

The boolean that halves a coverage figure

62.0% of households have an improved sanitation facility and 43.2% have basic service. The difference is one column — whether the latrine is shared — and the hygiene ladder has an equivalent that costs it half again.

Pierre Robentz Cassion

Lesson 2 of 8

WASH Analysis — The ladder is not the source

What this lesson covers

- Sanitation has the same shape as water
- Why sharing costs a rung
- Open defecation is the number to lead with
- The hygiene ladder, and the observation that carries it
- The three ladders cannot be averaged
- What comes next

Sanitation has the same shape as water

Rung	Definition
Safely managed	Improved, not shared, and excreta safely disposed of or treated
Basic	Improved, not shared with other households
Limited	Improved but shared
Unimproved	Pit latrine without a slab, hanging latrine, bucket
Open defecation	No facility

Sanitation has the same shape as water — In Python

```
IMPROVED_SANITATION = {  
    "flush-to-sewer", "flush-to-septic", "vip-latrine", "pit-latrine-with-slab",  
}  
  
improved = households["sanitation_facility"].isin(IMPROVED_SANITATION)  
shared = households["shared_sanitation"]  
  
print(f"improved facility: {improved.mean():.1%}")  
print(f"basic (not shared): {(improved & ~shared).mean():.1%}")  
print(f"limited (shared): {(improved & shared).mean():.1%}")
```

Sanitation has the same shape as water — In R

```
improved_sanitation <- c("flush-to-sewer", "flush-to-septic",  
                          "vip-latrine", "pit-latrine-with-slab")  
  
households |>  
  mutate(service = case_when(  
    sanitation_facility == "open-defecation" ~ "open defecation",  
    !sanitation_facility %in% improved_sanitation ~ "unimproved",  
    shared_sanitation ~ "limited",  
    TRUE ~ "basic"  
  )) |>  
  count(service) |>  
  mutate(share = n / sum(n))
```

Sanitation has the same shape as water

Rung	Households	Share
Basic	1,038	43.2%
Limited (shared)	452	18.8%
Unimproved	594	24.7%
Open defecation	319	13.3%

Sanitation has the same shape as water

- 62.0% have an improved facility and 43.2% have basic service — Eighteen points sit on a latrine that meets the...

Why sharing costs a rung

- This is the same argument as the thirty minutes — In both ladders the second condition exists because a facility that...

Open defecation is the number to lead with — In Python

```
od = households["sanitation_facility"].eq("open-defecation")
by_district = households.groupby("district")["sanitation_facility"].apply(
    lambda s: s.eq("open-defecation").mean()
)
print(f"{od.sum()} households, {od.mean():.1%}")
print((by_district * 100).round(1))
```

Open defecation is the number to lead with — In R

```
households |>
  summarise(open_defecation = mean(sanitation_facility == "open-defecation"),
            n = n(), .by = district)
```

The hygiene ladder, and the observation that carries it — In Python

```
has_facility = households["handwashing_facility"].ne("no-facility")
soap = households["soap_observed"]

print(f"any facility:                {has_facility.mean():.1%}")
print(f"basic (facility + soap):    {(has_facility & soap).mean():.1%}")
print(f"limited (no soap):          {(has_facility & ~soap).mean():.1%}")
print(f"no facility:                {(~has_facility).mean():.1%}")
```

The hygiene ladder, and the observation that carries it — In R

```
households |>
  mutate(hygiene = case_when(
    handwashing_facility == "no-facility" ~ "no facility",
    soap_observed ~ "basic",
    TRUE ~ "limited"
  )) |>
  count(hygiene) |>
  mutate(share = n / sum(n))
```

The hygiene ladder, and the observation that carries it

Rung	Households	Share
Basic — facility and soap	819	34.1%
Limited — facility, no soap	811	33.7%
No facility	773	32.2%

The hygiene ladder, and the observation that carries it

- 67.8% have somewhere to wash their hands and 34.1% have soap there — The hygiene ladder loses half its top rung to a . . .
- Observation is why this is trustworthy — Reported handwashing runs far above observed handwashing everywhere it has. . .

The three ladders cannot be averaged — In Python

```
summary = pd.DataFrame({
    "ladder": ["Drinking water", "Sanitation", "Hygiene"],
    "at_least_basic": [0.547, 0.432, 0.341],
    "denominator": ["all households"] * 3,
})
print(summary)
```

The three ladders cannot be averaged — In R

```
tibble::tribble(  
  ~ladder,          ~at_least_basic,  
  "Drinking water", 0.547,  
  "Sanitation",     0.432,  
  "Hygiene",        0.341  
)
```

The three ladders cannot be averaged

- Where a single composite is demanded, publish the three and the count of households that have all three — which is a . . .

What comes next

- The ladders classify.

Read the full lesson, with runnable code

```
https://data-analysis.cassion.dev/courses/wash-analysis/  
wash-02-sanitation-and-hygiene
```