

Homework 1

(due: 23:59, 29/Sep/2020)

Use the **Student_data.xlsx** (*Homework 1 - dataset* in Moodle) dataset to conduct a regression analysis and obtain a model explaining the student final grade (final column of the dataset).

1. Choose all or several variables which you hypothesize to affect the student's final grade.
2. Build a linear regression model by obtaining the model parameters using OLS method.
3. Evaluate the overall model, check internal validity (diagnose to see whether assumptions are met or not).
4. Extend the model (using transformation, selection or shrinkage methods) to obtain a better model, if possible.
5. Select the model which performs better on the test set and/or passes diagnostics best.
6. Interpret the results obtained from the final model.