Home-Assignment

This is an individual home assignments. You are allowed to support each other, but **identical or copied submissions will fail**. The home assignment (code and answers to exercise below) has to be submitted until Sunday August 1, 2021, by email to anthony.strittmatter@ensae.fr.

Task:

You take the role of a used car dealer. You have a database with prices and characteristics of used cars you sold previously (*use_car_database.csv*). You want to use this data to predict the prices of the used cars that are currently in your garage. For these cars you have a data set containing the characteristics (*new_used_cars.csv*), but not the prices. The characteristics that you observe in the data sets are described in the file *variable_description.xlsx*.

1. Predict the used car prices of the used cars that are currently in your garage using any method of your choice. Submit the *id* and the predicted *sales_price* as well as your code.

2. Submit your answers to the questions below.

Questions:

Short answers are appreciated (keywords or one short sentence).

1. Why do we use for predictions different samples to train and test models?

2. How can we select the tuning parameter of the Lasso?

3. Why is it difficult to interpret Lasso coefficients?

4. What is the difference between supervised and unsupervised machine learning?

5. Explain reinforcement learning intuitively?