

Quatre cinquièmes améliorés, la moitié servis

82,3 % de ces ménages sont sur une source améliorée et 56,3 % disposent au moins d'un service élémentaire d'eau de boisson. Les vingt-six points d'écart sont du temps de collecte, et ils sont toute la raison d'être des échelons.

Pierre Robentz Cassion

Leçon 1 sur 8

Analyse EAH — L'échelle n'est pas la source

Ce que couvre cette leçon

- Le nombre que donnent la plupart des rapports EAH
- L'échelle que l'indicateur ODD emploie réellement
- Pourquoi les trente minutes figurent dans la définition
- L'échelon que cette enquête ne peut pas calculer
- Rapportez la distribution, non l'échelon supérieur
- La suite

Le nombre que donnent la plupart des rapports EAH — En Python

```
import pandas as pd

households = pd.read_csv("wash-household-survey-2024.v1.csv")

IMPROVED = {                                # the JMP list, delivered water included
    "piped-into-dwelling", "piped-into-yard", "public-tap",
    "borehole", "protected-well", "protected-spring", "tanker-truck",
}

improved = households["water_source"].isin(IMPROVED)
print(f"improved source: {improved.mean():.1%} of {len(households)} households")
```

Le nombre que donnent la plupart des rapports EAH — En R

```
library(dplyr)

improved <- c("piped-into-dwelling", "piped-into-yard", "public-tap",
             "borehole", "protected-well", "protected-spring",
             "tanker-truck")

households |> summarise(improved = mean(water_source %in% improved), n = n())
```

Le nombre que donnent la plupart des rapports EAH

- 82,3 % — C'est le calcul exact d'une chose que personne n'a demandée

L'échelle que l'indicateur ODD emploie réellement

Échelon	Définition
Géré en toute sécurité	Source améliorée sur place, disponible en cas de besoin, et exempte de contamination
Élémentaire	Source améliorée, aller-retour de 30 minutes ou moins, attente comprise
Limité	Source améliorée, aller-retour de plus de 30 minutes
Non amélioré	Puits ou source non protégés
Eau de surface	Rivière, retenue, lac, mare, canal, canal d'irrigation

L'échelle que l'indicateur ODD emploie réellement — En Python

```
def ladder(row):  
    if row["water_source"] == "surface-water":  
        return "surface water"  
    if row["water_source"] not in IMPROVED:  
        return "unimproved"  
    minutes = row["round_trip_minutes"]  
    return "basic" if minutes <= 30 else "limited"  
  
households["service"] = households.apply(ladder, axis=1)  
print(households["service"].value_counts(normalize=True).round(3))
```

L'échelle que l'indicateur ODD emploie réellement — En R

```
households |>
  mutate(service = case_when(
    water_source == "surface-water" ~ "surface water",
    !water_source %in% improved ~ "unimproved",
    round_trip_minutes <= 30 ~ "basic",
    TRUE ~ "limited"
  )) |>
  count(service) |>
  mutate(share = n / sum(n))
```

L'échelle que l'indicateur ODD emploie réellement

Échelon	Ménages	Part
Élémentaire (sur place compris)	1 352	56,3 %
Limité	626	26,1 %
Non amélioré	326	13,6 %
Eau de surface	99	4,1 %

L'échelle que l'indicateur ODD emploie réellement

- Service d'eau de boisson au moins élémentaire : 56,3 % — Le chiffre par type de source valait 82,3 %, et chacun des...

Pourquoi les trente minutes figurent dans la définition

- les ménages collectent **moins** d'eau, si bien que la quantité passe plus souvent sous le minimum Sphère ;
- la collecte est faite par les femmes et les filles, si bien que le temps est pris sur la scolarité et le travail. . .
- les ménages complètent depuis une source plus proche et de moindre qualité, si bien que la source enregistrée n'est pas. . .
- Une définition qui paraît bureaucratique encode généralement un constat — Avant de soutenir qu'un seuil est arbitraire, . . .

L'échelon que cette enquête ne peut pas calculer — En Python

```
on_premises = (households["round_trip_minutes"] == 0)
tested = households["ecoli_cfu_100ml"].notna()

print(f"on premises: {on_premises.mean():.1%}")
print(f"water quality tested: {tested.mean():.1%}")
print(f"both: {(on_premises & tested).mean():.1%}")
```

L'échelon que cette enquête ne peut pas calculer — En R

```
households |> summarise(  
  on_premises = mean(round_trip_minutes == 0),  
  tested = mean(!is.na(ecoli_cfu_100ml))  
)
```

Rapportez la distribution, non l'échelon supérieur — Exemple

Drinking water service level, 2,403 households

At least basic	56.3%	1,352
of which on premises	27.7%	665
Limited	26.1%	626
Unimproved	13.6%	326
Surface water	4.1%	99

JMP ladder. Safely managed not computable: availability was not asked and water quality was tested on 33.4% of households.

- L'eau a une échelle et il en reste deux.

Lire la leçon complète, avec le code exécutable

```
https://data-analysis.cassion.dev/fr/cours/wash-analysis/  
wash-01-the-drinking-water-ladder
```