ASSIGNMENT 2 (10 POINTS)

• This assignment will be solved individually. You must upload your solution (**html and Rmd files**) at Aula Digital, be sure to include your full name at the top of your solution. NO NAME, NO GRADE!

• **Due date**: November 29th, 2023 at 23.55h.

• **NO late assignments will be allowed.**

To complete this assignment create a .Rmd file with your answers. Once completed, generate an .html document and upload both documents at Aula Digital in the space provided. Make sure your name is visible at the top of the document and show your R code.

This assignment uses the data frame `Prestige` included in the "car" package. Type "?Prestige" for a description of the data frame.

1.- Give a summary of the `Prestige` data frame. Indicate the size (number of observations and variables) and type of variables.

2.- The focus in this assignment is the regression model, where "income" will be the dependent variable. Draw different plots to see the relationship between the variables.

3.- Consider only variables "income" and "education". Draw a histogram for variable "income". Draw the boxplot for "education". Draw a scatterplot of both variables to see their relationship. Fit a linear regression and see how well this model fits the observed data. Draw the model over the scatterplot. What do you observe? Do the regression assumptions hold?

4.- Consider the polynomial regression model. How well does the polynomial model fit the data? Compare the R^2 of both linear and polynomial models. What do you conclude?

5.- Would some transformation procedure of the variables, such as log or square root, allow for a better model? Try out different transformations and plot the transformed variables. Which transformation is the "best"? For the best transformation, fit a regression model.