Recodage des codes postaux.

*Ce fichier explique les grandes lignes du processus.*

*Un script (recodage\_codes\_postaux.R) contient les manips que j’ai faites pour arriver à ça.*

*Deux fichiers textes contiennent le code à insérer pour recoder les codes postaux d’origine en codes postaux recodés. Il faudra compléter ça lorsqu’on aura tout recodé.*

Voilà ce que j’ai fait pour les codes postaux. Le but est d’attribuer à chaque répondant la tranche d’unité urbaine (et éventuellement d’autres caractéristiques de la commune) de sa commune de résidence et de travail.

* Recodage manuel

J’ai recodé, dans la liste des codes postaux déclarés par les enquêté·es que Milan m’a transmise, tous les codes postaux qui pouvaient l’être. Certains ne seront pas recodables (pas déclarés ou déclaration illisible), d’autres pourront peut-être l’être – un code postal de résidence 6000 peut être 60000 (Beauvais) ou 06000 (Nice), en regardant quel est le lieu de travail, on peut inférer le code postal. Ca marchera pas pour tous les codes postaux inconnus, mais on pourra quand même en rattraper quelques-uns. Par ailleurs, d’autres codes postaux demandent un recodage manuel, voir plus bas.

* Appariement

Le script « recodage\_codes\_postaux » contient tout ce qu’il faut pour attribuer à chaque code postal valide déclaré dans l’enquête les informations sur la commune correspondante.

Il a fallu :

Regrouper tous les arrondissements de Paris, Lyon et Marseille, qui deviennent une seule unité (75000, 69000, 13000), parce que c’était beaucoup plus simple comme ça vu la composition des bases INSEE et que ça m’étonnerais qu’on fasse quelque chose de la distinction entre arrondissements. Mais si jamais on veut le faire, alors on essaiera de faire quelque chose.

Corriger le fait que les codes postaux des départements avant le 10 sont lus à 4 et non 5 chiffres (le premier 0 de la ville de Foix disparaît pour donner 9000).

**Recodages manuels à poursuivre (et à noter ici, idéalement).**

Cinq codes postaux sont inconnus de la base insee.

98800 correspond d’après google à Nouméa. Je propose qu’on le passe en « étranger » sans autre forme de procès.

76821 correspond à Ballinger au Texas (ville qui m’est totalement inconnue), mais le gars écrit en clair saint étienne du rouvray, à côté de rouen, donc on va mettre ça (je comprends pas comment la déclaration d’une commune peut générer un code postal qui n’existe pas…)

76420 correspond à Bihorel, une commune de la banlieue de Rouen. Je sais pas pourquoi notre base insee la connaît pas. Je suis pour mettre le code postal que je trouve sur google pour bois-guillaume (la ville déclarée par le répondant) : 76230.

20144 c’est en Corse (un bled à côté de porto vecchio qui s’appelle zonza).

A suivre…