voyages selon les principales caractéristiques change également. En lien avec la ré-évaluation à la hausse du nombre de voyages professionnels, le poids des voyages à l'hôtel est significativement plus grand. À l'exception des nuitées en camping, les nuitées de tous les modes d'hébergement payant sont réestimées à la hausse, au détriment surtout de la résidence secondaire. S'agissant des transports, les voyages en avion progressent, de même que les voyages en train au détriment de la voiture. Cela est cohérent avec la progression plus marquée des voyages à l'étranger.

Nombre de nuitées selon l'année, le mode d'hébergement et la pondération voyage utilisée

Mode d'hébergement principal	pondération voyage ancienne méthode		pondération voyage nouvelle méthode		en milliers de nuitées	
	2019	2020	2019	2020	Écart	
	2019	2020	2019	2020	2019	2020
Hôtel	181 228	69 777	230 952	92 451	+27,4 %	+32,5 %
Camping	90 260	60 465	95 949	66 108	+6,3 %	+9,3 %
Location, gîte ou chambre d'hôte	176 857	121 420	202 115	143 324	+14,3 %	+18,0 %
Autre hébergement payant	69 146	44 379	78 286	52 130	+13,2 %	+17,5 %
Total hébergement payant	517 491	296 040	607 302	354 012	+17,4 %	+19,6 %
Résidence secondaire	147 371	132 933	130 689	116 290	-11,3 %	-12,5 %
Famille	328 350	269 982	351 915	300 880	+7,2 %	+11,4 %
Amis	70 959	54 972	73 102	56 664	+3,0 %	+3,1 %
Autre hébergement non payant	28 876	20 226	31 301	21 436	+8,4 %	+6,0 %
Total hébergement non payant	575 555	478 113	587 006	495 270	+2,0 %	+3,6 %
Total	1 093 046	774 153	1 194 307	849 281	+9,3 %	+9,7 %

Nombre de nuitées selon l'année, le mode de transport et la pondération voyage utilisée

Mode de transport principal	pondération voyage ancienne méthode		pondération voyage nouvelle méthode		Écart	
	2019	2020	2019	2020	2019	2020
	en milliers de nuitées					
Voiture	715 900	595 919	766 531	646 563	+7,1 %	+8,5 %
Camping-car	19 376	13 666	17 740	12 324	-8,4 %	-9,8 %
Deux roues	2 257		2 467		+9,3 %	
Autocar	19 856	7 065	20 547	8 124	+3,5 %	+15,0 %
Train	127 606	87 662	136 513	96 222	+7,0 %	+9,8 %
Bateau	11 167	3 956	11 098	4 715	-0,6 %	+19,2 %
Avion	192 631	63 062	234 815	77 860	+21,9 %	+23,5 %
Autre	4 254	2 822	4 597	3 472	+8,1 %	+23,0 %
Total	1 093 046	774 153	1 194 307	849 281	+9,3 %	+9,7 %

Nombre de nuitées selon l'année, la destination et la pondération voyage utilisée

Destination	pondération voyage ancienne méthode		pondération voyage nouvelle méthode		Écart	
	2019	2020	2019	2020	2019	2020
	en milliers de nuitées					
France métropolitaine	850 969	701 631	908 257	761 681	+6,7%	+8,6%
DOM	12 223	5 739	13 510	6 456	+10,5%	+12,5%
étranger	229 854	66 783	272 540	81 144	+18,6%	+21,5%
Total Résultat	1 093 046	774 153	1 194 307	849 281	+9,3%	+9,7%

Les dépenses associées sont également revalorisées. La dépense totale réalisée par les résidents français en 2019, toutes destinations confondues, est réévaluée à 167 milliards d'euros en 2019, soit + 17 % par rapport à l'estimation initiale. L'augmentation du nombre total de voyage explique ce résultat (contribution de 19 points), la stratification des voyages qui tend à repondérer les voyages courts (- 2 points) influence peu le niveau de la dépense finale.

Dépenses réalisées au cours des voyages touristiques selon l'année, le type de dépense et la pondération dépense utilisée

Type de dépense	pondération dépense ancienne méthode		pondération dépense nouvelle méthode		Écart	
	2019	2020	2019	2020	2019	2020
	en milliards d'euros					
Hébergement	30,7	20,7	35,6	24,9	16 %	20 %
Restauration	16,2	9,7	18,7	11,3	15 %	17 %
Transport	29,0	16,4	33,9	19,5	17 %	19 %
Autre	66,4	32,7	79,0	37,5	19 %	14 %
Total	142,4	79,5	167,2	93,2	17 %	17 %