Problem n.3

You and your closest friends have decided to rent a boat for your next holiday in Greece. The file `boats.txt` reports data on boat rental collected from a website. The dataset reports, for 80 search results, the price of the boat rental [€], the length of the boat [meters], the engine power [KWatt], the draught (depth of water needed to float a ship) [meters], the number of crew members, the year of construction and the deck material (wood or fiberglass).

a) Formulate a linear regression model for the price of the boat rental, as a function of all the other variables. Include in the model a possible dependence of the price of the boat rental on the categorical variable 'deck material', but only in the intercept. Report the model and its parametrization, together with the estimates of all its parameters. Verify the model assumptions.

b) Using the appropriate statistical test, state if you deem necessary to include into the model the variables related to the dimension of the boat (i.e. the length of the boat, the engine power and the draught).

c) Using the appropriate statistical test, state if you deem necessary to include into the model the variables related to accessory features (i.e. number of crew members and deck material).

d) Based on appropriate test(s), reduce the model and update the model parameters.

e) You want to rent a boat with the following characteristics: length of the boat = 10 meters, engine power = 1070 KWatt, draught = 1.5 meters, number of crew members = 1, year of construction = 2015 and deck material = fiberglass. Using the last model, compute a pointwise estimate and a prediction interval of level 95% for the price of the boat rental.

Upload your results here:
https://forms.office.com/Pages/ResponsePage.aspx?id=K3EXCvNtXUKAjjCd8ope612LHtvIHvFEsEi2L6mhPg
lUNkxGTzRJRDVRVE5DV0xROUdHMTBVUjlUNC4u