

# **{sampler} : une grammaire tidy pour exprimer les plans de sondage en R**

Ahmadou Dicko<sup>1</sup>

Dans les rapports méthodologiques et les supports académiques, les plans de sondage sont décrits de manière structurée avec une logique explicite entre strates, grappes, degrés de tirage et probabilités de sélection. Cette structure se perd souvent dans le code R, où le design se disperse entre objets intermédiaires, appels successifs et détails techniques. Plus le plan est complexe, plus l'écart se creuse, compliquant la relecture, la maintenance et la vérification.

{sampler} a été conçu pour réduire cet écart. Le package propose une grammaire déclarative inspirée de l'écosystème tidy, où chaque verbe correspond à une opération du plan d'échantillonnage: `stratify_by()`, `cluster_by()`, `draw()`, `execute()`. Des plans d'échantillonnage complexes y sont codés de façon plus structurée, plus lisible et plus proche du raisonnement statistique.

La présentation illustrera cette approche à partir de plusieurs types de plans usuels en statistique publique et en enquêtes appliquées. Nous montrerons comment une écriture plus explicite du design peut améliorer la reproductibilité, limiter les erreurs d'implémentation et faciliter la transition entre conception, tirage et analyse. Le but est de montrer comment {sampler} peut contribuer à rendre l'échantillonnage plus accessible sans sacrifier la rigueur méthodologique.

**Mots-clefs** : Sondages - Échantillonnage - Tidyverse - Enseignement - Reproductibilité

## **Lien**

<https://dickoa.gitlab.io/sampler/>

<sup>1</sup>mail@ahmadoudicko.com