

# Le débat sur le dénominateur, et les découpages que vous promettez

Trois dénominateurs pour un numérateur, la règle qui tranche entre eux, et la désagrégation qui fait passer de vingt-quatre cellules à quarante-huit, dont trois trop petites pour être publiées.

---

Pierre Robentz Cassion

Leçon 4 sur 8

Conception d'indicateurs et cadre logique — La fiche de référence

## Ce que couvre cette leçon

- Toute dispute sur un chiffre est une dispute sur le dénominateur
- Trois dénominateurs, un registre
- La règle
- Dites ce qu'il exclut, dans la fiche
- La désagrégation est une promesse sur la taille d'échantillon
- Fixez un effectif minimal, dans la fiche, à l'avance
- Quels découpages changent réellement une décision
- La suite

## Toute dispute sur un chiffre est une dispute sur le dénominateur

- Le numérateur se règle d'ordinaire en une minute.

## Trois dénominateurs, un registre — En Python

```
measured = muac["muac_mm"].notna()
assessed = measured | muac["oedema"].notna()
cases = (muac["muac_mm"] < 125) | (muac["oedema"] == True)

for label, mask in [("all screened", pd.Series(True, index=muac.index)),
                    ("with a measurement", measured),
                    ("assessed either way", assessed)]:
    print(f"{label:22} n={mask.sum():5}  GAM={cases.sum() / mask.sum():.2%}")
```

## Trois dénominateurs, un registre — En R

```
muac |>
  summarise(
    all_screened = n(),
    measured     = sum(!is.na(muac_mm)),
    assessed     = sum(!is.na(muac_mm) | !is.na(oedema)),
    cases       = sum(muac_mm < 125 | oedema, na.rm = TRUE)
  ) |>
  mutate(across(c(all_screened, measured, assessed), ~ cases / .x, .names = "gam_{.col}"))
```

## Trois dénominateurs, un registre

| Dénominateur                        | n     | MAG    |
|-------------------------------------|-------|--------|
| Tous ceux qui sont venus            | 4 218 | 8,58 % |
| Enfants avec une mesure de PB       | 4 146 | 8,73 % |
| Enfants évalués par mesure ou œdème | 4 216 | 8,59 % |

Le dénominateur est la population qui était réellement susceptible de figurer au numérateur, sur la même période, au même endroit.

- **Couverture.** Pas tout le district — la tranche d'âge cible du bassin, sur la période.
- **Taux de guérison.** Pas tous les admis — ceux ayant atteint une issue. Les enfants encore en traitement à la date. . .
- **Aboutissement des référencements.** Pas tous les cas — les cas ayant consenti à être référés, ce que le jeu de. . .
- **Assiduité.** Pas les enfants inscrits multipliés par tous les jours de classe — les enfants inscrits multipliés par. . .

## Dites ce qu'il exclut, dans la fiche — Exemple

|             |                                                                                                                                                                                                              |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Denominator | Children with a MUAC measurement recorded (n = 4,146 of 4,218).                                                                                                                                              |
| Excludes    | 72 children (1.7%) screened but not measured, coded -99. Their nutritional status is unknown; if they were systematically the most distressed children, GAM is understated. Not testable from this register. |

## La désagrégation est une promesse sur la taille d'échantillon — En Python

```
banded = muac[muac["age_months"].notna()].assign(
    band=lambda d: (d["age_months"] >= 24).map({True: "24-59", False: "6-23"})
)

by_two = banded.groupby(["commune", "band"]).size()
by_three = banded.groupby(["commune", "band", "sex"]).size()

print(f"commune x band:          {len(by_two)} cells, smallest {by_two.min()}")
print(f"commune x band x sex: {len(by_three)} cells, smallest {by_three.min()}, "
      f"{(by_three < 30).sum()} below 30")
```

## La désagrégation est une promesse sur la taille d'échantillon — En R

```
banded <- muac |>
  filter(!is.na(age_months)) |>
  mutate(band = if_else(age_months >= 24, "24-59", "6-23"))

banded |> count(commune, band) |> summarise(cells = n(), smallest = min(n))
banded |> count(commune, band, sex) |> summarise(cells = n(), smallest = min(n),
  under_30 = sum(n < 30))
```

## La désagrégation est une promesse sur la taille d'échantillon

| Désagrégation                  | Cellules | Plus petite | Cellules sous 30 |
|--------------------------------|----------|-------------|------------------|
| Commune × tranche d'âge        | 24       | 49          | 0                |
| Commune × tranche d'âge × sexe | 48       | 16          | 3                |

## La désagrégation est une promesse sur la taille d'échantillon

- Ajouter un seul découpage binaire double les cellules et divise leur taille par deux
  - Passer de deux à trois. . .

## Fixez un effectif minimal, dans la fiche, à l'avance — En Python

```
MIN_CELL = 30

table = (
    banded.assign(case=cases)
    .groupby(["commune", "band", "sex"])
    .agg(n=("case", "size"), cases=("case", "sum"))
)

table["rate"] = (table["cases"] / table["n"]).where(table["n"] >= MIN_CELL)
table["note"] = table["n"].lt(MIN_CELL).map({True: "suppressed: n < 30", False: ""})
```

## Fixez un effectif minimal, dans la fiche, à l'avance — En R

```
MIN_CELL <- 30

table <- banded |>
  summarise(n = n(), cases = sum(case), .by = c(commune, band, sex)) |>
  mutate(rate = if_else(n >= MIN_CELL, cases / n, NA_real_),
         note = if_else(n < MIN_CELL, "suppressed: n < 30", ""))
```

## Quels découpages changent réellement une décision

- **Sexe** — oui, presque toujours. Cela change le ciblage, les effectifs et les messages.
- **Tranche d'âge** — oui en nutrition, où les critères d'admission diffèrent selon l'âge.
- **District** — oui, cela déplace des ressources.
- **Type de formation** — parfois ; cela change la supervision, ce qui est réel.
- **Mois** — d'ordinaire une tendance et non une désagrégation, et moins chère en graphique.

- Vous savez désormais définir un indicateur avec précision.

Lire la leçon complète, avec le code exécutable

[https://data-analysis.cassion.dev/fr/cours/indicator-design-logframe/  
indicators-04-denominators-and-disaggregation](https://data-analysis.cassion.dev/fr/cours/indicator-design-logframe/indicators-04-denominators-and-disaggregation)