\*\*Submission is in pairs –both of your exercises should be submitted in your moodle profiles. Only one of them will be tested. \*\*

The exercise should be submitted in the relevant place in the Moodle. You can submit the exercise in one of two ways:

1. Using a word file to answer the text questions and an R code for the code. Make sure to note which question exactly you are answering. You can skip the words file and add the answers as notes in the code.

2. Using R Markdown, submitting the .Rmd file and the output in html format. Answer the text question between blocks in the code.

-The first lines in the code and word file should specify your names and IDs.

-In any case, for a perfect score - all parts of code you used to answer the questions should be submitted and it should be clear which code part refers to which question. We highly recommend adding as many comments as possible to your code.

-Before running the code, we will change only the setwd() line to set a new working directory. Other than that, all code should run smoothly.

-Answers to text questions should be concise and to the point. Unless instructed otherwise, use no more than 75 words for each answer (3-4 lines in word with font 11). Answers that do not answer the question directly or that include a lot of unnecessary information will be marked as wrong.

**Analyzing factorial data**

In this exercise, you will analyze the data from:

Lande, A., Mayo, R. (2022). (Dis)Trust in Science: A Correction Based on Scientific Sources Backfires Unless a Self-Reasoning Strategy Is Used (submitted to review)

(We thank Anat Lande and Ruti Mayo for sharing their data)

In the study, subjects were presented with disinformation about the COVID-19 vaccine (that the vaccine can cause sterility in women). They were assigned to one of 6 groups, determined by two factors: (1) A treatment (named ThinkFRE) that was hypothesized to reduce belief in false information and increase trust in facts and experts / Control group without treatment, and (2) the type of correction given for the false information: correction with scientific source/correction without a source/no correction (baseline). The main DV was the belief that vaccine can cause sterility: mainDV\_sterility, and the columns for the two factors are “Group” and “Correction”.

It is a known finding that correction with scientific sources causes disbelief, and the manipulation was designed to reduce this effect. The research hypothesis was that the influence of thinkFRE (compared to control group) manipulation will be stronger for subjects who were given a correction with a scientific source.

1. Perform a power analysis for the interaction effect (not the main effects) using a simulation:
   1. Simulate a 2(group: control/thinkFRE) x3(correction: source/no source/no correction) two-way ANOVA design with 6 groups, distributed normally with equal SD of 1.5. Use the following population means and SD with a sample size of n=47 per sub-group:  
      ***Control***

Source: mean = 2.9, SD=1.3

No-source: mean = 2.3, SD = 1.3

No-correction: mean = 2, SD=1.3

***thinkFRE***

Source: mean = 1.8, SD=1.3

No-source: mean = 2, SD = 1.3

No-correction: mean = 1.8, SD=1.3

* 1. Run a 2-way ANOVA and save the p-value for the interaction. Before moving on, make sure your code works properly and that you save the desired p-value.  
     (Note - if “s” is the variable containing the summary of your anova, the code s[[1]]$`Pr(>F)`[3] will give you the p-value)
  2. Repeat this process 2,000 times to calculate and report the power of the study.

1. Use this value (the power) to calculate the PPV for 4 different values of R (between 0.2-0.8 in jumps of 0.2) as demonstrated in class. Plot the resulting PPV graph and explain **shortly** in your words what it shows.
2. Change the number of subjects in each sub-group. What is the approximate group size you need for getting a power of 95%?
3. Load the data and inspect it. Is there a reason to suspect a problem in the way we did the power analysis?
4. Clean the dataset by the following criteria:
   1. Delete all rows without data in the columns for Group, Correction and mainDV\_sterility.
   2. Only subjects who have the value “0” in the “exclude” column
   3. No outliers more than 3 SDs (in mainDV\_sterility) from the mean of each subgroup

How many subjects were excluded?

1. Calculate the variance within each sub-group (6 numbers). What do you learn from these numbers and how does it relate to assumptions we make for the analysis?
2. Run a 2-way ANOVA to test for the main effects and interaction and write a short report of the results with the relevant effect sizes (up to 130 words).
3. Create an informative plot that allows understanding of the results (use informative factor names and labels!). Can the researchers reject the null hypothesis described above? Describe the meaning of those findings in your terms (up to 130 words).
4. Assuming that the p-values are valid, what is missing from the current analysis to conclude that the findings you described are statistically significant?
5. Check the two main effects with two 1-way ANOVAs, each time addressing only one factor. Describe shortly the differences from the results of 2-way ANOVA and the reasons for them. (NOTE: this is done for the goal of demonstration. In real life, there is no logic in re-analyzing 2-way ANOVA with 1-way ANOVA)

**Note**: Please save a csv version of the data after your changes – we will keep working on it in the next exercise!