

The report that changes the next round

Findings, root cause, corrective action, owner, deadline — five columns, one row per finding. Plus the follow-up round that is the only evidence any of it worked.

Pierre Robentz Cassion

Lesson 8 of 8

Data Quality Assessment — Turning findings into change

What this lesson covers

- Who actually reads it
- The order
- The finding table
- Findings a reader can check
- Score the dimensions, do not grade the district
- The follow-up round is the point
- Presenting it to the people in it
- Where this course, and this module, leave you

Who actually reads it

- **The programme manager** reads the summary and wants to know whether the figures in last quarter's report can stand.
- **The HMIS or M&E focal point** reads the finding table and wants to know what they are being asked to do.
- **The donor's monitoring adviser**, six months later, reads the methods section and wants to know whether your findings. . .

The order

- **What was assessed.** Dataset, period, reporting units in scope, indicators examined. Four lines.
- **How.** Which dimensions, which desk checks, which facilities were visited and *why those* — the `reason` column from...
- **The scorecard.** Five dimensions, five measures, with the one you could not measure shown as a finding rather than a...
- **Findings.** The table below.
- **What this means for published figures.** The section everyone skips and the programme manager needs.
- **Limitations.** What the sample cannot support, what you could not verify.

The finding table

#	Finding	Root cause	Corrective action	Owner	By
1	Reporting rate fell to 29% in August and 45% in September, district-wide	System: form supply interrupted, all facilities simultaneously	Confirm cause with logistics; hold two months' buffer stock of registers	District HMIS officer	30 Sep
2	SCH09 under-reported attendance by 42% (VF 1.42)	Tool: register prints Y/N boxes; the extract recognises only true/false	Accept Y/N in the extract's value mapping; reprint register at next order	HMIS data manager	15 Aug
3	Survey team 2 rounds heights to the nearest 0.5 cm (69% vs 18–23%)	Capacity: board read to the half centimetre	Retrain team 2 on board reading; supervisor to observe 10	Survey coordinator	Before next round

The finding table

- Every finding has a number in it — "Reporting was poor in August" is not a finding; "29%, against 82–92% in other. . .
- Every owner is a role, and a person holds it — "The district" cannot be chased
- Every deadline is a date — "Ongoing" and "as soon as possible" both mean never, and everyone in the room knows it when. . .

Findings a reader can check — In Python

```
findings = pd.DataFrame([
    {
        "id": 2,
        "dimension": "Accuracy",
        "unit": "SCH09",
        "measure": "verification factor",
        "value": 1.416,
        "comparator": "district 1.012, 23 of 24 schools exactly 1.000",
        "evidence": "outputs/verification-factors.csv",
    },
])
findings.to_csv("outputs/dqa-findings.csv", index=False)
```

Findings a reader can check — In R

```
findings <- tibble::tribble(  
  ~id, ~dimension, ~unit, ~measure, ~value, ~comparator,  
  2L, "Accuracy", "SCH09", "verification factor", 1.416, "district 1.012"  
)  
  
readr::write_csv(findings, here::here("outputs", "dqa-findings.csv"))
```

Score the dimensions, do not grade the district — In Python

```
scorecard = pd.DataFrame({
    "dimension": ["Completeness", "Timeliness", "Accuracy", "Consistency", "Integrity"],
    "measure": ["Reporting rate", "Not assessable", "Verification factor",
               "Outlier rate", "Round-number share"],
    "value": ["76.5%", "-", "1.012 district / 1.42 worst unit",
             "11 of 2,094 below half own median", "10.4%, chance level"],
    "finding": [1, 4, 2, None, None],
})
```

Score the dimensions, do not grade the district — In R

```
scorecard <- tibble::tribble(  
  ~dimension,    ~measure,          ~value,              ~finding,  
  "Completeness", "Reporting rate",    "76.5%",             1L,  
  "Timeliness",  "Not assessable",      "-",                4L,  
  "Accuracy",    "Verification factor", "1.012 district / 1.42 max", 2L,  
  "Consistency", "Outlier rate",        "11 of 2,094",       NA_integer_,  
  "Integrity",   "Round-number share",  "10.4%, chance level",  NA_integer_  
)
```

The follow-up round is the point — In Python

```
history = pd.concat([
    pd.read_csv(p).assign(round=p.stem)
    for p in sorted(Path("outputs/dqa").glob("*.csv"))
])
print(history.pivot_table(index="measure", columns="round", values="value"))
```

The follow-up round is the point — In R

```
history <- purrr::map_dfr(  
  list.files(here::here("outputs", "dqa"), full.names = TRUE),  
  ~ readr::read_csv(.x) |> dplyr::mutate(round = tools::file_path_sans_ext(basename(.x)))  
)  
  
tidyr::pivot_wider(history, id_cols = measure, names_from = round, values_from = value)
```

The follow-up round is the point

Measure	Q3	Q4	Movement
Reporting rate	76.5%	88.1%	+11.6 pt
Schools outside VF tolerance	1 of 24	0 of 24	closed
Team 2 rounded heights	69%	24%	closed

The follow-up round is the point

- Carry every open action forward — An action that appears in three consecutive reports without closing is itself a . . .

Presenting it to the people in it

- **Show it to them first.** A finding a facility sees for the first time in a donor report will be contested, and the. . .
- **Lead with the system findings.** Starting with the August form supply establishes that the assessment is looking at. . .
- **Name what was clean.** Twenty-three of twenty-four schools were exact. Say it before you say the twenty-fourth was. . .
- **Ask them for the cause.** The people who filled in the register usually know why it looks like that, and they will. . .

Presenting it to the people in it

The measure of a data quality assessment is not how many findings it produced. It is whether next quarter's data is better, and defensive reporting is the one outcome that guarantees it will not be.

Where this course, and this module, leave you

- You can profile a fresh export and log every change you make to it.

Read the full lesson, with runnable code

```
https://data-analysis.cassion.dev/courses/data-quality-assessment/  
dqa-08-the-dqa-report
```