

Règles d'attribution des caractéristiques sociodémographiques des contacts

Variables n_sexe_ci ; n_age_ci ; n_pcs_ci ; n_nat_ci

Cas 1 : la personne est enquêtée

- On prend ses caractéristiques déclarées.
- La variable source= « directe »

Cas 2 : la personne n'est pas enquêtée (n=1510)

Cas 2.1 : Le contact n'est cité qu'une fois (945 contacts).

Cas 2.1.1 : Et les variables G_C sexe ; G_C age ; G_C CAT (ou CAT2) et G_C PN sont renseignées.

- On prend toutes les infos données par le contact.
- La variable source= « indirecte1 »

Cas 2.1.2 : Et les variables G_C sexe ; G_C age ; G_C CAT (ou CAT2) et G_C PN sont vides.

- Cela signifie qu'il s'agit d'une autre configuration qu'une personne seule (les informations n'ont alors pas été demandées)
- Les variables prennent la variable 66. Question non posée.
- La variable source= « indirecte1 »

Cas 2.2 : Le contact est cité plusieurs fois. (565 contacts qui correspondent à 236 codes différents)

191 contacts sont de même sexe et d'âge similaire ; 45 sont soit d'âge, soit de sexe différent (on en déduit que ce sont deux personnes différentes à l'adresse qui ont été enquêtées, et on ne conservera dans ce cas que le pays de naissance ou le sexe si toutes les valeurs correspondent).

Source= « indirecteX », X valant le nombre de personnes qui citent

Âge :

- Si l'âge déclaré par tous les citants correspond à la même tranche d'âge, celle-ci est attribuée.
- Si plus de deux tranches d'âges sont citées et qu'une est majoritaire, on attribue cette variable.
- Si deux tranches d'âge proche sont citées, on attribue la plus âgée des deux (suivant la tendance observée chez les contacts ayant été enquêtés qui ont souvent été rajeunis).
- Si toutes les citations sont vides (autres configurations qu'une personne seule) ou NSP, on attribue la valeur « 66. Question non posée ».
- Si le sexe des contacts est différent et/ou les tranches d'âge sont trop différentes (plus d'une tranche d'écart) entre les citations, on attribue la valeur « 77. Plusieurs contacts cités »

Pays de naissance :

- 191 contacts pour lesquels toutes les déclarations de pays de naissance sont les mêmes. Cette valeur est attribuée (il peut s'agir de contacts différents à la même adresse, mais ayant les mêmes pays de naissance, donc on conserve l'information).
- 11 cas pour lesquels le pays est différent et on avait identifié que c'était des personnes différentes, donc nat=77. Plusieurs contacts cités

-
- 25 cas pour lesquels toutes les déclarations sont "NSP". Donc nat=NSP.
- 6 cas avec une réponse plus fréquente. On code la réponse la plus fréquente.
- 5 cas à deux citations avec un pays France, un pays étranger. On privilégie le pays étranger.
- Un cas Algérie/Tunisie : on code Tunisie pour pouvoir faire des analyses sur la région
- Deux cas avec deux pays de deux régions du monde, on code nat=NSP.

PCS :

- Parmi les 191 cas où on estime que c'est le même contact, 137 ont la même PCS et 54 ont des déclarations différentes.
- Si la même PCS déclarée, elle est attribuée.
- Si plusieurs déclarations et une est majoritaire, elle est attribuée.
- Sinon, les valeurs sont choisies en fonction de l'analyse des bonnes déclarations réalisées sur la base des contacts interrogés (voir annexe - Correction des incohérences dans les professions des contacts doublons
- Si les contacts sont différents, pcs = 77. Plusieurs contacts différents.
- Si toutes les valeurs sont vides ou NSP, pcs=66. Question non posée
- Si les déclarations ne permettent pas de trancher, pcs=99.

Sexe :

- Si le sexe est le même dans toutes les déclarations, on le conserve (même s'il peut s'agir de contacts différents).
- Si deux sexes différents : sexe=77. Plusieurs contacts différents.
- Si toutes les valeurs sont vides ou NSP, sexe=66. Question non posée.

Annexe : Correction des incohérences dans les professions des contacts doublons

Comparaison des réponses 4 et 5, à partir de 653 enquêtés

Quand le citant dit employé : la moitié sont bien estimés, sinon 16 % sont des professions intermédiaires.

Quand le citant dit professions intermédiaires : seuls 32 % sont bien estimés et 22 % sont des employés.

Quand un citant dit 4 et l'autre dit 5, on a donc 36 % de chance que ce soit 5 et 24 % de chance que ce soit 4. Si on admet que, vu les deux réponses, la bonne réponse est forcément entre 4 et 5, on tombe à 58 % de chance que la bonne réponse soit 5 versus 42 % qu'elle soit 4.

Conclusion : Pcs = 4 ou 5 → Correction à Pcs=5

9 corrections

Comparaison des réponses 3 et 4, à partir de 446 enquêtés

Quand le citant dit cadres : 51 % de bonne réponse, sinon 27 % de professions intermédiaires

Quand le citant dit professions intermédiaires : seuls 32 % sont bien estimés et 17 % sont des cadres.

Quand un citant dit 3 et l'autre dit 4, on a donc 34 % de chances que ce soit 3 et 30 % que ce soit 4. Si on admet que, vu les deux réponses, la bonne réponse est forcément entre 3 et 4, on tombe à 51 % de chance que la bonne réponse soit 3 versus 49 % qu'elle soit 4.

Conclusion : PCS=3 ou 4 → Correction à Pcs=3

8 corrections

Autres corrections :

PCS=2 ou 3 → Correction à PCS=3, parce que quand un citant dit 2, ça peut être 3 mais pas l'inverse. (4 corrections)

PCS=1 ou 5 → Les deux autres agriculteurs cités enquêtés sont des ouvriers. Donc Pcs=5 (1 correction)

PCS=2 ou 4 → Correction à PCS=4, parce que quand un citant dit 2, ça peut être 4 mais pas l'inverse. (1 correction)

PCS=2 ou 5 → Correction à PCS=5, parce que quand un citant dit 2, ça peut être 5 mais pas l'inverse. (1 correction)

PCS=3 ou 5 → Inconnu (99) (27 ou 29 % sur deux catégories très éloignées, 19 % sur professions intermédiaires...) (7 corrections)

PCS=3, 4, 5 → Inconnu (99) (1 correction)

PCS=5 ou 6 → Correction à PCS=5 (plus probable et catégorie plus nombreuse) 2 corrections

PCS=2 5 6 → Correction à PCS=5. Il existe beaucoup de 2 et de 6 qui sont en réalité des 5, l'inverse est rare. (1 correction)

PCS= 84 ou autres → Correction à PCS=Autres

PCS = 8 ou autres → Correction à PCS=Autres