**Instructions - reminder**

-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. **Use no more than 75 words for each answer (3-4 lines in word with font 11).** We will not count answers over the word count. Answers that do not answer the question directly or that include a lot of unnecessary information will be marked as wrong.

**Part 1 (70 points)**

Load the dataset “ex4\_data.csv”. The data was already cleaned and ready for analysis. It includes measurements from subjects who gave a speech in front of a crowd (each row is one subject):

***PS\_score*** – a rating of an external judge (0-100) of how convincing the participant’s speech was.  
***Personality*** – a categorical variable describing each participant as “extroverted”/”introverted”  
***Caffeine*** – the amount of caffeine measured in the blood of the participants.  
***Arousal*** – the average change in the pupil diameter of the subject during the speech, compared to a baseline level.  
***Age*** – subjects’ age

1. The first hypothesis you will test is that personality modulates the effect of arousal. Specifically, we hypothesize that while controlling for the effect of age, the relationship between caffeine level and public speaking (1) is positive on average, and (2) stronger for extraverted subjects than for introverted people. (50 points)   
   1. In this part we will learn how to conduct a power analysis for multiple regression model. For that, please read the appendix at the end.  
      Here, you are asked to **perform a power analysis for the crucial interaction effect**. Thus, your null model should be the same model you would use for testing the hypothesis but without the modulation of the caffeine effect by personality. Assume that the R^2 of the null model is 0.18 and of the alternative model is 0.24.

What is the N needed for a power of 80%?

* 1. Perform the necessary transformations/coding to the categorical and continuous variables. Use a coding scheme where coefficients reflect the difference between levels of categorical variables and the grand mean when the continuous variables are set to 0.
  2. Run the relevant multiple regression and report it in a table. Include the estimate, t, score, df, and p-value for each predictor. Also report adjusted and unadjusted R2 for the model.
  3. Print the VIF score of each predictor in the model. Should you be concerned about multicollinearity? Why? (Briefly - up to 30 words)
  4. Explain briefly the influences of your predictors on public speaking. Add a measure for the size of the hypothesized effects described above (in adjusted R2).
  5. Assuming that the effect of age, by itself, is not theoretically interesting, why did we add it to the model? Suggest two possible reasons.
  6. Create an informative graph visualizing the model predictions (using ggpredict as shown in the tutorial). The graph should be focused on the two effects that were hypothesized, while considering the controlled variable (so if there was an effect, the graph should make it visible).

1. (20 points) A colleague interested in the effect of caffeine (alone, regardless of personality and age) suggested that the reason caffeine influenced public speaking was because it was confounded with arousal. That is, according to your colleague, caffeine did not directly influence public speaking, but rather through its indirect effect on arousal.
   1. Run the relevant models to test for this hypothesis and report the relevant measures for this hypothesis (including the direct and indirect effect, and the coefficients of the model/s you have used), including the relevant estimates and p-values.
   2. Was the colleague correct? Briefly explain the way in which caffeine influences performance.

**­Part 2 (30 points)**

In this part you will analyze data from a single subject performing The Muller-Lyer task. The subject saw a target line (with six different lengths) with arrows pointing inward (>---<) or outward (<--->) and was tasked with identifying whether the line is longer or shorter than a reference line.

The relevant columns are:

***Left arrow*** – whether the arrows pointed inwards or outwards.  
***length diff*** – The length of the target line minus the length of the reference line (positive value indicate that the target line was longer)  
***response*** – 1 for “target line longer”, 0 for “target line shorter”  
***RT*** - response time

1. Preprocessing:

* Load the “Muller\_Lyer\_S2.csv” file.
* Multiply length\_diff by 100.
* Use effect-coding for the “left arrow” variable.
* Subset the data to clean for RTs which were more than 2.5 SDs below or above the overall mean. How many trials are left?

1. Run the relevant analysis to test the hypothesis that the probability to identify the target line as longer is modified by the difference in lengths and the arrows direction (without interaction). Report the coefficients of the regression, their SE, and p-values, alongside the odds ratio for each predictor.
2. Explain the meaning of your results, including the meaning of the odds ratio.
3. Visualize the results you got. You can use the predictions of the model with ggpredict or the results themselves with ggplot(…)+geom\_smooth(…) or any other plotting function you want, as long as you make sure the graphs are understandable and coherent with complete labels.

**Appendix**

In essence, this power analysis is based on Cohens’ f2 (using the pwr.f2.test function) – an effect size based on the change in explained variance. f2 is calculated using the hypothesized (unadjusted) R^2 of two models – the null model **R2A**, including all the relevant predictors except those the hypothesis is focused on, and **R2AB**, an alternative model that includes all predictors. Cohen’s f2 is then calculated as the change in the ***unadjusted*** R2 (the regular, uncorrected explained variance) divided by the unexplained variance:  
Text

Description automatically generated

According to Cohen’s (1988) guidelines, f2≥ 0.02, f2≥ 0.15, and f2 ≥ 0.35 represent small, medium, and large effect sizes, respectively. The distribution of f2 for a given set of degrees of freedom is known, and therefore it is possible to calculate the probability to get a significant result.  
The degrees of freedom are **u** = the number of predictors added (1 for adding one factor, 2 for one factor and interaction, 3 for two factors and their interaction… etc.) and . N is the number of subjects. If you are interested in more details on how to conduct a power analysis given a hypothesized effect size f2,  [please read the tutorial here](https://dstanley4.github.io/psyc4780bookdown/sample-size-for-multiple-regression.html).