

# Analysis of Student Performance Dataset

## Goals

1. Build a classifier to predict the probability of a student failing or passing Maths and Portuguese
2. Predict which major attributes affect student results.
3. Classify student pass and fail rates by the attributes of their guardians (father or mother, guardian's job, education etc.)

## Source Data

<https://archive.ics.uci.edu/ml/datasets/Student+Performance>

## Analysis methods

Both predictive and descriptive methods. Use descriptive methods to classify the students' pass and fail rates by the attributes of their guardians and use predictive methods in building classifiers that can predict the probability of a student failing or passing Maths and Portuguese. Some possible task are:

- Exploratory analysis
- Linear classification
- Machine learning techniques such as linear regression, random forest, etc.

## Expected results

1. Having a clear idea of the probabilities of whether the student would pass or fail.
2. Visual representation demonstrating the attributes that mostly affect a student's performance.
3. Predict student pass or fail rate by the attributes of their guardians

## Estimated team size / required time

A team of four is required with estimated completion time of three weeks.