

Assignment 1

The assignments are completed (mainly) outside of class in groups of three and will be graded. It is advised that you complete the assignments in the same group as the exercise.

For the assignment, you turn in (1) the completed excel file, (2) your Rmarkdown file (.Rmd), and (3) the pdf or word document created by Rmarkdown containing your organized code and answers to the questions posed. The assignment will be turned in by uploading these 3 files to Blackboard before the next lecture. The titles of the files should include each group members last name and assignment name (e.g., ArtsGrandfieldAssign1.Rmd).

Note 1: In the assignment, please include a sentence stating each student's contribution towards the end product (e.g., which student completed what tasks). For collaborative efforts, tasks can be repeated over multiple students. In the extreme case of very diverging unique contributions, the coordinator has the possibility to differentiate the grading over students within a group.

Note 2: The Rmarkdown file should include the group members names, the necessary organized code (no output) and nicely formulated answers to each of the questions.

Assignment 1 Details

The file exam.sav holds data from pupils in English Schools. The dependent variable is exam-scores, and the independent variables are reading ability (LRT, London Reading Test) and average reading ability in school (AvsLRT).

Answer the questions given below regarding the exam.sav dataset. All analysis results (e.g., regression coefficients and their statistical significance) should be interpreted and reported in APA style. Please use R to perform the multilevel analyses, and - in addition to the answers to the questions -also hand in the filled in summary excel sheet to provide an overview of the results of all analyses conducted.

1. Why should, in theory, a multilevel analysis be performed on these data?
2. Provide descriptive statistics and check for outliers (don't perform analyses, just look in the scatterplots).
3. Answer the following questions
 - a. Should you perform a multilevel analysis or not?
 - b. What are the null- and alternative hypotheses?
 - c. What are the separate level 1 and 2 model equations, and the mixed model model equation?
 - d. Provide and interpret the relevant results (don't just copy the output, report the relevant results in APA style).

- e. What is the intraclass correlation?
 - f. What is your conclusion considering the overall question regarding the necessity of performing a multilevel analysis?
4. Which variables are predictors of *exam-score*?
Provide and interpret the relevant results and provide your overall conclusion.
5. Is the relation between *reading-score* and *exam-score* the same in all schools?
- a. What are the null- and alternative hypotheses?
 - b. Provide and interpret the relevant results.
 - c. Provide an overall conclusion.
6. Let's take a closer look at the regression coefficient of the *reading-score* obtained in the model under question 5.
- a. Construct a 95% confidence interval for the regression coefficient of *reading-score*
 - b. Construct the 95% predictive interval for the regression coefficients of *reading-score*
 - c. Explain the different meanings of 6b. and 6c.
7. Can *average reading-score* explain (a part of) the different relations between *reading-score* and *exam-score* in different schools?
Provide and interpret the relevant results and provide your overall conclusion.
8. Choose a final model and provide the separate level 1 and 2 model equations, as well as the mixed model equation.
9. Check the normality assumption for the level 1 residuals.
- a. In R: The level 1 and 2 residuals can be accessed via the function `residuals(object, ...)`. Check the help file for more information and note the argument `level` and test what it does!
 - b. Using the obtained residuals, (visually) inspect the normality assumption.