

SEMI RURAL

CALCUL DES PONDERATIONS

L'enquête Semi-rural 2012 repose sur un sondage à 4 degrés.

Pour chaque degré, nous allons calculer le poids de sondage des unités échantillonnées puis procéder à une correction de la non-réponse si cela est nécessaire et possible.

Plusieurs jeux de poids vont être produits, correspondant aux différents types de questionnaires collectés :

- Une pondération commune aux QIL & QAA
- Une pondération spécifique à l'exploitation des QAA
- Une pondération spécifique à l'exploitation des QIL.

A. Pondération commune QIL & QAA

1) 1^{er} niveau : Agglomération

La sélection des agglomérations s'est faite de façon raisonnée, donc non aléatoire.

1) Poids de sondage

La sélection n'étant pas aléatoire, il n'existe pas de probabilités de sondage, donc pas de poids de sondage. Les résultats de l'enquête ne sont extrapolables qu'aux agglomérations enquêtées.

2) Correction de la non-réponse

Il n'existe pas de non-réponse au niveau agglomération. Il n'y a donc pas de correction à réaliser.

2) 2^{ème} niveau : Service

Afin d'assurer la représentativité de l'enquête, tous les services pour les sans-domicile implantés dans les agglomérations sélectionnées ont été sollicités pour répondre à l'enquête. A ce niveau-là, il s'agit donc d'un recensement.

a) Poids de sondage

Dans le cas d'un recensement, toutes les unités ont une probabilité égale à 1. Le poids de sondage des services vaut donc : $W_{\text{service}}=1$.

b) Correction de la non-réponse

Les informations disponibles sur les services non-répondants étant très rares, nous préférons ne pas extrapoler les résultats obtenus des services répondants aux services non répondants. L'enquête sera donc représentative uniquement des services enquêtés.

3) 3ème niveau : Visite

Si la fréquentation quotidienne estimée d'un centre était inférieure à 33 individus, une seule visite était prévue. Si la fréquentation estimée était égale ou supérieure à 33 individus, deux visites étaient prévues.

a) Poids de sondage

Le poids de sondage des visites dépend du nombre de visites prévues dans le service :

- Si la fréquentation estimée < 33 : **WS_VISITE** = 1
- Si la fréquentation estimée ≥ 33 : **WS_VISITE** = 0,5

b) Correction de la non-réponse

Pour être cohérent avec la volonté de ne pas appliquer les résultats d'un service à un autre, la seule correction de la non-réponse au niveau visite que nous nous autorisons est de remplacer une visite par une autre si elle est réalisée dans le même service.

Parmi les 81 services ciblés par l'enquête, seuls 6 devaient être enquêtés deux fois (deux visites prévues). Parmi ces 6 services, un seul n'a honoré qu'une seule des deux visites prévues (pour les autres, soit le service est non-répondant, soit les deux visites ont eu lieu).

Pour ce service visité une seule fois au lieu de 2, on applique à la visite ayant eu lieu un coefficient de correction de 2 noté **C_VISITEQILQAA**.

Pour toutes les autres visites, il n'y a pas de correction.

Pour obtenir le poids de sondage corrigé de la non-réponse au niveau visite, on réalise le calcul suivant :

$$\mathbf{W_VISITEQILQAA = WS_VISITE \times C_VISITEQILQAA.}$$

4) 4ème niveau : Individu

En fonction de la fréquentation estimée lors de la pré-visite, un nombre de personnes à solliciter a été fixé en amont de la visite.

a) Poids de sondage

Le poids de sondage correspond donc au ratio du nombre de personnes présentes lors de la visite sur le nombre de personnes effectivement sollicitées lors de la visite.

$$\mathbf{WS_INDQILQAA = Nb_resents / Nb_sollicites}$$

b) Correction de la non-réponse

Parmi les personnes sollicitées, on compte des refusants. Afin de corriger la non-réponse à l'échelle individu, on multiplie le poids de sondage par le ratio du nombre de personnes sollicitées sur le nombre de questionnaires collectés.

$$C_INDQILQAA = Nb_sollicites / Nb_quest_visite$$

Pour obtenir le poids de sondage corrigé de la non-réponse au niveau individu, on réalise le calcul suivant :

$$W_INDQILQAA = WS_INDQILQAA \times C_INDQILQAA$$

5) Calcul du poids de sondage corrigé de la non-réponse

Une fois le poids de sondage calculé et les corrections de la non-réponse réalisées à chaque degré, on peut calculer le poids de sondage final en multipliant les différents poids calculés.

$$W_QILQAA = W_VISITEQILQAA \times W_INDQILQAA$$

6) Redressement - Correction de la proportion de francophones

Une fois ces étapes de calcul réalisées, on contrôle si la proportion de francophones parmi les répondants (en utilisant la pondération) est identique à celle calculée parmi les personnes présentes lors de l'enquête.

Dans les services enquêtés, 620 individus dont 432 francophones ont été recensés. La proportion de francophones est donc de 69,68%. Via l'échantillon des répondants (QIL+QAA) pondéré par le poids obtenu dans le point 5), la proportion de francophones est estimée à 87.75%. Cette surreprésentation des francophones parmi les répondants est l'effet conjugué du plan de sondage (qui conduit effectivement à surreprésenter les francophones car ciblés par QIL & QAA) et d'un potentiel taux de réponse plus favorable parmi les francophones.

	Répondants Effectifs bruts		Répondants Effectifs pondérés avec W_{QILQAA}		Effectifs recensés lors des visites (cible)	
Fr.	281	86.46	544	87.75	432	69.69
Non fr.	44	13.54	76	12.25	188	30.31
Total	325	100.00	620	100.00	620	100.00

Afin de corriger ce biais, un coefficient correctif a été appliqué en fonction du statut francophone/non francophone.

- Questionnaires francophones : $C_FRQILQAA = 432/544 = 0.79$
- Questionnaires non francophones : $C_FRQILQAA = 188/76 = 2.47$

On appellera $W_QILQAAFR$ la pondération résultant de cette opération :

$$W_QILQAAFR = W_QILQAA \times C_FRQILQAA$$

7) Statistiques des poids

	Total	Moyenne	CV	Min	Max	Max/min
W_QILQAA	620	1.908	52.587	0.5	5.0	10
W_QILQAAFR	620	1.908	83.906	0.397	10.955	27.594

B. Pondération QIL uniquement

Le calcul de la pondération QIL diffère de celui de la pondération QIL+QAA à partir du niveau individu uniquement. Les calculs relatifs aux 3 premiers niveaux sont identiques à ceux présentés dans la partie précédente.

1) 1^{er} niveau : Agglomération

Cf. §A.1.

2) 2^{ème} niveau : Service

Cf. §A.2.

3) 3^{ème} niveau : Visite

Cf. §A.3.

4) 4^{ème} niveau : Individu

Pour chaque visite, un maximum de 4 QIL était attendu. Si, lors de la visite, on rencontrait 4 francophones ou moins, on proposait alors à chacun d'eux de participer à l'enquête et de remplir un QIL. A contrario, si l'effectif de francophones rencontrés était supérieur, seuls 4 d'entre eux étaient sollicités pour remplir un QIL, les autres se voyaient proposer un QAA.

a) Poids de sondage

Le poids de sondage correspond donc au ratio du nombre de personnes francophones présentes lors de la visite sur le nombre de personnes effectivement sollicitées pour réaliser un QIL lors de la visite.

Le poids des QIL dépend du nombre de francophones sollicités lors de la visite :

- Si $Nb_Fr_Sollicites > 4$: $WS_INDQIL = Nb_fr_Presents / 4$
- Si $Nb_Fr_Sollicites \leq 4$: $WS_INDQIL = Nb_fr_Presents / Nb_Fr_Sollicites$

b) Correction de la non-réponse

Là encore, le coefficient de correction de la non-réponse dépend du nombre de francophones sollicités lors de la visite :

- Si $Nb_Fr_Sollicites > 4$: $C_INDQIL = 4 / Nb_QIL_visite$
- Si $Nb_Fr_Sollicites \leq 4$: $C_INDQIL = Nb_Fr_Sollicites / Nb_QIL_visite$

Pour obtenir le poids de sondage corrigé de la non-réponse au niveau individu, on réalise le calcul suivant :

$$W_INDQIL = WS_INDQIL \times C_INDQIL$$

5) Calcul du poids de sondage corrigé de la non-réponse

Une fois le poids de sondage calculé et les corrections de la non-réponse réalisées à chaque degré, on peut calculer le poids de sondage final en multipliant les différents poids calculés.

$$W_QIL = W_VISITEQILQAA \times W_INDQIL$$

6) Redressement – Correction de la proportion de francophones

Pour la pondération des QIL, cette étape n'est pas nécessaire puisque les QIL ne sont réalisés qu'auprès d'individus francophones.

7) Statistiques des poids

	Total	Moyenne	CV	Min	Max	Max/min
W_QIL	432	2.5714	56.694	0.502	8.035	16.006

C. Pondération QAA uniquement

Tout comme pour le calcul QIL, la méthode de calcul QAA change uniquement au niveau individu. Les calculs relatifs aux 3 premiers niveaux sont identiques à ceux présentés dans la première partie.

1) 1^{er} niveau : Agglomération

Cf. §A.1.

2) 2^{ème} niveau : Service

Cf. §A.2.

3) 3^{ème} niveau : Visite

Mêmes calculs que ceux décrits en §A.2 mais le champ et les résultats sont différents.

Les variables portent donc les noms suivants : **WS_VISITEQAA**, **C_VISITEQAA** et **W_VISITEQAA**.

4) 4ème niveau : Individu

Pour chaque visite, un nombre de QAA à réaliser était fixé en amont de la visite.

a) Poids de sondage

Le poids de sondage correspond donc au ratio du nombre de personnes présentes lors de la visite sur le nombre de personnes effectivement sollicitées pour réaliser un QAA lors de la visite. Le nombre de personnes sollicitées pour réaliser un QAA correspond à la différence entre le nombre de questionnaires total attendu et l'objectif en termes de QIL.

- **Nb_sollicites_QIL = min(Nb_Fr_sollicites, 4)**
- **Nb_sollicites_QAA = Nb_sollicites - Nb_sollicites_QIL**
- **WS_INDQAA = Nb_presents / Nb_sollicites_QAA**

b) Correction de la non-réponse

On procède de la même façon que précédemment pour corriger la non-réponse à l'échelle individu.

$$C_INDQAA = Nb_Sollicites_QAA / Nb_QAA_visite$$

$$W_INDQAA = C_INDQAA * WS_INDQAA$$

5) Calcul du poids de sondage corrigé de la non-réponse

Une fois le poids de sondage calculé et les corrections de la non-réponse réalisées à chaque degré, on peut calculer le poids de sondage final en multipliant les différents poids calculés.

$$W_QAA = W_VISITEQAA \times W_INDQAA$$

6) Redressement – Correction de la proportion de francophones

Comme pour la pondération commune QIL et QAA, on contrôle ensuite si la proportion de francophones parmi les répondants (en utilisant la pondération) est identique à celle calculée parmi les personnes présentes dans les services pour lesquels des QAA ont été collectés.

Dans les services enquêtés par QAA, 554 individus dont 384 francophones ont été recensés. La proportion de francophones est donc de 69,31%. Via l'échantillon des répondants QAA pondéré par le poids W_{QIL} , la proportion de francophones est estimée à 68,24%. En appliquant un coefficient différent aux QAA en fonction de la langue parlée par le répondant, on va redresser l'échantillon pour que les effectifs et la répartition francophones/non francophones collent à la cible.

	Répondants Effectifs bruts		Répondants Effectifs pondérés avec W_{QAA}		Effectifs recensés lors des visites avec QAA (cible)	
Fr.	113	71.97	370	75.82	333	68.24
Non fr.	44	28.03	118	24.18	155	31.76
Total	157	100.00	488	100.0	488	100.0

Afin de corriger ce biais, les coefficients correctifs sont les suivants :

- Questionnaires francophones : $C_{FRQAA} = 333/370 = 0.9$
- Questionnaires non francophones : $C_{FRQAA} = 155/118 = 1.31$

On appellera W_{QAAFR} la pondération résultant de cette opération : $W_{QAAFR} = W_{QAA} \times C_{FRQAA}$

8) Statistiques des poids

	Total	Moyenne	CV	Min	Max	Max/min
W_{QAA}	488	3.11	86.05	1	23	23
W_{QAAFR}	488	3.11	83.43	1.05	20.7	19.71