IDR4000 Portfolio exam - 2022

Overall requirements for the portfolio exam

Your report should be a completely reproducible document, built from a source file (e.g. R markdown) with data and code stored in a repository. A reference/link to your github repository should be provided in the report. To make your life easier, I suggest that you compile your report from one repository.

Part 1: Reliability and tools for reproducible data science

The purpose of this assignment is to present estimates of reliability of measures collected in the physiology lab. A second purpose is to use tools for reproducible data science. The report that you are expected to hand in therefore has some strict requirements in its format (see below). The assignment is a group assignment and at least three students are expected to contribute to each report.

Elements of the report

The report should describe one test that you have performed in the physiology-lab. Select the test that is most interesting to you. The test should be described with a detailed protocol, including preparations of the participant (that is being tested), standardization of the test, and post-test data preparation. Post-test data preparation refers to steps needed to get data from e.g. equipment used during the test. This section should take into account and reference [@RN1739; @RN2511]

The next section should contain descriptive data from the test. This could be measures of central tendency and variability in the measures you have collected. If possible, try to find similar estimates in the scientific literature.

Finally, we are interested in reliability. Here you need to calculate an estimate of reliability of the test. Use (and reference) [@RN130]. Try to be clear with what measure of reliability you are using and what it is telling you.

The format of the report

The report should be uploaded to github with both a source file (`.Rmd` or `.qmd`) and report file as output (`html`, `pdf` or `docx`-file). The github folder should also contain the dataset being used in the calculations. Work on your assignment as a R project. Contributions from members of the group can be made directly to the github repository.

Part 2: Laboratory report from DNA, RNA or protein extraction/analysis labs

This is a group collaboration. Each group is expected to present background, methods (including principles for measurements) and results from one of the labs (DNA, RNA or protein). You are free to shape the report as you see fit.

Part 3: Philosophy of science

This is an individual report. The report should be updated based on comments from Finnur. See a detailed description on Canvas.

Part 4: Study designs

Overview

Choose an area of interest (e.g. protein supplementation for muscle hypertrophy or the effect of block periodization on VO2max). Find at least five *original research studies*[1] in your selected area and describe strength and weakness of these studies (see below). The report should focus on the design of the studies and selection of statistical tests to answer study aims. Conclude your report with a recommendation, how should future studies in your area be designed to best answer similar questions?

The report should be handed in on canvas as a link to a github folder containing a reproducible report. This is an individual assignment.

Details

When **analyzing** your studies you can use the QALMRI method[2].

Note that the report **should not** contain your QALMRI table but should instead be focused on describing differences and similarities in all studies together (see also below)!

[1] Avoid using review articles or meta-analyses

[2] See Teaching undergraduate students to read empirical articles: An evaluation and revision of the QALMRI method, this advice was also heavily influenced by this website

The questions may help you analyze your studies

1. What was the broader problem the authors are trying to resolve in the study?
2. What are the specific questions the authors are trying to answer?

   The first point should be similar in all your studies, this could be e.g. the effect of age on physical functioning, effect of certain training protocols on VO2max, etc. The second point is potentially different between your studies.

Alternative explanations

3. Is the specific question framed as an hypothesis or a question?
4. If the authors have formulated a hypothesis, what alternative explanations can you think of that could potentially explain the data that the authors hypothesize?

Logic

5. What is the logic of the hypothesis or the question. Try to create a "line of logic" between the introduction and the question/hypothesis.

Methods

6. Describe the study design. Use Hulley, (2013), Chapters 7-13[3] in your analysis.
7. Describe the sample and if the study defines the population.
8. Describe the method of recruiting participants to the study. Did the authors justify their sample size (i.e. did they do a power calculation)?
9. Describe how the study was conducted (what tests was performed when etc.)
10. Describe the variables in the study, what variables relate to the question/hypothesis?
11. What methods did the authors use to make claims (what statistical tests were used)

Results

12. What were the main results of the study, did the authors answer their question/address their hypothesis?

Inference

13. What could the authors conclude from the study?
14. What did the authors conclude about the study population?

[3] Hulley, S. B. (2013). Designing clinical research. Philadelphia, Wolters Kluwer/Lippincott Williams & Wilkins.

Performing your literature review and writing the report

Your report should not contain a detailed summary of all studies for all these questions, instead you should summarize your results. Highlight differences and similarities between studies. *As the main point of this review is study designs and statistical analyses, this should be your main focus.* When doing your literature review, it is however a good idea to structure it in a table with the above mentioned headings:

- Question
- Alternative
- Logic
- Methods
- Results
- Inference

Part 5: Analyzing repeated measures experiments

In this assignment you will analyze and report on trial investigating the effect of resistance training volume on lean mass and muscle strength. The data are part of the `exscidata` package and can be accessed as `data("strengthvolume")` and `data("dxadata")`. Read the instructions carefully!

Organizing the report

Your report should consist of the sections Introduction, Methods, Results and Discussion. Each part of the report should be written as a reproducible document and a link or reference to the repository containing the source document(s) should be included in the report. Below follows detailed descriptions and requirements for each section.

Introduction

This section should consist of a description of the field, resistance-training volume and muscle strength and mass. Use at least five to ten references to introduce your audience and explain why you are doing the analysis/study. A tip is to use the QALMRI method, introduced in Assignment 4 to structure the reading of background information. It is up to you how you motivate the study and how you phrase the purpose of the study. It could be a hypothesis based on previous studies, it could also be question to fill a knowledge gap that you have identified in your literature review.

Structure the introduction in paragraphs. A first paragraph could contain a general introduction to the field, why is it of interest to investigate resistance-training? A second paragraph could specifically describe the specific field of resistance-training volume, why is important to

know more about how we are likely to respond to different training volumes. The second paragraph should incorporate definitions important for your report, e.g., training volume, muscle mass and strength. Try to incorporate these definition as a fluid part of the text.

A third (or last) paragraph of the introduction should contain a statement regarding the purpose of the study. The purpose could be descriptive, hypothesis-driven or guided by a question. Although it could be considered a bit backward, you should explore the data sets before you select your question/hypothesis/purpose for it to be possible to answer.

Methods

The method should give a thorough overview of the study and specific details regarding data collection. You can read about the details of this specific study in [@RN2358]. Use your own words to describe the study based on this description. A nice way to structure the methods section is to include subheadings:

- Participants and study overview: Describe the participants and give an overview of all tests/measurements. Participants should be described in the first table of the report (Table 1). The overview of the tests/measurements should be done without double presentation as details should be presented in subsequent sections.
- Specific descriptions (e.g. strength tests): Describe in detail how tests/measurements that you mentioned in the overview where conducted.
- Data analysis and statistics: Describe how you treated the data prior to statistical tests or procedures and what tests/procedures were used to draw inference (or more generally, to answer your purpose). Describe how you present data (e.g. descriptive data with mean (SD), inference with confidence intervals etc.).

Results

Describe the results of your analysis. This description should make use of table and figures as well as a text that guides and structures the content to the reader. Think about it this way, the text should describe when and how to read the figures and tables. This means that all aspects of the results should be covered in the text. The figures and tables should also be "self explanatory", this means that you have to include descriptive figure captions and descriptions of tables (see below for tips).

As the main purpose of the analysis should concern the effect of training volume on muscle mass and strength, it is natural that the comparison of training outcomes between volume conditions is the main analysis in the results. You may also have questions regarding the relationship between muscle strength and mass gains, if there are differences between men and women etc. Selection of statistical/analysis techniques should reflect the study question/purpose.

Discussion

Structure the discussion with a first paragraph describing the main results of the analysis, this could be the answer to your question or a statement regarding the study hypothesis. In the following paragraphs discuss all results that you have presented in the light of previous studies. It is your job to give the reader plausible interpretations and explanations of your results. This is how single scientific results are incorporated in our collective understanding. These interpretations can later be challenged, however if you give the reader good arguments and clear descriptions, your insights will be valuable to collective reasoning even if they turn out to be wrong in light of new data.

End the discussion with a summary or conclusion. Some journals request that you state your conclusions under a specific heading in the end of the report/article.

Organizing the data analysis

Data preparation

The data is already structured in the `exscidata` package. To access the data, use the following code:

```
library(exscidata)
data("dxadata"); data("strengthvolume")
```

To get and overview of the variables in each data set use `?strengthvolume` and `?dxadata`. In the `dxadata` the variables of interest are organized in a more convenient way using the code below:

```
library(tidyverse)

dxadata %>%
  select(participant:include, lean.left_leg, lean.right_leg) %>%
  pivot_longer(names_to = "leg",
               values_to = "lean.mass",
               cols = lean.left_leg:lean.right_leg) %>%
  mutate(leg = if_else(leg == "lean.left_leg", "L", "R"),
         sets = if_else(multiple == leg, "multiple", "single")) %>%
  select(participant, time, sex, include, sets, leg, lean.mass) %>%
  print()
```

```
# A tibble: 160 x 7
   participant time  sex    include sets     leg   lean.mass
   <chr>       <chr> <chr>  <chr>   <chr>    <chr>     <dbl>
```

```
 1 FP28        pre   female incl    multiple L         7059
 2 FP28        pre   female incl    single   R         7104
 3 FP40        pre   female incl    single   L         7190
 4 FP40        pre   female incl    multiple R         7506
 5 FP21        pre   male   incl    single   L        10281
 6 FP21        pre   male   incl    multiple R        10200
 7 FP34        pre   female incl    single   L         6014
 8 FP34        pre   female incl    multiple R         6009
 9 FP23        pre   male   incl    single   L         8242
10 FP23        pre   male   incl    multiple R         8685
# ... with 150 more rows
```