

109 % scolarisés, 97 % scolarisés

Deux taux issus d'un même registre. Le brut vaut 109,2 % et le net 97,4 %, les deux sont exacts, et les douze points d'écart sont des enfants scolarisés au mauvais âge plutôt que des enfants absents.

Pierre Robentz Cassion

Leçon 1 sur 8

Analyse de programmes d'éducation — L'inscription et son dénominateur

Ce que couvre cette leçon

- Deux taux, un registre
- Pourquoi un taux peut dépasser 100 %
- Le dénominateur est une projection
- Lequel rapporter
- Ce que les taux ne peuvent pas dire
- La suite

Deux taux, un registre — En Python (suite)

```
import pandas as pd

enrolment = pd.read_csv("school-enrolment-2024.v1.csv")
population = pd.read_csv("school-age-population-2024.v1.csv")

current = enrolment[enrolment["school_year"] == 2024]
primary = current[current["grade"].between(1, 6)]

denominator = population.loc[
    (population["reference_year"] == 2024)
    & population["age_years"].between(6, 11),
    "projected_population",
].sum()

gross = len(primary) / denominator
net = len(primary[primary["age_years"].between(6, 11)]) / denominator
```

Deux taux, un registre — En Python (suite)

```
print(f"gross enrolment ratio: {gross:.1%}")  
print(f"net enrolment ratio:   {net:.1%}")
```

Deux taux, un registre — En R

```
library(dplyr)

primary <- enrolment |> filter(school_year == 2024, between(grade, 1, 6))
denominator <- population |>
  filter(reference_year == 2024, between(age_years, 6, 11)) |>
  summarise(n = sum(projected_population)) |> pull(n)

tibble(
  gross = nrow(primary) / denominator,
  net = sum(between(primary$age_years, 6, 11)) / denominator
)
```

Deux taux, un registre

Taux	Numérateur	Dénominateur	Valeur
Brut	Tous les inscrits au primaire, tout âge	Enfants de 6 à 11 ans	109,2 %
Net	Inscrits au primaire de 6 à 11 ans	Enfants de 6 à 11 ans	97,4 %

Deux taux, un registre

- Les numérateurs diffèrent ; le dénominateur est identique — C'est toute la définition, et c'est pourquoi les deux taux. . .

Pourquoi un taux peut dépasser 100 %

- Le retard et l'avance d'âge — Les enfants hors de la tranche d'âge officielle sont au numérateur et pas au dénominateur
- Un dénominateur trop petit — Le fichier de population est une projection et non un dénombrement, et une projection qui. . .
- Le double comptage — Un enfant inscrit dans deux écoles figure deux fois

Pourquoi un taux peut dépasser 100 % — En Python

```
duplicates = current.duplicated(subset=["student_id", "school_year"], keep=False)
print(f"student-years appearing more than once: {duplicates.sum()}")

unique = current.drop_duplicates(subset=["student_id", "school_year"])
unique_primary = unique[unique["grade"].between(1, 6)]
print(f"gross ratio after deduplication: {len(unique_primary) / denominator:.1%}")
```

Pourquoi un taux peut dépasser 100 % — En R

```
enrolment |> filter(school_year == 2024) |>  
  count(student_id) |> filter(n > 1) |> nrow()
```

Le dénominateur est une projection — En Python

```
CENSUS_YEAR, GROWTH = 2015, 0.021
factor = (1 + GROWTH) ** (2024 - CENSUS_YEAR)
print(f"nine years compounded at {GROWTH:.1%}: factor {factor:.3f}")
print(f"so about {1 - 1 / factor:.0%} of the denominator is an assumption")
```

Le dénominateur est une projection — En R

```
(1 + 0.021)^(2024 - 2015)
```

Le dénominateur est une projection

- Un facteur de 1,21, donc environ un sixième du dénominateur n'a jamais été compté — C'est le dénominateur vaccinal du...

Le dénominateur est une projection — En Python

```
for error in (-0.05, 0.0, 0.05):
    adjusted = denominator * (1 + error)
    print(f"projection {error:+.0%}: gross {len(primary) / adjusted:.1%}, "
          f"net {len(primary[primary['age_years'].between(6, 11)]) / adjusted:.1%}")
```

Le dénominateur est une projection — En R

Vary the denominator and see which conclusions survive.

Le dénominateur est une projection

- Montrez la sensibilité plutôt que l'estimation ponctuelle — là où le dénominateur est projeté

- Le taux net répond à « les enfants d'âge scolaire sont-ils scolarisés » — C'est le cadrage de l'ODD 4.1 et le bon taux. . .
- Le taux brut répond à « quelle quantité de scolarité primaire est délivrée » — C'est le bon taux pour une question de. . .
- L'écart entre les deux est le constat — et c'est la raison de rapporter les deux : douze points de retard d'âge sont un. . .

Lequel rapporter — Exemple

Primary enrolment, 2024, four districts

Gross enrolment ratio	109.2%	1,052 enrolees / 963 children aged 6-11
Net enrolment ratio	97.4%	938 in-age enrolees / same denominator
Over-age share of enrolment	36.4%	above the official age for their grade

Denominator is a projection from the 2015 census at 2.1% a year. A 5% error in it moves the gross ratio by about five points and changes no conclusion.

12 student-years were recorded twice after transfers and are deduplicated.

Ce que les taux ne peuvent pas dire

- Aucun n'est la présence — Un enfant inscrit et jamais présent figure aux deux numérateurs
- Aucun n'est l'achèvement — L'inscription est un stock à un moment de l'année ; savoir si ces enfants terminent est une. . .
- Aucun n'est l'apprentissage — Un système peut scolariser chaque enfant d'âge scolaire et n'en instruire aucun, et la. . .

- Douze points de l'écart entre les deux taux sont du retard d'âge.

Lire la leçon complète, avec le code exécutable

[https://data-analysis.cassion.dev/fr/cours/education-programme-analysis/
edu-01-gross-and-net-enrolment](https://data-analysis.cassion.dev/fr/cours/education-programme-analysis/edu-01-gross-and-net-enrolment)