When Learners Leave School: A Survival Analysis of Dropout

A basic research proposal · Basic Education Research Fund 2026

[Researcher 1], Senior Education Program Specialist (Planning and Research)
[Researcher 2], Education Program Supervisor · [Researcher 3], Teacher III (Guidance Designate)
[Schools Division Office], Department of Education

A sample with invented data, for teaching: nash.com.ph/blog

Introduction and rationale: who leaves, and when

148,222

junior high school dropouts in the Philippines in SY 2022–2023

EDCOM II (2026), p. 590

5.7%

of Grades 7 to 10 learners left school in SY 2023–2024: 7.4% of boys, 3.8% of girls

Albert et al. (2025), p. 23

0

dropouts: the aim BEDP 2030 sets for DepEd's schools, which lose the most learners in Grades 7 and 8

DepEd (2022), pp. 131, 133



Our division counts its dropouts once a year. It does not know in which weeks learners leave, or how well the first month's absences, kept on every School Form 2, foretell leaving.

Literature review: who leaves, and when

Albert et al. (2025)

Junior high leaver rates of 3.8% to 5.2% before the pandemic, 6.2% in SY 2020–2021 and 5.7% in SY 2023–2024; boys leave at about twice the rate of girls.

David et al. (2018)

Absences grow from September; overage children are at high risk; distance and travel cost make learners start a year but not finish it.

EDCOM II (2026)

No policy requires routine, data-driven risk analysis; it recommends risk profiling from attendance records and other data.

So: the records to warn early exist; no one reads them week by week.

Literature review: early signs, and studying when

Balfanz et al. (2007)

Attending 80% of the time or less in Grade 6 was one of four flags that found 60% of later non-graduates; 13% of flagged students graduated on time.

Olson (2014)

Students who missed more than four days in September were 16 times as likely to be chronically absent over the year.

Willett & Singer (1991); Cox (1972)

Ask not whether learners leave but when; a hazard model relates the risk at each moment to a learner's traits.

So: we follow every learner through the year and ask when, not only whether.

Four questions, three null hypotheses

1. How many junior high learners left school during SY 2025–2026, and in which weeks?
2. Do the shares still in school differ by early absences, distance, age for grade, 4Ps membership and sex?
3. With the others held constant, how is each associated with the hazard of leaving?
4. Does the association of early absences hold through the whole year?

Null hypotheses (α = .05)

- $H_0 1$: The survival curves do not differ by any of the five traits.
- $H_0 2$: With the others held constant, none of them is associated with the hazard of leaving.
- $H_0 3$: The hazard ratio of early absences is the same before and after the year-end break.

What the study covers, and what it cannot show

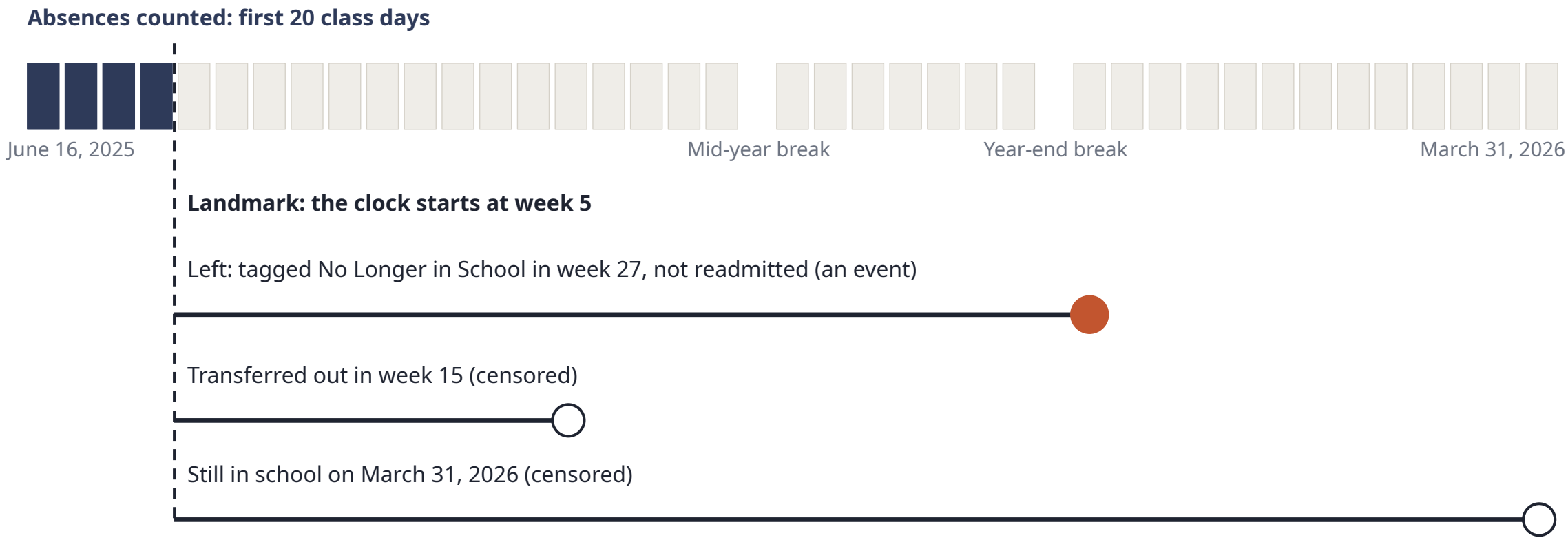
Scope

- All Grades 7 to 10 learners of the division's 18 public secondary schools, SY 2025–2026
- About 10,600 learners' records: no sampling
- School Form 2, the LIS and the enrollment form
- Records read after the school year ends, May to June 2026

Limitations

- Observational: associations, not causes
- Records are only as good as their keeping
- Distance from the barangay, not the home
- One school year; leaving between years is outside it

Research design: a cohort followed week by week



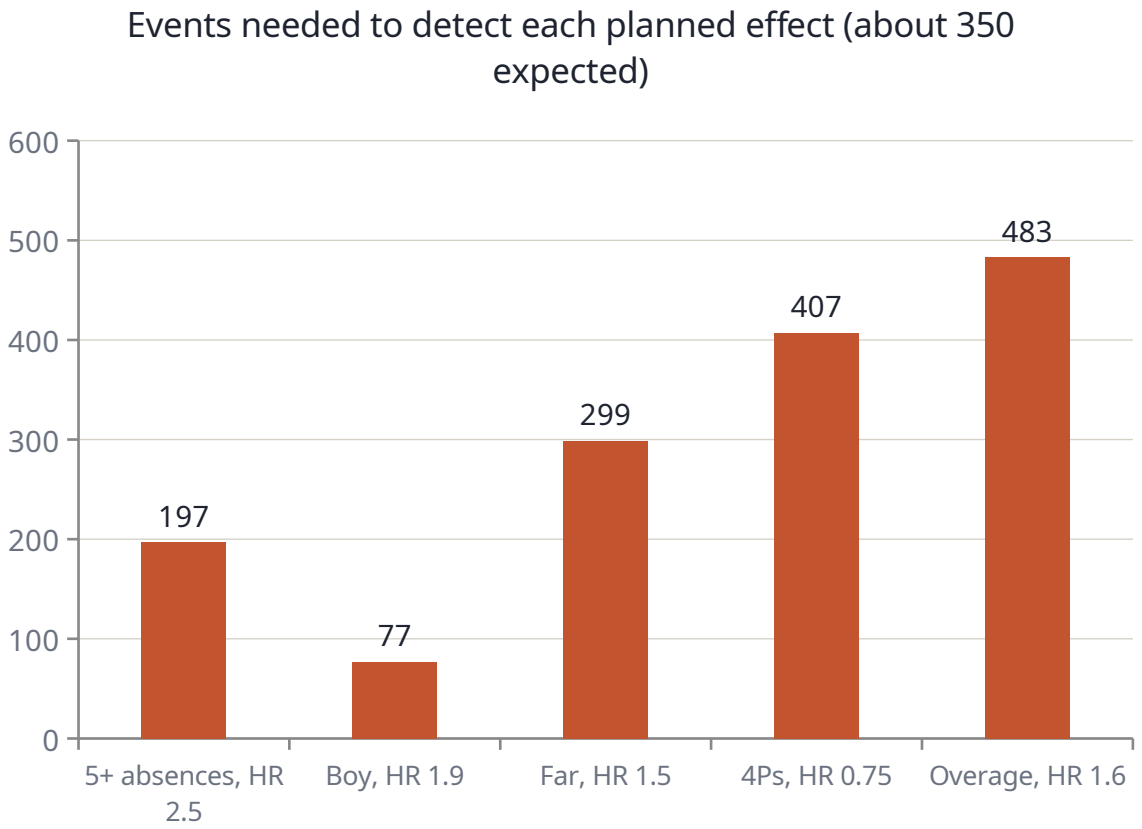
A retrospective cohort: every learner's record followed through the 39 school weeks of SY 2025–2026.

Sampling: every record, and the events that count

18 public secondary schools in the division

10,615 Grades 7 to 10 learners at the beginning of SY 2025–2026: all of them

≈ 350 learners expected to leave, at the division's dropout rate of about 3.3% the year before



Schoenfeld's (1983) formula, α = .05, power .80: enough events for early absences, sex and distance, not for overage or 4Ps at the planned sizes.

Data collection: records the schools already keep

The records extract (LIS)

Grade, sex, birth date, status at the end of the year, and the effective date of any transfer or No Longer in School tag; 4Ps from the enrollment form; the barangay from School Form 1. Codes in place of names.

The School Form 2 abstraction sheet

Each learner's absences in the first 20 class days, June 16 to July 11, 2025, and the reason School Form 2 records for a learner who left.

Derived variables

The school week of leaving (1 to 39, breaks skipped); absences 0 or 1, 2 to 4, 5 or more; overage by two years or more; road distance from the barangay, far beyond 5 km.

Data collection: validity, reliability, schedule

A record check

A second researcher re-abstracts a random 10% without seeing the first: the ICC for absences, Cohen's kappa for status, the same week for leavers.

Definitions fixed in advance

Leaving: still No Longer in School at the end of the year. Transfers are censored; learners readmitted later are not counted as leaving.

A landmark at week 5

The models start after the first 20 class days, so early absences never predict leaving that happened while they were counted.

Checks of the model

Schoenfeld residuals for proportional hazards; a discrete-time model and a competing-risks estimate as checks.

Approval and data request April 2026 · Ethics review April to May · LIS extract May 11 to 15 · School Form 2 abstraction May 18 to June 5 · Record check June 15 to 19 · Analysis July to August.

Ethical issues: consent, privacy, and review

Consent

Families cannot practicably be asked one at a time: an ethics committee reviews the use of coded records without individual consent.

Privacy

A coded extract with no names or Learner Reference Numbers; the key stays with the Planning Officer (RA 10173).

No decisions

The study's file is never used to decide anything about a learner; results name no learner or school.

Approval

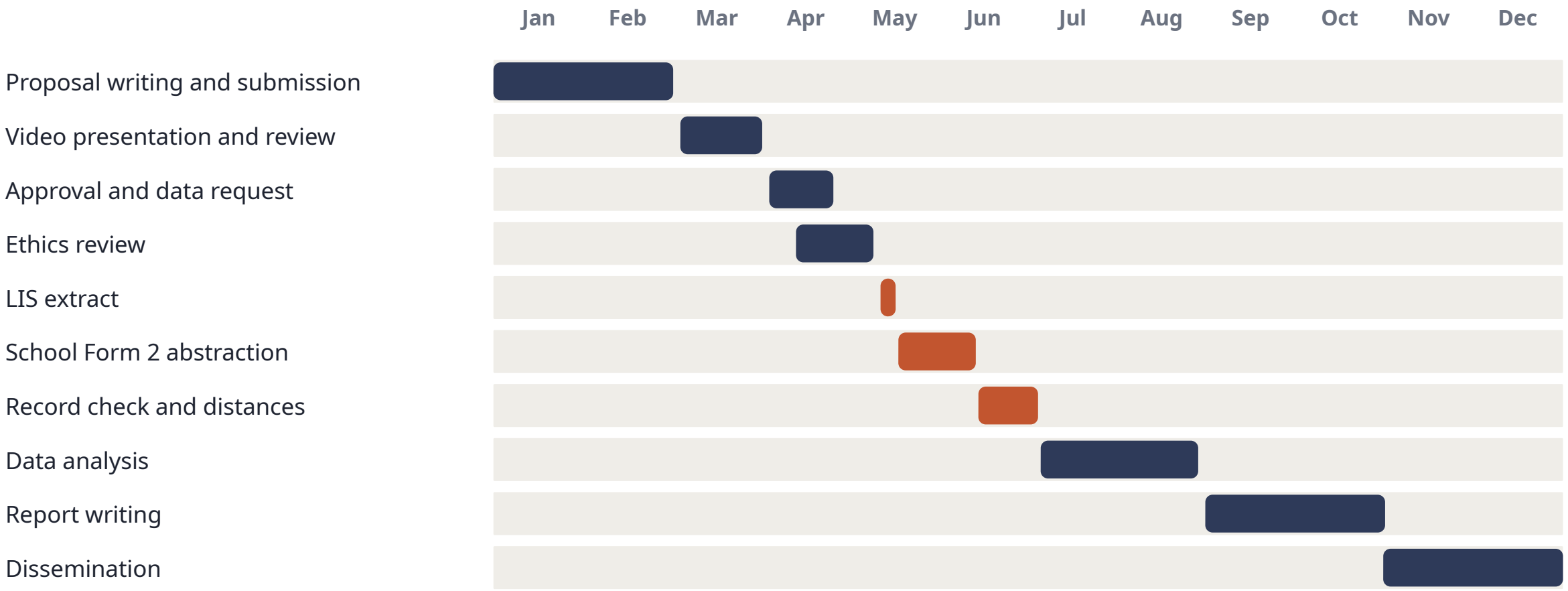
The Schools Division Superintendent's approval and the Division Research Committee's review.

Plan for data analysis: Kaplan-Meier and Cox

Question 1	Counts and the simple dropout rate by grade and sex; leavers by week; the Kaplan-Meier share still in school; transfers as a competing event.
Question 2	Kaplan-Meier curves by group from the landmark at week 5; log-rank tests.
Question 3	Cox regression, stratified by grade, standard errors clustered by school, Efron's method for tied weeks.
Question 4	Scaled Schoenfeld residuals against time; if a ratio changes, let it differ before and after the year-end break.

Python 3.13 with lifelines and statsmodels; Excel recomputes the life table, a log-rank test and every statistic it can.

Timetable: January to December 2026



Terracotta: work with the records in the division office and the schools. Months are of 2026.

Cost estimates: PhP 15,000 in eligible items

Item	Particulars	Amount (PhP)
Food	Snacks for the abstraction days	2,700.00
Reproduction	School Form 2 abstraction sheets (about 600 pages)	1,200.00
Transportation	36 trips to the 18 schools	3,600.00
Communication	Mobile data, three researchers, five months	4,500.00
Supplies	Folders, envelopes, an encrypted flash drive	1,600.00
Dissemination	Policy brief	1,400.00
Total		**15,000.00**

Not charged to the grant

- Statistical software: Python and its libraries are free
- Honoraria for the record check: the researchers do it
- Equipment, utilities and office rent

Dissemination and advocacy: who needs the findings

School heads and guidance designates

Division school heads' conference and a two-page policy brief: act on the first month and the weeks after the breaks.

November 2026

Planning and monitoring units

Division management committee: build the early warning list into routine reports.

December 2026

Class advisers

Learning action cell sessions: when follow-up matters most.

January to February 2027

Researchers

Regional research congress and the division's research journal.

March to June 2027

References

Albert, J. R. G., Monterola, S. L. C., Cabalfin, D. L. D., & Mahmoud, M. A. (2025). Out-of-school children in the Philippines: Post-pandemic realities and emerging challenges at the educational spectrum extremes (PIDS Discussion Paper No. 2025-34). PIDS.

Balfanz, R., Herzog, L., & Mac Iver, D. J. (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Early identification and effective interventions. Educational Psychologist, 42(4), 223–235.

Cox, D. R. (1972). Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological), 34(2), 187–202.

David, C. C., Albert, J. R. G., & Vizmanos, J. F. V. (2018). Out-of-school children: Changing landscape of school attendance and barriers to completion (PIDS Discussion Paper No. 2018-25). PIDS.

Davidson-Pilon, C. (2019). lifelines: Survival analysis in Python. Journal of Open Source Software, 4(40), 1317.

Department of Education. (2014). Adoption of the modified school forms (SFs) for public elementary and secondary schools effective end of school year 2013-2014 (DepEd Order No. 4, s. 2014).

Department of Education. (2022). Adoption of the Basic Education Development Plan 2030 (DepEd Order No. 024, s. 2022).

Department of Education. (2025). Multi-year implementing guidelines on the school calendar and activities (DepEd Order No. 012, s. 2025).

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. Journal of the American Statistical Association, 53(282), 457–481.

Olson, L. S. (2014). Why September matters: Improving student attendance. Baltimore Education Research Consortium.

Orbeta, A. C., Jr., Melad, K. A. M., & Araos, N. V. V. (2021). Reassessing the impact of the Pantawid Pamilyang Pilipino Program: Results of the third wave impact evaluation (PIDS Discussion Paper No. 2021-05). PIDS.

Schoenfeld, D. A. (1983). Sample-size formula for the proportional-hazards regression model. Biometrics, 39(2), 499–503.

Second Congressional Commission on Education. (2026). Turning point: A decade of necessary reform (2026–2035) [EDCOM II final report].

Willett, J. B., & Singer, J. D. (1991). From whether to when: New methods for studying student dropout and teacher attrition. Review of Educational Research, 61(4), 407–450.

Thank you

[Researcher 1] · [DepEd email address] · [mobile number]
[Researcher 2] · [DepEd email address]
[Researcher 3] · [DepEd email address]
[Schools Division Office], [address]