

Appendix: Data definitions

1. **Schools.csv:** This table is a record of schools.

Column name	Description
school_id	Unique identifier, assigned for each school
name	Name of the school
location	Suburb, state, and postcode of the school
ASGS-RA_2016	Australian Statistical Geography Standard (ASGS) remoteness structure code 1 = major cities of Australia 2 = Inner Regional Australia 3 = Outer Regional Australia 4 = remote areas of Australia 5 = very remote parts of Australia

2. **Students.csv:** This table is a record for each student.

Column name	Description
student_id	Unique identifier for the student
student_name	Name of the student, pseudonymous/anonymous
gender	Gender of the student

3. **Enrolments.csv:** This table indicates which student is enrolled where.

Column name	Description
enrolment_id	Unique enrolment identifier for a student enrolled at a given school in a given calendar year
student_id	Unique identifier of the enrolled student
year_of_enrolment	Calendar year when the student enrolled
enrolled_year_level	Level of education for which the student enrolled
enrolled_school_id	Unique identifier of the school where the student is enrolled
enrolled_class	The student's class associated with the enrolment
flag_current_enrolment	1 if the student was currently enrolled and 0 otherwise

4. **year_level_mapping.csv:** This table maps year_level_id to its corresponding year level.

Column name	Description
year_level_id	Identifier for the level of education
year_level	Level of education in 'words' corresponding to year_level_id

5. **ae_assessment.csv:** This table gives the A–E assessment results for students, across all schools.

Column name	Description
id	Unique identifier of a student's particular A–E assessment
assessment	Name of the assessment (here, only A–E)
assessment_id	Identifier of the subject for which the student appeared for assessment (consult ae_assessment_mapping.csv for details)
student_id	Unique identifier of the student who appeared for assessment
school_id	Unique identifier of the school where the student was enrolled
year_level_when_assessed	Level of education for which the student appeared for assessment
year_when_assessed	Calendar year in which the assessment took place
term_when_assessed	School term in which the assessment took place
result	The score the student obtained in the assessment

6. **ae_assessment_mapping.csv:** This table maps ae_assessment_id to its corresponding domain name.

Column name	Description
ae_assessment_id	Unique identifier of a particular A–E assessment domain
assessment	Name of the A–E assessment domain corresponding to the ae_assessment_id

7. **pat_assessment.csv:** This table presents the PAT results for students in years 1–10. You can find further details about the assessment [here](#).

Column name	Description
id	Unique identifier of a student's particular PAT assessment
assessment_code	Name of the assessment (here, only PAT)

domain_id	Unique identifier of a domain of PAT assessment
assessment_name	Name of the domain corresponding to the domain_id
student_id	Unique identifier of the student who appeared for assessment
school_id	Unique identifier of the school where the student was enrolled
year_level_when_assessed	Level of education for which the student appeared for assessment
year	Calendar year in which the assessment took place
scale_score	The score the student obtained in the assessment (the max. possible score is 50, and the min. score is 0)

8. **naplan_assessment.csv:** The table contains the NAPLAN results for students in years 3, 5, 7, and 9. This assessment is done once a year. The mapping table `naplan_pat_mapping.csv` shows the domains for NAPLAN. You can find more information about the assessment [here](#).

Column name	Description
id	Unique identifier of a student's particular NAPLAN assessment
assessment_name	Name of the assessment (here, only NAPLAN)
domain_id	Unique identifier of a domain of NAPLAN assessment
student_id	Unique identifier of the student who appeared for assessment
school_id	Unique identifier of the school where the student was enrolled
grade_level	Level of education for which the student appeared for assessment
year_assessed	Calendar year in which the assessment took place
score	The score the student obtained in the assessment (the max. possible score is 700, and the min. score is 0)

9. **naplan_mapping.csv:** This table maps a domain ID from NAPLAN assessment to its corresponding domain name (in words).

Column name	Description
assessment	Name of the assessment (NAPLAN)
domain_id	Unique identifier of the domain pertaining to NAPLAN assessment (as in <code>naplan_assessment.csv</code>)
domain	Name of the domain corresponding to the assessment and domain_id

10. **pat_mapping.csv:** This table maps a domain ID from PAT assessment to its corresponding domain name (in words).

Column name	Description
assessment	Name of the assessment (PAT)
domain_id	Unique identifier of the domain pertaining to PAT assessment (as in pat_assessment.csv)
domain	Name of the domain corresponding to the assessment and domain_id

11. **Attendance.csv:** This table is a record of a given student's attendance.

Column name	Description
enrolment_id	Unique enrolment identifier for a given student enrolled at a given school in a given calendar year
attendance_date_id	Unique identifier for the attendance date (it is further explained at the end of the table)
day_of_week	The given date's day of the week
session_no	The given session number of the day
year_level_when_attended	The student's level of education associated with the given attendance record
school_id_when_attended	Unique identifier of the school attended by this student
class_when_attended	The student's class associated with the given attendance record
participation_code	Numeric code indicated whether this student was present, absent, or late. Here, 0 = absent, 1 = present, and 2 = late

Note:

The column **attendance_date_id** represents the generalised date* for a school and it is in the form of <year>_<term>_<week>_<day_in_week>.

In your example, the attendance_date_id = 2021_2_10_4 can be translated as:

- Year 2021
- Term 2
- Week 10
- Day 4 (Thursday) or Day 3 (Wednesday)

*Generalised date is used instead of a normal date (2/2/21) due to how there are some schools that would start on different dates. Example Public schools may start Term 1 on 21st January whilst Private schools may start the week after on the 28th of January.