109% enrolled, 97% enrolled

Two ratios from one register. The gross is 109.2% and the net is 97.4%, both are correct, and the twelve points between them are children who are in school at the wrong age rather than children who are missing.

Pierre Robentz Cassion

Lesson 1 of 8

Education Programme Analysis — Enrolment and its denominator

What this lesson covers

- Two ratios, one register
- Why a ratio can exceed 100%
- The denominator is a projection
- Which one to report
- What the ratios cannot tell you
- What comes next

Two ratios, one register — In Python (cont.)

```python
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
```

Two ratios, one register — In Python (cont.)

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

Two ratios, one register — In R

```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
)
```

Two ratios, one register

Ratio	Numerator	Denominator	Value
Gross	All primary enrolees, any age	Children aged 6–11	**109.2%**
Net	Primary enrolees aged 6–11	Children aged 6–11	**97.4%**

Two ratios, one register

- The numerators differ; the denominator is identical — That is the whole definition, and it is why the two ratios are...

Why a ratio can exceed 100%

- Over-age and under-age enrolment — Children outside the official age range are in the numerator and not in the...
- A denominator that is too small — The population file is a projection, not a count, and a projection that undershoots...
- Double counting — A child enrolled at two schools appears twice

Why a ratio can exceed 100% — In 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%}")
```

Why a ratio can exceed 100% — In R

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

The denominator is a projection — In 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")
```

The denominator is a projection — In R

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

The denominator is a projection

- A factor of 1.21, so roughly a sixth of the denominator was never counted — This is the immunisation denominator from...

The denominator is a projection — In 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%}")
```

The denominator is a projection — In R

```
# Vary the denominator and see which conclusions survive.
```

The denominator is a projection

- Show the sensitivity rather than the point estimate — where the denominator is projected

Which one to report

- Net enrolment answers "are children of school age in school" — It is the SDG 4.1 framing and the right ratio for a...
- Gross enrolment answers "how much primary schooling is being delivered" — It is the right ratio for a capacity question...
- The gap between them is the finding — and it is the reason to report both: twelve points of over-age enrolment is a...

Which one to report — Example

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.

What the ratios cannot tell you

- Neither is attendance — A child enrolled and never present is in both numerators
- Neither is completion — Enrolment is a stock at a point in the year; whether those children finish is a cohort question...
- Neither is learning — A system can enrol every child of school age and teach none of them, and the last unit is the...

What comes next

- Twelve points of the gap between the two ratios is over-age enrolment.

Read the full lesson, with runnable code

https://data-analysis.cassion.dev/courses/education-programme-analysis/edu-01-gross-and-net-enrolment