

The “StudentPerformance” dataset hosted on UCI contains approximately 1000 observations, with 33 attributes. It includes student performance data of two courses: Math and Portuguese language. Math data set includes 395 students. And Portuguese language includes 649 students. They all have the same attributes. The output target variable is the final grade of student.

**Attributes(33):**

- 1 school – student’s school (**binary**: “GP” – Gabriel Pereira or “MS” – Mousinho da Silveira)
- 2 sex – student’s sex (**binary**: “F” – female or “M” – male)
- 3 age – student’s age (**numeric**: from 15 to 22)
- 4 address – student’s home address type (**binary**: “U” – urban or “R” – rural)
- 5 famsize – family size (**binary**: “LE3” – less or equal to 3 or “GT3” – greater than 3)
- 6 Pstatus – parent’s cohabitation status (**binary**: “T” – living together or “A” – apart)
- 7 Medu – mother’s education (**categorical**: 0 – none, 1 – primary education (4th grade), 2 – 5th to 9th grade, 3 – secondary education or 4 – higher education)
- 8 Fedu – father’s education (**categorical**: 0 – none, 1 – primary education (4th grade), 2 – 5th to 9th grade, 3 – secondary education or 4 – higher education)
- 9 Mjob – mother’s job (**categorical**: “teacher”, “health” care related, civil “services” (e.g. administrative or police), “at\_home” or “other”)
- 10 Fjob – father’s job (**categorical**: “teacher”, “health” care related, civil “services” (e.g. administrative or police), “at\_home” or “other”)
- 11 reason – reason to choose this school (**categorical**: close to “home”, school “reputation”, “course” preference or “other”)
- 12 guardian – student’s guardian (**categorical**: “mother”, “father” or “other”)
- 13 traveltime – home to school travel time (**categorical**: 1 – <15 min., 2 – 15 to 30 min., 3 – 30 min. to 1 hour, or 4 – >1 hour)
- 14 studytime – weekly study time (**categorical**: 1 – <2 hours, 2 – 2 to 5 hours, 3 – 5 to 10 hours, or 4 – >10 hours)
- 15 failures – number of past class failures (**categorical**: n if  $1 \leq n < 3$ , else 4)
- 16 schoolsup – extra educational support (**binary**: yes or no)

17 famsup - family educational support (**binary**: yes or no)

18 paid - extra paid classes within the course subject (Math or Portuguese) (**binary**: yes or no)

19 activities - extra-curricular activities (**binary**: yes or no)

20 nursery - attended nursery school (**binary**: yes or no)

21 higher - wants to take higher education (**binary**: yes or no)

22 internet - Internet access at home (**binary**: yes or no)

23 romantic - with a romantic relationship (**binary**: yes or no)

24 famrel - quality of family relationships (**categorical**: from 1 - very bad to 5 - excellent)

25 freetime - free time after school (**categorical**: from 1 - very low to 5 - very high)

26 goout - going out with friends (**numeric**: from 1 - very low to 5 - very high)

27 Dalc - workday alcohol consumption (**numeric**: from 1 - very low to 5 - very high)

28 Walc - weekend alcohol consumption (**numeric**: from 1 - very low to 5 - very high)

29 health - current health status (**numeric**: from 1 - very bad to 5 - very good)

30 absences - number of school absences (**numeric**: from 0 to 93)

# these grades are related with the course subject, Math or Portuguese:

31 G1 - first period grade (**numeric**: from 0 to 20)

31 G2 - second period grade (**numeric**: from 0 to 20)

32 G3 - final grade (**numeric**: from 0 to 20, **output target**)