CASSION · PROJECT REPORT

Seventy-two home visits before term end

School attendance and dropout early warning

Organisation	Education programme covering 24 schools
Prepared for	Education programme officer and head teachers
Updated	July 26, 2026
Sectors	Education
Standards and methodologies	Sustainable Development Goals (SDG)
Data	`school-attendance-2024`

DECISION THIS INFORMS

Which students receive a home visit before the end of term.

Every dataset behind this report is synthetic. No record traces to a real person, and no figure here describes a real programme.

Read the project online at data-analysis.cassion.dev/projects/school-attendance-analytics

Contents

Context

The problem

Dropout was only measured at year end, by which point the students had already left and the programme could report the number but never change it.

1 Recommendation

Visit **72 students** before the end of term, starting with the 15 who have missed the most school days since they were last present.

Head teachers receive their own school's list; the programme officer receives the school counts. A school whose flag rate sits well above the others is a conversation about the school, not about fifteen separate students.

This ranks risk, not dropout. A student on the list may have transferred, be ill, or be helping at home during harvest. The output is a conversation to have before the term ends — and a home visit that arrives treating the student as a dropout will be the last one that family accepts.

2 Why the measurement changed

Dropout was previously counted at year end, by which point the student had already left and the programme could report the number but not change it.

Attendance decays over roughly three weeks before a student stops coming, and that decay is the only reason early warning is possible. The flag is therefore a recent run of absence rather than a low total: **more than 15 school days missed since the student was last present.**

A student who missed a fortnight in February and came back is not the same case as one who has missed the last fortnight, and an attendance-rate threshold cannot tell them apart.

Full feature definitions and the threshold's provenance are in `docs/feature-definitions.md`.

3 The scale

	Value
Students on the roster	1,200
Student-days analysed	69,974
District attendance	88.4%
Students flagged	**72 (6.0%)**

4 Three data decisions that changed the list

A missing row is a closure, not an absence. SCH07 and SCH18 were closed for 15 school days, 11 to 29 March. Those student-days were never recorded. Building the features on a full-calendar grid invents them, and the obvious fill is "absent":

Treatment	Students flagged	Share from the two closed schools
Closure filled as absence	130	44.6%
Recorded days only	72	18.1%

Those two schools hold 9.8% of the roster. Filling closures as absence would have made them 45% of the visit list — a list of children whose school was shut.

Attendance coding had to be harmonised first. One school recorded `Y/N` while the rest recorded `true/false`. A boolean cast turns `Y/N` into missing values silently, and that school then supplies most of the top of any risk ranking — for a coding accident rather than for dropout.

Two students were on the roster twice, after an unregistered transfer left both registrations live. Unresolved, their attendance counts twice and both are double-weighted in every school rate.

5 The residual over-representation is reported, not adjusted

After the fix, the two closed schools still account for 18.1% of flags against 9.8% of students. That gap is left in place: **a three-week closure is a plausible trigger for real disengagement**, and adjusting it away would remove exactly the students the programme most needs to visit.

School	Students	Flagged	Rate
SCH07	61	7	11.5%
SCH24	54	6	11.1%
SCH18	56	6	10.7%
SCH15	47	5	10.6%
SCH09	49	4	8.2%
SCH14	48	4	8.3%
SCH01	55	4	7.3%
SCH22	56	4	7.1%

SCH07 and SCH18 are the two that closed. SCH24 and SCH15 are not, and their rates are comparable — those two are the schools to ask about.

6 The visit order

Sorted by how long the student has been gone, then by how they attended before. A student who was at 95% and stopped three weeks ago is a different case from one at 40% all term, and the second is a longer conversation.

Roughly 3% of roster rows have no grade recorded. Those students are still ranked — grade is not used in the score. It appears on the list only so a head teacher knows which class to ask about, and a missing grade must never remove a student from a visit list.

7 What this does not establish

There is no outcome label in this data, so the flag has no measurable sensitivity or specificity. The threshold is a programme judgement, and the honest framing is "students the programme wants to ask about", not "students who will drop out".

Transfers are invisible. A student who moved to another school looks identical to one who stopped attending, and the two roster duplicates are evidence that transfer recording is imperfect.

Absence has no reason code. Illness, labour, cost and disengagement are indistinguishable here and call for different responses.

A student who disengages in the final fortnight cannot be flagged, because they cannot accumulate 15 missed days before the term ends.

8 Recommended next steps

1. **Run the visits before term end.** The whole value of the measure is that it arrives while the school year can still act on it.
2. **Run the flag mid-term as well as at term end**, so students who disengage late are not invisible to it.
3. **Ask SCH24 and SCH15 about their flag rates.** They are as high as the two schools that closed and have no closure to explain them.
4. **Add a reason code to the attendance form.** Illness and labour need different responses, and nothing in the current data separates them.

About this edition

This report is generated from the same source as the project page, so the two cannot drift. The online edition carries the analysis notebook, the dataset and any corrections issued since this file was produced.

Read the project online at data-analysis.cassion.dev/projects/school-attendance-analytics

Generated July 26, 2026.