

Informations sur l'échantillonnage

Projet FNS Ambizione „The neighborhood in the cloud”

Christoph Zangger

Situation de base

Le projet "The neighborhood in the cloud" (<https://data.snf.ch/grants/grant/201694>), soutenu par le Fonds national suisse, s'intéresse à la question de savoir comment se forment les réseaux ayant une référence spatiale, aussi bien en ligne que hors ligne, et comment ceux-ci influencent, avec d'autres contextes géographiques et sociaux, les chances individuelles de formation et d'emploi ainsi que la participation politique. Afin d'obtenir des résultats aussi robustes et généralisables que possible, le projet fait appel à différentes sources de données. Parmi celles-ci figure une enquête menée auprès d'environ 9000 personnes résidant en Suisse, qui seront interrogées à deux reprises. Les données existantes ne permettent pas d'identifier les voisinages à petite échelle, ce qui est nécessaire pour le projet, et elles ne présentent pas non plus le contenu correspondant, ce qui rend cette enquête supplémentaire nécessaire.

Comment les communes tirées ont-elles été sélectionnées ?

Le nombre visé de 9'000 personnes interrogées (ou contactées) doit permettre de tirer des conclusions aussi larges que possible et généralisables à l'ensemble de la Suisse. Pour cela, il est essentiel que ces personnes soient sélectionnées de manière aléatoire. Comme les relations sociales à petite échelle entre voisin·e·s directs sont d'un intérêt central, une simple sélection aléatoire des personnes n'est pas envisageable. Au lieu de cela, 30 coordonnées/adresses sont tirées au hasard et les 30 ménages les plus proches géographiquement sont invités à participer. Une telle procédure n'est cependant pas réalisable avec le cadre d'échantillonnage de l'Office fédéral

de la statistique et les données de contact correspondantes doivent être demandées directement aux offices de la population.

Afin de réduire le nombre d'offices à contacter et de garantir en même temps la représentativité de l'échantillon tiré, 86 communes ont été sélectionnées au total, dont 20 sont des localités à caractère urbain comptant au moins 20'000 habitant·e·s, 20 autres sont des localités à caractère urbain comptant moins de 20'000 habitant·e·s, 25 sont des communes rurales et 21 communes présentent un caractère "intermédiaire" selon la classification de l'Office fédéral de la statistique. La probabilité de tirage au sein de la typologie spatiale a été effectuée proportionnellement à la taille des communes. La figure 1 montre les 86 communes sélectionnées de manière aléatoire.

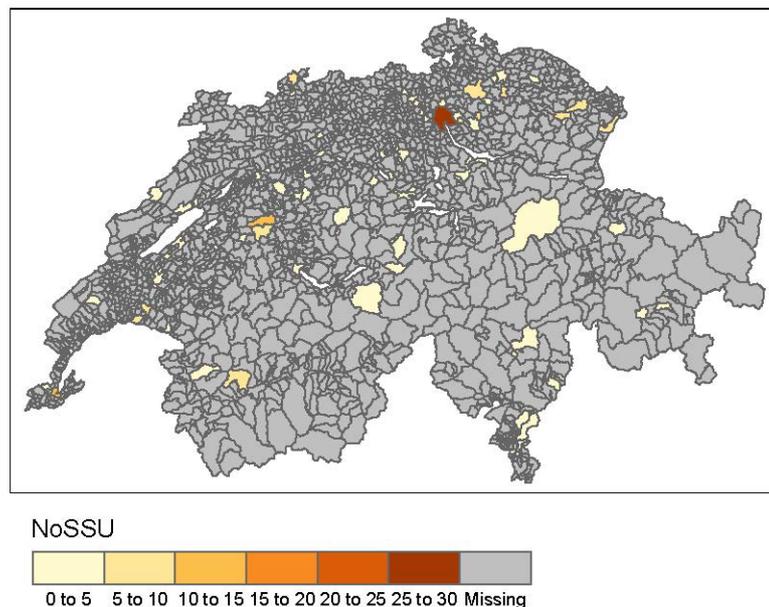


Figure 1 : Communes sélectionnées au hasard & nombre de coordonnées à tirer au sein de la commune

Au sein des 86 communes, le nombre de 300 coordonnées/adresses à tirer a été déterminé dans un deuxième temps. Pour ce faire, on a utilisé une sélection aléatoire avec remise, là encore avec des probabilités de tirage proportionnelles à la taille de la localité. On a veillé à ce que chaque commune soit tirée au moins une fois. La coloration de la figure 1 indique le nombre de coordonnées/adresses à tirer.

Comment les différentes personnes interrogées ont-elles été sélectionnées ?

Après avoir déterminé la commune à contacter et le nombre de personnes à y interroger (nombre de coordonnées à tirer x 30), l'étape suivante a consisté à choisir au hasard, au sein des communes,

les coordonnées (points cibles de l'échantillon) autour desquelles les 30 ménages suivants seront tirés. Pour tenir compte de la densité de population dans les communes concernées lors du tirage des points cibles, on a utilisé à cet effet le registre officiel des bâtiments. Seules les adresses des bâtiments à usage exclusif d'habitation (GKAT 1020) et celles des bâtiments à usage annexe (GKAT 1030) sont prises en compte. Les bâtiments à usage partiel d'habitation (p. ex. foyers ou hôpitaux ; GKAT 1040) ont toutefois été exclus, de même que toutes les adresses de bâtiments non résidentiels (GKAT 1060), de bâtiments spéciaux (GKAT 1080) et de logements provisoires (p. ex. conteneurs de chantier ; GKAT 1010).

Un simple échantillonnage aléatoire de cette base de données a pour conséquence que les adresses situées dans des zones à forte densité de population (nombreux bâtiments sur la même route) sont plus susceptibles d'être tirées que, par exemple, un chalet d'alpage isolé. Pour chacune des communes tirées lors de la première étape, le nombre correspondant de points cibles a ensuite été tiré. Autour de ceux-ci, 29 autres adresses ont été progressivement déterminées en fonction de leur distance euclidienne par rapport à l'adresse cible tirée (procédure automatisée par l'utilisation des coordonnées des bâtiments respectifs).

Ces groupes de 30 adresses ont ensuite été envoyés aux offices des habitant·e·s des communes concernées, en leur demandant de les remplir progressivement avec 30 ménages. Il arrive souvent que seules quelques-unes des 30 adresses cibles de chaque groupe soient nécessaires, car il est possible, par exemple, que la cible tirée soit un immeuble/bloc d'habitation comptant plusieurs ménages. Si le nombre de ménages à une adresse donnée dépasse le nombre de 30 unités, celles-ci ont été sélectionnées de manière aléatoire. Les données de contact correspondantes ont ensuite été mises à notre disposition par les offices des habitant·e·s pour la prise de contact.

Comment les données de contact sont-elles utilisées et la protection des données est-elle garantie ?

Les données de contact ainsi que les données d'enquête collectées sont stockées sur un lecteur de serveur physique de l'Université de Berne. Seules les personnes employées par l'Université de Berne et directement impliquées dans la collecte des données (5 personnes) ont accès à ce lecteur. Les données de contact sont en outre sauvegardées sur ce lecteur de serveur dans un lecteur virtuel crypté, de sorte que même en cas de vol physique du serveur, les données restent protégées. Enfin, les données de contact ne sont conservées que le temps de l'enquête (2 mois) et sont ensuite détruites sans possibilité de récupération.

Pour toute autre question, veuillez-vous adresser à christoph.zangger@unibe.ch