RX-04

Nous avons compté les associations d'un territoire dans un fichier qui n'en voit qu'un quart

Retour d'expérience. Un chiffre publié, faux d'un facteur quatre, et rien dans la donnée ne le signalait.

Famille	*Généricité*	*Statut*
Retours d'expérience	GÉNÉRIQUE	EXPLORÉ — production du corpus, non validée par une décision
Écrit le	*Licence*	
15 septembre 2026	CC BY-SA 4.0	

L'attendu

Dénombrer les associations actives dans un rayon de 45 km, pour pouvoir dire ce que contient le milieu associatif qu'on prétend soutenir. Source retenue : l'**annuaire des entreprises et associations** (`recherche-entreprises.api.gouv.fr`), en filtrant sur le code d'activité 94.99Z, « autres organisations fonctionnant par adhésion volontaire ». API publique, à jour, requête géographique native, CORS ouvert. Le choix allait de soi.

Résultat publié : **1 952 associations actives**, soit une pour 122 habitants, avec le commentaire « le tissu associatif est dense ».

Le déroulé

En cherchant à obtenir l'**objet déclaré** de chaque association — que cet annuaire ne fournit pas — il a fallu descendre jusqu'au Répertoire national des associations, tenu par le ministère de l'Intérieur. Le même périmètre, les mêmes 284 communes, le même filtre « active » y donne :

8 340 associations.

	annuaire des entreprises	Répertoire national des associations
Associations comptées dans la zone	1 952	**8 340**
Une association pour	122 habitants	**29 habitants**
Objet déclaré disponible	non	oui, renseigné à 100 %

L'écart n'est pas une erreur de requête. Il est structurel, et il a un nom : **une association n'a de numéro SIREN que si elle emploie, perçoit des subventions publiques, ou exerce une activité assujettie.** Dans le dump du RNA pour la Corrèze, **197 lignes sur 9 434 portent un SIRET — 2 %**. L'annuaire des entreprises ne voit que la minorité immatriculée.

Autrement dit : nous avions publié le nombre d'associations **qui ont une existence financière**, en l'appelant le nombre d'associations. Ce n'est pas la même population, et la différence est exactement celle qui compte pour un projet rural — les associations sans salarié, sans subvention, sans comptabilité, qui sont l'immense majorité du tissu qu'on voulait décrire.

Pourquoi rien ne l'a signalé

La requête était juste. Le filtre était le bon, la géolocalisation fonctionnait, le dédoublonnage était fait, les établissements fermés étaient exclus, et le siège hors zone était correctement remplacé par l'établissement local. Techniquement, le comptage était propre.

Le résultat était plausible. Une association pour 122 habitants est un chiffre crédible — c'est à peu près la moyenne nationale telle qu'elle circule. Rien n'appelait la vérification. Un chiffre faux d'un facteur quatre a passé le contrôle de vraisemblance parce qu'il ressemblait à un chiffre connu.

Le champ manquant n'était pas vide, il était absent. Si l'annuaire avait renvoyé une liste avec des trous, on l'aurait vu. Il renvoie une liste complète — complète de ce qu'il contient. Une source ne dit jamais ce qu'elle ne contient pas.

C'est la recherche d'un autre champ qui a révélé le problème. Nous ne cherchions pas à vérifier le total ; nous cherchions l'objet déclaré. La correction est un effet de bord. Sans ce besoin, le chiffre serait encore en ligne.

Les leçons

1. **Un registre a un critère d'entrée, et ce critère est une définition cachée.** SIRENE n'immatricule pas les associations : il immatricule celles qui font quelque chose d'économiquement visible. Compter dans SIRENE, c'est donc compter *cette* population-là. La question à poser à toute source de dénombrement n'est pas « est-elle à jour ? » mais **« qu'est-ce qui fait entrer une ligne dans ce fichier ? »** — et cette question se pose avant la première requête.

2. **Pour un dénombrement, chercher le registre dont l'objet *est* de dénombrer.** Le RNA existe pour recenser les associations ; SIRENE existe pour immatriculer des unités économiques. Les deux contiennent des associations ; un seul les compte. À chaque fois, préférer le registre dont la population cible est exactement la population cherchée, même s'il est plus pénible d'accès — ici, 408 Mo à télécharger contre une API avec CORS.

3. **Croiser deux sources sur le même comptage, avant de publier.** Le contrôle qui aurait tout révélé tient en une comparaison de deux totaux, et prend une heure. Il n'a pas été fait parce que la première source semblait suffire. **Un chiffre unique n'est pas un chiffre vérifié**, aussi propre soit sa provenance.

4. **Le rapport entre les deux sources est lui-même une information.** 1 952 sur 8 340, soit 23 %, est une mesure de la part du tissu associatif qui a une existence financière. Elle vaut d'être publiée, et elle n'existait qu'une fois l'erreur trouvée. Une erreur corrigée laisse souvent un indicateur derrière elle.

Les changements

- Tout dénombrement publié nomme, dans la méthode, **le critère d'entrée du registre utilisé** — pas seulement sa source et sa date.
- Tout dénombrement est **confronté à une seconde source** de population avant publication, et l'écart est publié quand il est notable.
- Quand un registre dédié existe pour la population cherchée, il l'emporte sur une API plus commode portant une population voisine.

- Les chiffres déjà publiés sur le site sont repassés à ce filtre, par ordre d'exposition.

Portée

Générique, et fréquent. Le même piège guette : compter les exploitations agricoles dans le registre parcellaire graphique (il ne voit que les déclarants PAC), les entreprises dans les greffes (pas les auto-entrepreneurs), les logements dans la taxe d'habitation, les habitants dans les listes électorales, les cours d'eau dans les stations hydrométriques (voir `tdb-eau` : 43 km² mesurés, le chevelu invisible). À chaque fois, la source est officielle, la requête est juste, et le résultat décrit une population plus étroite que celle qu'on nomme.

Un registre ne se trompe jamais sur ce qu'il contient. Il se tait sur ce qu'il ne contient pas.

Fiches liées

RX-01 · RX-02 · RX-03 · PA-02

CoMons, Collectif du Mons — RX-04, Nous avons compté les associations d'un territoire dans un fichier qui n'en voit qu'un quart. Fiche du Grain, CC BY-SA 4.0. Version en ligne : comons.actitude.org/grain-rx-04-source-hors-de-son-domaine.html