

Theoretical and Methodological Foundations of Visual Computing (Assignment 13)

Please submit your solution to ILIAS by 2025-02-02, 23:55. Solutions will be discussed in the tutorial on Thursday, 2025-02-06. Please note that source code will only be accepted if it compiles.

Exercise 13.1 (10 points)

You did an online study, within-subjects, examining how fast people can interpret data using different visualization types. Additionally, you wanted to know if color would have an impact. Hence, the independent variables are color and visualization type. The dependent variable, affected by color and visualization type, is reaction time.

Each participant was anonymised with a randomly assigned ID number ("ID" in the dataset). The participants tested all the combinations of color and visualization type four times.

The raw results of your study are in the 'vis_type.csv' file. You are encouraged to use the "help" tab in RStudio or Google to solve the questions.

- Import the data into RStudio and inspect the data using `summary()`. You'll notice that ID is numeric and the independent variables are characters. In order to do analysis with the data you have to convert the independent variables and the participants' ID into factors. Convert these variables to factors.
- Which visualization type leads to the lowest reaction time? Which to the highest? Use summary statistics as well as visualizations to find out the answer.
- Which color leads to the lowest reaction time? Which to the highest? Use summary statistics as well as visualizations to find out the answer.
- Create a 95% confidence interval plot of each combination of color and visualization type. To do so, first you'll need to get the following summary statistics of the reaction time per combination of color and visualization type: mean, standard deviation, upper- and lower bound of the 95% confidence interval.
- Based on the summary statistics, which combination of color and visualization type leads to the lowest RT? Which to the highest RT?

Exercise 13.2 (10 points)

You continue your analysis on the data and investigate if there is a difference in reaction time between the visualization types and each color.

- (a) A two-way ANOVA would be the best way to analyse the data if the data meets the assumptions. Name these assumptions and use R to check them.
- (b) Which of the ANOVA assumption(s) is not met? What can you do in order to make the data meet this assumption? Perform your suggested solution(s) in R.
- (c) Does the data meet the ANOVA assumptions now? Use R to check and reason your answer.
- (d) Carry out the appropriate statistical test, either ANOVA or equivalent. What are the results?
- (e) Carry out a posthoc analysis too. Create a 95% confidence interval plot of the results.