QuantHCI Assignment C2: Analyzing data with linear models

Firstname Lastname
12-345-678
alice@example.com

Firstname Lastname
12-345-678
bob@example.com

1. ASSIGNMENT INSTRUCTION*

Form a team of 2–3 students. In this assignment, you will apply generalized linear model and generalized linear hypotheses to analyze a dataset from a study setup similar to P1 to answer two research questions (RQ) mentioned in the next two sections. Write up a short reproducible report (max. 3 pages using this template) as typically seen in the "Results" and "Discussion" section of academic papers. The given dataset is clean (equal number of trials, no missing data, no outliers). So, you can focus on model building and estimation right away.

As with typical data analysis tasks, this assignment is *deceptively* simple. It will also not take much time to complete the analysis. However, there are many pitfalls that will compromise your analysis, interpretation, and reasoning. Here are some hints (which you won't have in real data analysis):

- You will definitely need to apply knowledge from the first three linear model lectures (9 – 30 Apr.). You may proceed working on the assignment straight away. Then, after the lecture on 30 Apr., you may need to make minor modifications before handing in the results.
- You will also need to consider statistical assumptions and reexpression
- You will need to re-read page 252 of MacKenzie (2012) [1], which is provided with the assignment P1.

1.1 Deliverable

Hand in one zip file of the whole directory (as you have received this assignment). Make sure that the `report/report.pdf` is the version you intended to hand in. Hand in the assignment on OLAT "C2 Submission" Sunday 13 May 2018, 23:59. Create one submission per team.

1.2 Grading

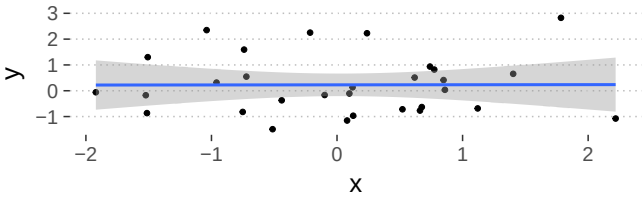

Figure 1: Example figure caption.

This part of assignment contributes to 10% of your final grade and will be evaluated in the following aspects:

- *Thoroughness* in the assessment of statistical assumptions
- *Correctness* of the code used in the analysis
- *Readability* of the code and the reasoning in the text
- *Completeness and conformity* of statistical reports, e.g., degrees of freedom, decimal points
- *Effective use* of statistical graphics to communicate your results
- *Reproducibility* of the code used to generate the report.

1.3 About this template

You can directly create reproductible PDF document conforming to ACM template from R Markdown. Check the package `rticles` for details. In fact, this PDF is generated from the files in the `report` folder.

During the data exploration, we recommend you to test your code in R Markdown Notebook (`.Rmd`) or simple R files (`.R`) because they are relatively quicker to run and debug. After you have a solid analysis, port your code into this template and use the **Knit** button to generate PDF.

Note how the example figures shown in Figure 1 use the package `ggpubr` to generate plots that are printer-friendly. Also note in the source code how to refer to figure with `\ref` and `\\label` in the figure caption. You may change the height of each figure (`fig.height`) as appropriate.

In the `{r setup ...}` chunk above (only visible in the source code), we also use `import::from` to selectively import functions from the packages. This prevents functions of similar names from crashing with each other. To use `import::from`, you'll need to install the `import` package.

You may want to selectively hide some code chunks with `echo=FALSE` (Show only output) or `include=FALSE` (Show nothing, but run code)

Sometimes, you can easily format the number with decimal points using the `format_num()` provided in code, such as this: 1.23. However, at other times (e.g., $p < .001$), formatting would be easier if you just type in the number.

The footer on the bottom-left of the page is determined in the template. Don't bother wrestling with it.

Please remove the instruction sections from your submission.

Change the title of the submission to a short clause that highlight your findings.

1.4 Dataset

Each group will use a unique subset of data from the given dataset. Use the following code block (visible in the source code) to set `group_seed` and generate `my_data` to generate the dataset for your group.

2. EFFECT OF INPUT TECHNIQUES

RQ1: To which extent does the choice of input techniques influence the pointing performance?

Describe your claim, evidence and reasoning. Make sure to justify your choice of analysis methods.

3. LEARNING EFFECT

RQ2: To which extent does the learning effect influences the pointing performance?

4. CONCLUSION

In this section, you will write about what you learned by drawing from both RQ1 and RQ2

(For grading `group_seed`: (write down the number of your `group_seed` here))

References

[1] MacKenzie, I.S. 2012. *Human-computer interaction: An empirical research perspective.* Newnes.