Homework 3: Linear Regression

This homework is aimed at practicing linear regression with model selection and with interactions, as well as interpreting the output of such models.

The data is subset from a well-known experimental study of 925 English nouns by Paivio et al. 1968 (see paper). The subjects were asked to rate the nouns on concreteness, imagery and so-called meaningfulness. Concreteness is the most transparent parameter: concrete words are those that denote tangible objects, materials or persons that can be easily perceived with the senses. Imagery scores relect how quickly and easily the word arouses a mental image deined as sensory experience, such as a mental picture or sound. The subjects were asked to rate concreteness and imagery on a scale from 1 (minimum) to 7 (maximum). There is also information about word length in syllables and letters. A detailed description of the data frame can be found in Levshina 2015:89-91 and in the original paper. Please familiarize yourself with the study before coding. Our research question is what characteristics of a word predict how many associations (on average) a person makes.

Complete the instructions in the R file. Make sure to answer questions in full sentences as a comment. Note that I have combined two data frames into one and that we will use them for a different analysis than is presented in the text book - thus don't get confused when you get different numbers for summary statistics and don't copy/paste from the text book.

In a separate file, answer the following questions based on your results:
  a) In about 2 paragraphs summarize your answer to the research question based on the analysis. Mention which predictors turned out to be significant and which not. Offer possible explanations of why that might be.
  b) For each significant predictor or interaction, describe the kind of effect they have on the dependent variable.
  c) Finally, mention anything that was surprising or striking to you and suggest one question or issue for further research.

You can work in groups of up to two people – in that case you will turn in the <u>same file</u> and receive the same grade. Please state clearly who you worked with.

Submission format:
  • code as .R file; text responses as .pdf
  • your files must include:
    ◦ your name(s) – if you worked with someone else, this has to be explictly stated!

Upload on ILIAS by 10:15 on November 13, 2024

Rubric for evaluation:

There is usually more than one correct answer and points will be given for all versions that work/make sense.

Code	Points
1	1
2	2 each
3a	4
3b & 3c	2 each
Text answers	
a)	4
b)	2
c)	1
Max. points possible	**20**
Late submission	−3 points per day