Exercise 1: Feature Effects
Make yourself familiar with the **iml** package in R and the method `FeatureEffects`. Answer the following questions for your chosen dataset

(a) Create PDP and ALE plots for at least 6 features of your dataset you are most interested in how they affect your target variable. Can you see any differences between PDP and ALE plots? Can you explain those differences?
Based on the PDPs: Which of the 6 features would you consider most and which one least important?

(b) Create ICE plots for the same features as above and interpret the difference to the PDPs. Are your chosen features interacting with any other features?

Exercise 2: Feature Importance
Answer the following questions for your chosen dataset

(a) Make yourself familiar with the `FeatureImp` method in the **iml** package and calculate the permutation feature importance (PFI) for all features in your dataset. Calculate the PFI on your training and on your test dataset. Do you observe any differences? Interpret your results.

(b) Write a function which calculates the LOCO feauture importance on the test dataset. How does it differ to the PFI results? Explain your observation.
Hint: Refit your model on the training data without the feature of interest. Predict with the refitted model on the test dataset and calculate the loss difference between your original model and your refitted model. Repeat for all features.

(c) Compare your feature importance values to your feature effect results of the first exercise sheet. Are they conclusive?

Exercise 3: Local Methods
Answer the following questions for your chosen dataset

(a) Choose 2-3 observations from your dataset which you would like to explain. How do the different features influence the predictions of the regarded observations? Use LIME for your analysis and interpret your results.
Hint: Use `LocalModel` from the **iml** package.

(b) Your task is to choose an observation in your dataset of which you want to change the label (e.g. credit accepted vs. credit declined in a classification task or change a wine with a rating of 5 to a predicted wine rating of 6 in a regression task). Try to find such a counterfactual by changing as few features as possible (e.g., 1-2 features) and with rather small changes.
Bonus: Try the **counterfactuals** package (`https://github.com/dandls/counterfactuals`), e.g. with the method `MOCRegr` or `MOCClassif`, and compare your results.

Exercise 4: Shapley Values
Make yourself familiar with the method `Shapley` in the **iml** package. Answer the following questions for your chosen dataset

(a) Choose one or two observations of your dataset which might be worth to explain via Shapley values. Calculate and visualize the Shapley values for these observations and interpret your results.

(b) Sample 100 observations from your dataset at hand and calculate the Shapley values for these observations

- Calculate the Shapley feature importance for your sample. How does it differ from the PFI?
- Create a Shapley dependency plot for the 6 features you have chosen in the feature effects exercise and compare your plots with the PDPs.