Mappeeksamen IDR4000

Student Studentson

2024-09-10

Table of contents

Introduksjon til mappetemplat og Lærinsgrefleksjon

Emnet IDR4000, kvantitativ metode og statistikk bruker mappeeksamen som evalueringsform. Dette er et "templat" for skriving og innlevering av eksamen. Lag en egen versjon av templatet ved å laste ned det (fork) fra GitHub og bruk denne som utgangspunkt. Ta vekk alle instruksjoner før du leverer den siste versjonen av din mappe. Du kan levere den samme mappen med aktuelle deler oppdaterte i løpet av emnet.

Mappeeksamen består av følgende deler:

- Rapport: "Deskriptiv statistikk, reliabilitet og validitet og verktøy for reproduserbar vitenskap".
- Laborasjonsrapport fra molekylærlabb
- Arbeidskrav i vitenskapsteori
- Rapport: "Statistisk inferens, statistiske modeller og statistisk styrke"
- Rapport: "Studiedesign"
- Rapport: "Analyse av eksperimenter med repeterte målinger"

I templatet organiseres hver del som et kapittel.

Referanser finner du sist i dokumenetet (eks. (Spiegelhalter 2019))

1 Assignment 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 assignment description). The assignment is a group assignment and at least three students are expected to contribute to each report.

1.1 Elements of the report

Importantly, the report should contain:

- At least one table (created from your data)
- At least one figure (created from your data), and
- data presented in the text.
- The report should use a bibliography file to manage references.

2 Assignment 2: Regression models, predicting from data

The assignment has three parts:

- Part 1: Lactate thresholds
- Part 2: Predicting sizes of DNA fragments
- Part 3: Interpreting a regression table

3 Assignment 3: Drawing inference from statistical models, and statistical power

This assignment is set up as a statistical laboratory, we will perform simulations and your assignment is to interpret and explain the results. Create a report based on the code used in the lab and make sure you answer the specified questions (1-8). You can be as creative as you want and explore the results further.

4 Assignment 4: Study designs

4.1 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. 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?

[1] Avoid using review articles or meta-analyses

5 Assignment 5: Analyzing repeated measures experiments

5.1 Assignment overview

In this assignment you will analyse 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!

> Below you will find a basic outline of the report and example code that we worked on in class.

5.2 Introduction

5.3 Methods

5.3.1 Participants and study overview

5.3.2 Muscle strength and hypertrophy

5.3.3 Data analysis and statistics

5.4 Results

The average difference in lean mass changes between sets were 122.8, 95% CI: [8.6, 237], p = 0.036.

```
## Time points in strength data set

strengthvolume %>%
  distinct(exercise)
```

```
# A tibble: 6 x 1
  exercise
  <chr>
1 legpress
2 legext
3 isok.60
4 isok.120
5 isok.240
6 isom
```

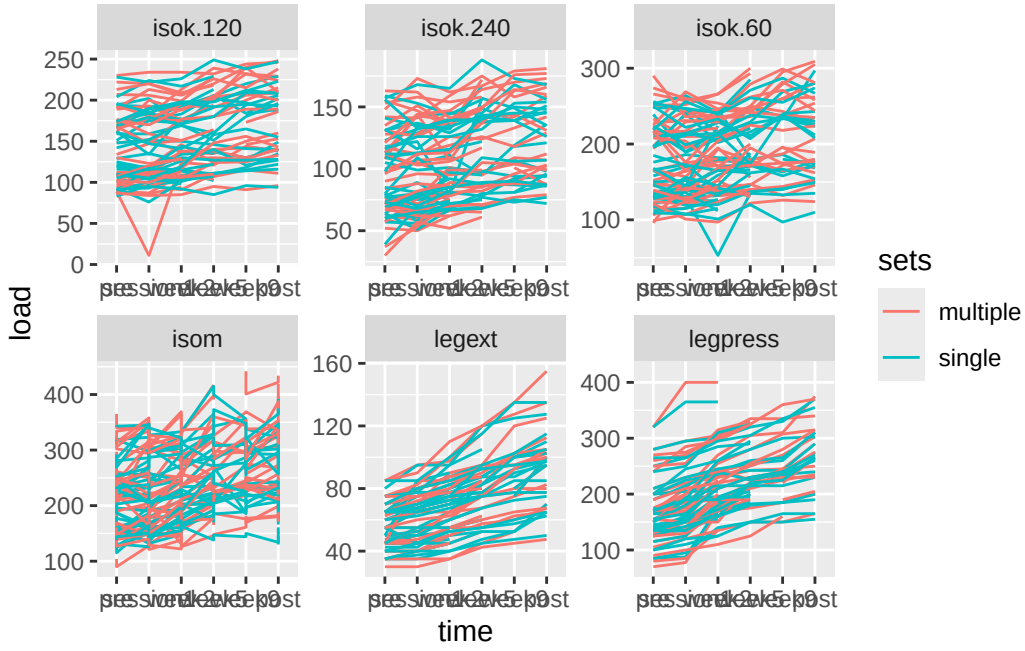
```r
## Exploratory plot of strength data

str <- strengthvolume %>%
  filter(include == "incl") %>%
  mutate(time = factor(time, levels = c("pre", "session1",
                                        "week2", "week5",
                                        "week9", "post"))) %>%
  print()
```

```
# A tibble: 2,856 x 8
   participant sex   include time     sets     leg   exercise  load
   <chr>       <chr> <chr>   <fct>    <chr>    <chr> <chr>    <dbl>
 1 FP13        male  incl    pre      single   R     legpress   115
 2 FP13        male  incl    pre      multiple L     legpress   115
 3 FP13        male  incl    pre      single   R     legext      55
 4 FP13        male  incl    pre      multiple L     legext      55
 5 FP13        male  incl    session1 single   R     legpress   125
 6 FP13        male  incl    session1 multiple L     legpress   125
 7 FP13        male  incl    session1 single   R     legext      55
 8 FP13        male  incl    session1 multiple L     legext      55
 9 FP13        male  incl    week2    single   R     legpress   185
10 FP13        male  incl    week2    multiple L     legpress   175
# i 2,846 more rows
```

```r
str %>%
  ggplot(aes(time,
             load,
             group = paste(participant, sets),
             color = sets)) +
  geom_line() +
  facet_wrap(~ exercise, scales = "free")
```

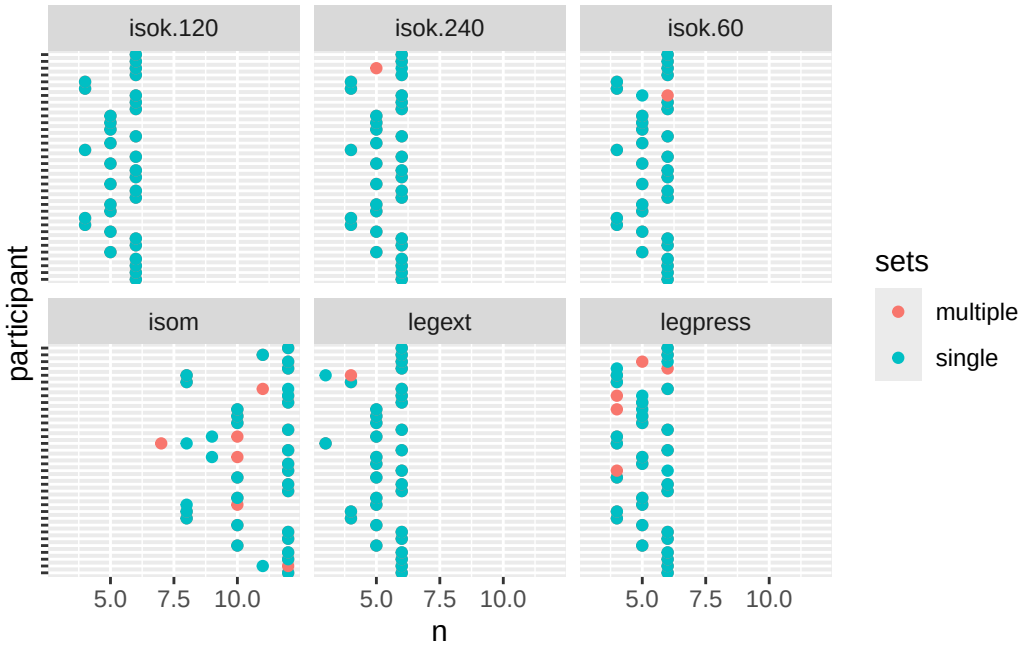
```
Warning: Removed 5 rows containing missing values or values outside the scale range
(`geom_line()`).
```




```
## How many measurements per participant

str %>%
  filter(!is.na(load)) %>%
  group_by(participant, exercise, sets) %>%
  summarise(n = n() ) %>%
  ggplot(aes(n, participant, color = sets)) +
  geom_point() +
  facet_wrap(~ exercise) +
  theme(axis.text.y = element_blank())
```

```
`summarise()` has grouped output by 'participant', 'exercise'. You can override
using the `.groups` argument.
```




```
## Use pre and post data
# Combine pre data prior to data analysis
# per exercise, leg, participant, and sets

str %>%
  mutate(time = if_else(time %in% c("pre", "session1"), "pre", time)) %>%

  filter(time %in% c("pre", "post")) %>%

  summarise(load = max(load, na.rm = TRUE),
            .by = c(participant,
                    sex,
                    time,
                    sets,
                    exercise,
                    leg)) %>%

  print()
```

```
Warning: There were 7 warnings in `summarise()`.
The first warning was:
i In argument: `load = max(load, na.rm = TRUE)`.
```

```
i In group 62: `participant = "FP6"`, `sex = "female"`, `time = "post"`, `sets
  = "multiple"`, `exercise = "legpress"`, `leg = "L"`.
Caused by warning in `max()`:
! no non-missing arguments to max; returning -Inf
i Run `dplyr::last_dplyr_warnings()` to see the 6 remaining warnings.
```

```
# A tibble: 816 x 7
   participant sex    time  sets     exercise leg    load
   <chr>       <chr>  <chr> <chr>    <chr>    <chr> <dbl>
 1 FP13        male   pre   single   legpress R     125
 2 FP13        male   pre   multiple legpress L     125
 3 FP13        male   pre   single   legext   R      55
 4 FP13        male   pre   multiple legext   L      55
 5 FP13        male   post  single   legpress R     230
 6 FP13        male   post  multiple legpress L     235
 7 FP13        male   post  single   legext   R      97.5
 8 FP13        male   post  multiple legext   L     100
 9 FP16        female pre   single   legpress R      95
10 FP16        female pre   multiple legpress L      85
# i 806 more rows
```

5.5 Discussion

5.6 Conclusion

6 Philosophy of science

See instructions on canvas.

7 Molecular Laboratory report

Select one laboratory assignment and write a detailed report.

References

Spiegelhalter, D. J. 2019. *The Art of Statistics : How to Learn from Data*. Book. First US edition. New York: Basic Books.