

Reach, coverage, and the people counted twelve times

Three indicators that all sound like "how many people did we help", and the monthly series that turns 4,200 children into 50,000 the moment somebody sums the column.

Pierre Robentz Cassion

Lesson 5 of 8

Indicator Design and the LogFrame — Counting people

What this lesson covers

- Three questions that sound like one
- The cumulative sum that manufactures people
- Deduplicating across months, when you can
- Reach is not coverage, and the difference is the denominator
- Cumulative counts in proposals
- Which one belongs in your LogFrame
- What comes next

Three questions that sound like one

- **Reach** — how many *distinct people* received something, over a stated period.
- **Service volume** — how many *services* were delivered. Larger than reach whenever anyone comes twice.
- **Coverage** — what *share* of those who needed it got it. Requires a denominator and is the only one that says whether. . .

The cumulative sum that manufactures people — In Python

```
monthly = (  
    muac.assign(month=muac["screening_date"].dt.strftime("%Y-%m"))  
    .groupby("month")  
    .agg(screenings=("child_id", "size"),  
         distinct_children=("child_id", "nunique"))  
)  
monthly["cumulative_naive"] = monthly["screenings"].cumsum()  
  
print(monthly)  
print("sum of monthly counts: ", monthly["screenings"].sum())  
print("distinct children, year:", muac["child_id"].nunique())
```

The cumulative sum that manufactures people — In R

```
monthly <- muac |>
  mutate(month = format(screening_date, "%Y-%m")) |>
  summarise(screenings = n(), distinct_children = n_distinct(child_id), .by = month) |>
  arrange(month) |>
  mutate(cumulative_naive = cumsum(screenings))

c(sum_of_months = sum(monthly$screenings),
  distinct_year = n_distinct(muac$child_id))
```

The cumulative sum that manufactures people

	Correct	Wrong
Monthly reach, January	distinct people in January	—
Annual reach	distinct people across the year	sum of twelve monthly counts
Cumulative reach to date	distinct people since programme start	running total of monthly counts

The cumulative sum that manufactures people

- A cumulative reach figure can only be computed from person-level data — If your reporting is monthly aggregates, you...

The cumulative sum that manufactures people — In Python

```
annual_reach = muac["child_id"].nunique()  
service_volume = len(muac)  
print(f"reach {annual_reach:,} children; {service_volume:,} screening events")
```

The cumulative sum that manufactures people — In R

```
c(reach = n_distinct(muac$child_id), volume = nrow(muac))
```

Deduplicating across months, when you can — In Python

```
first_contact = (  
    muac.sort_values("screening_date")  
    .drop_duplicates(subset=["child_id"], keep="first")  
    .assign(month=lambda d: d["screening_date"].dt.strftime("%Y-%m"))  
)  
new_by_month = first_contact.groupby("month").size()  
print(new_by_month.cumsum())
```

Deduplicating across months, when you can — In R

```
first_contact <- muac |>
  arrange(screening_date) |>
  distinct(child_id, .keep_all = TRUE) |>
  mutate(month = format(screening_date, "%Y-%m"))

first_contact |> count(month) |> arrange(month) |> mutate(cumulative = cumsum(n))
```

Reach is not coverage, and the difference is the denominator — In Python

```
reached = muac["child_id"].nunique()
under5_population = 21500          # from the administrative frame, projected

print(f"screening coverage: {reached / under5_population:.1%}")
```

Reach is not coverage, and the difference is the denominator — In R

```
c(coverage = n_distinct(muac$child_id) / 21500)
```

Reach is not coverage, and the difference is the denominator

- Admissions over expected caseload is not coverage — Dividing the children admitted to treatment by the caseload you...
- A reach figure with a population denominator is not coverage either — unless the population is the population *in need*

Cumulative counts in proposals

- **State the deduplication level.** "Unique individuals, deduplicated within year, not across years" is a defensible and. . .
- **Never sum reach across sectors.** A household receiving water, food and a protection service is one household, and. . .
- **Say when you cannot.** "Aggregate figures are not deduplicated across partners; the true unique reach is lower" is a. . .

Which one belongs in your LogFrame

If the row is about	Use
Programme effort and cost	Service volume
People served	Reach, with the deduplication level stated
Whether the response is sufficient	Coverage, with a needs-based denominator
Whether people who start finish	A completion or dropout rate

Which one belongs in your LogFrame

Reach tells you what you did. Coverage tells you whether it was enough. A report with only the first is a description of activity, and every reader knows it.

What comes next

- Almost every indicator in this lesson already has an official definition published by somebody — UNICEF, WHO, the cluster, the SDG framework.

Read the full lesson, with runnable code

[https://data-analysis.cassion.dev/courses/indicator-design-logframe/
indicators-05-reach-and-coverage](https://data-analysis.cassion.dev/courses/indicator-design-logframe/indicators-05-reach-and-coverage)