As of when, from what, and how often

The three questions asked of every number on a dashboard, every month, by someone who was not in the analysis. Answering them on the page costs one line each and converts a recurring argument into a recurring glance.

Pierre Robentz Cassion

Lesson 3 of 8

Dashboards and Reporting — What travels with the number

What this lesson covers

- The three questions
- What the answers look like
- Refresh, stated as a cadence and an owner
- Completeness belongs next to the number
- What the platform's dashboards state
- Where provenance is thin, say so
- Report it whole
- What comes next

The three questions

- "As of when?" — A coverage figure with no date is a figure that will be quoted next March
- "From what?" — Which export, which version, which extract of the database — and whether it is the same one the last…
- "How often does it change?" — Whether the number in front of them is this morning's or last quarter's, and whether…
- Answer all three on the page — Each is one line, and the alternative is answering them by email eleven times

What the answers look like — Example

```
Source: wash-household-survey-2024.v1.csv
Data as of: 2024-11-30 (survey round 1)
Refreshed: when a new survey round is published; this is a single round
```

What the answers look like

- Name the file, not the system — "From DHIS2" is where it came from; the export's filename and version is what somebody. . .
- Version by filename — `.v1.csv` is what makes "the same file the March report used" a checkable statement
- Distinguish the data date from the build date — They are different and only one of them matters to a reader

What the answers look like

	Means	Put it on the page?
Data as of	The last observation in the file	**Yes, prominently**
Extract date	When the export was pulled	If they differ materially
Build date	When the HTML was written	**No**

What the answers look like

- The build date is the one everybody puts on and nobody needs — It changes on every rebuild, tells a reader nothing...
- None of the seven dashboards here carries a generation date — and that is why rebuilding them produces byte-identical...

Refresh, stated as a cadence and an owner — Example

```
Refreshed monthly, on the 5th, from the screening database export.
Owner: the M&E officer. If the export has not arrived by the 10th, the
page still shows the previous month and says so.
```

Refresh, stated as a cadence and an owner

- A cadence without an owner is a wish — The useful line names who does it and what happens when it does not happen
- Say what a stale page looks like — The failure mode is a dashboard silently showing last month while everyone assumes. . .

Completeness belongs next to the number — In Python

```
submitted = reports[reports["report_submitted"]].shape[0]
print(f"reporting completeness: {submitted / len(reports):.1%}")
```

Completeness belongs next to the number — In R

```r
mean(reports$report_submitted)
```

Completeness belongs next to the number

- A coverage figure computed from 76% of facilities is a different number from one computed from 99% — and the...
- So completeness goes on the page — beside any indicator whose denominator depends on who reported
- And a month with no reports is drawn as a gap, never as a zero — Zero is a claim that nothing happened

What the platform's dashboards state — Example

Source: protection-referrals-2024.v1.csv
How these numbers are produced: computed from the committed dataset file named above by scripts/export/build-project-dashboards.mjs. Nothing is typed by hand and the page contains no script, no external request and no generation date, so rebuilding it on unchanged data produces an identical file.

What the platform's dashboards state

- The last clause is the strongest provenance claim available — and it is checkable by anyone with the repository:. . .
- "Nothing is typed by hand" is the part that matters to a reviewer — A number in a caption that was typed can drift from. . .

Where provenance is thin, say so

- A number from another agency — "Cluster 4W, February submission" is provenance; "cluster data" is not
- A denominator from a projection — Population figures are usually projected from a census, and the projection's base…
- A number somebody adjusted — If a facility's obviously-wrong entry was corrected before analysis, that is a decision…

Report it whole — Example

Provenance panel

Source	wash-household-survey-2024.v1.csv 2,403 households, 18 communities, 3 districts
Data as of	2024-11-30, survey round 1
Refresh	single round; no monthly cadence. A second round is planned for 2025 and will make a trend meaningful.
Completeness	water quality tested on 33.4% of households; every quality figure states that denominator separately
Produced by	scripts/export/build-project-dashboards.mjs, from the committed file. No hand-entered numbers, no generation date, byte-identical on rebuild.
Definitions	docs/indicator-definitions.md

Report it whole

- The last line is the handover to the next lesson — A provenance panel says where the number came from; a definition…

What comes next

- Two artefacts on this platform computed improved water coverage from the same file and got different answers.

Read the full lesson, with runnable code

https://data-analysis.cassion.dev/courses/dashboards-and-reporting/dash-03-provenance-on-every-figure