Exercises
(Linear Regression)

Exercise 1: plotting regression lines

(easy warm-up exercise)

Using the RT-data example from the lecture (RT as a function of lexical frequency and spelling of words; http://www.psy.gla.ac.uk/~christop/MScStats/2018/Regress/RTs.csv), create a scatterplot with the following features:

- Red symbols for "lowercase" and blue symbols for "uppercase" data points
- Two regression lines (red for the linear relationship between *logfreq* and *RT* in the "lowercase" spelling condition; blue for the linear relationship between *logfreq* and *RT* in the "uppercase" spelling condition)
- Extra points if you also include 95% confidence intervals around the lines

Question (briefly explain your answer): What kind of effect would one use such a plot as a visualisation for?

- The main effect of *logfreq*
- The main effect of *spelling*
- The *spelling* × *logfreq* interaction

Exercise 2: `lm()` with categorical predictors

(slightly tricky – clever predictor coding required!)

The following link takes you to a new (fabricated) independent-measures dataset:
http://www.psy.gla.ac.uk/~christop/MScStats/2018/Regress/HW_1_2.csv

It contains three orthogonally manipulated categorical predictor variables, namely *A* (with levels *a1* and *a2*), *B* (with levels *b1* and *b2*), and *C* (with levels *c1* and *c2*). The dependent variable (DV) is continuous and normally distributed, and there are 160 cases altogether.

(1) Using the function `lm()` in R, run the equivalent of a 2 × 2 × 2 ANOVA to determine which ***overall main effects and interactions*** are significant*. Report corresponding *F*- and *p*-values.

(2) Using the function `lm()` in R, run ***follow-up tests*** to decompose any significant* **two-way** interactions from the previous omnibus analysis into simple effects. Again, report corresponding *F*- and *p*-values.

(3) Using the function `lm()` in R, ***hierarchically decompose the three-way interaction*** (*A* × *B* × *C*) into simpler effects. Specifically, determine separate "simple" two-way interactions (*B* × *C*) for each level of *A*; if any of these simple *B* × *C* interactions is significant*, decompose it further into simple effects of *C* for each level of *B*. Again, report *F*- and *p*-values throughout.

(*Note: "significant" means $p \leq .05$)

Exercise 3: determine `lm()` predicted values "by hand"

(easy again)

From the omnibus analysis in part (1) of the previous exercise, use the model coefficients to determine the predicted mean DV for the following design cell: *A*=*a2*, *B*=*b2*, *C*=*c1*.