

Mappeeksamen IDR4000

Ole-Kristian Kjølnør

2024-09-10

Table of contents

Introduksjon	3
1 Assignment 1: Reliability and tools for reproducible data science	4
1.1 Detailed protocol	4
1.1.1 Metode	4
1.1.2 Utstyr som ble benyttet	4
1.1.3 Preparations of participant being tested	5
1.2 Standardization of test	5
1.3 Post test data preperation	6
1.3.1 Typical error	9
1.3.2 referanse Hopkins	10
1.3.3 Data i tekst	10
2 Assignment 2: Regression models, predicting from data	11
3 Assignment 3: Drawing inference from statistical models, and statistical power	12
4 Assignment 4: Study designs	13
4.1 Overview	13
5 Assignment 5: Analyzing repeated measures experiments	14
5.1 Assignment overview	14
5.2 Introduction	14
5.3 Methods	14
5.3.1 Participants and study overview	14
5.3.2 Muscle strength and hypertrophy	14
5.3.3 Data analysis and statistics	14
5.4 Results	14
5.5 Discussion	18
5.6 Conclusion	18
6 Philosophy of science	19
7 Molecular Laboratory report	20
References	21

Introduksjon

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”

1 Assignment 1: Reliability and tools for reproducible data science

1.1 Detailed protocol

1.1.1 Metode

1.1.2 Utstyr som ble benyttet

- Biosen (for måling av laktat)
 - lansett finger
 - rør for oppsamling av blod
 - blandingscontainer for rør med blod
 - papir for å tørke av finger
- Vyntus (for måling av metabolsk respons)
 - miksekammer
 - slange fra miksekammer til munnstykke
 - neseklype (teipbit på nese for at neseklype skulle sitte godt)
 - gassbeholder med standardgass
- Lode-sykkel
 - Samme oppsett på sykkel ved hver test (høyde sete, avtsand sete til styre, osv.). Oppsettet lagres på data etter innstilling ved t1.
 - MTB pedaler
 - 172,5 mm lengde krankarmer
- Gulvvifte (samme innstilling ved hver test)
- Pulsbelte garmin
- Hansker

1.1.3 Preparations of participant being tested

- Informasjon fra testansvarlig til testdeltaker vedrørende hvordan test skal gjennomføres:
 - VO2max test der vi starter på en gitt watt (motstand). Watten vil øke hvert minutt (hvor mye avhenger av protokoll) og testdeltaker sykler helt til han/hun ikke greier mer.
 - Test avsluttes dersom tråkkfrekvens < 60 rpm.
 - Test skal gjennomføres sittende
 - Borg skala benyttes når utmattelse inntreffer hos testdeltaker. Deltaker vil bli spurt om å gi et tall fra 6-20 på hvor sliten han/hun er, der 6 er “ingen anstrengelse” og 20 er “maksimal anstrengelse”.
 - Underveis i test vil testdeltaker ha tilgang på hvor lenge han/hun har syklet.
 - Testdeltaker informeres om at testleder kun vil informere om hvor lenge det er til neste økning og hvilken watt som sykles på for øyeblikket. Annen data som puls, VO2, osv., vil ikke utøver ha tilgang på underveis i test.
 - Testleder vil mot slutten av test bidra til å pushe testdeltaker slik at han/hun får ut sitt ytterste. Hvordan denne “pushingen” gjennomføres vil variere fra testleder til testleder. Det viktige er at testleder oppfører seg tilnærmet likt ved hver test ovenfor den aktuelle testdeltaker.

1.2 Standardization of test

- Matinntak:
 - siste store måltid senest 2 timer før teststart. Tillatt å spise en mindre karbohydratkilde senest 30 min før test, dvs. en banan, en bar, en gel, osv.
- Koffeininntak:
 - inntak som normalt.
- Test gjennomføres på samme tidspunkt på dagen ved alle tester
- Søvn/døgnrytme:
 - forsøk å opprettholde normal døgnrytme gjennom testperioden.
- Trening:
 - ingen hard trening på underkropp dagen før test.
- Oppvarmingsprotokoll:
 - 5 min oppvarming (2 min 150W, 2 min 175W, 1 min 220W)
- Samme testprotokoll ved alle testene:
 - start på 200W - 20W økning hvert minutt til utmattelse.

1.3 Post test data preperation

Data er hentet ut og formatert i excel. Deretter benyttet vi read_xlsx for å importere datafilen til R.

```
df %>%
  # velger de variablene jeg ønsker å ha med fra datasettet
  select(id, timepoint, weight, w.max, vo2.max, vco2.max, rer.max, ve.max, bf.max, hr.max, la.max)
  # hr.max og la.max var <chr>, så gjør om de til <dbl>

  mutate(hr.max = as.numeric(hr.max)) %>%
  mutate(la.max = as.numeric(la.max)) %>%

  #formaterer tbl slik at karakterene weight:borg.max samles i én kolonne og tilhørende verdier
  pivot_longer(names_to = "variables",
               values_to = "values",
               cols = weight:borg.max) %>%

  # grupperer etter id, timepoint og variables
  group_by(id, timepoint, variables) %>%

  # slår sammen karakterene id og timepoint
  mutate(id_timep = paste(id, timepoint, sep = "_")) %>%

  # fjerner gruppering
  ungroup() %>%

  # fjerner id og timepoint siden de nå er slått sammen til id_timep
  select(-id, -timepoint) %>%

  #ordner tabell så variabler weight:borg.max havner lengst til venstre og id_timep får hver sin kolonne
  pivot_wider(names_from = id_timep,
              values_from = values) %>%

  # printer tabell i "viewer"
  gt() %>%

  # lager egen kolonneoverskrift for å sortere de ulike id'ene
  tab_spanner(label = "id13", columns = c("id13_t1", "id13_t2", "id13_t3")) %>%

  tab_spanner(label = "id14", columns = c("id14_t1", "id14_t2", "id14_t3", "id14_t4")) %>%
```

```

tab_spanner(label = "id15", columns = c("id15_t1", "id15_t2", "id15_t3")) %>%

tab_spanner(label = "id16", columns = c("id16_t1", "id16_t2")) %>%

#endrer navn på kolonnene
cols_label(variables = " ",
            id13_t1 = "t1",
            id13_t2 = "t2",
            id13_t3 = "t3",
            id14_t1 = "t1",
            id14_t2 = "t2",
            id14_t3 = "t3",
            id14_t4 = "t4",
            id15_t1 = "t1",
            id15_t2 = "t2",
            id15_t3 = "t3",
            id16_t1 = "t1",
            id16_t2 = "t2") %>%

fmt_number(decimals = 2) %>%

tab_options(table.font.size = 9)

```

Table 1.1: Data from reliability tests sorted by id and timepoint

	id13			id14				id15			id16	
	t1	t2	t3	t1	t2	t3	t4	t1	t2	t3	t1	t2
weight	88.80	90.10	90.00	68.50	68.80	69.20	69.00	103.90	103.60	101.30	82.10	82.80
w.max	341.00	344.33	354.33	385.33	380.00	371.67	381.00	305.00	280.33	293.33	383.67	379.33
vo2.max	4,634.50	4,606.50	4,540.50	4,556.50	4,561.50	4,437.00	4,545.00	4,350.50	3,818.50	4,134.00	4,872.00	5,050.00
vco2.max	5,453.00	5,312.00	5,324.50	5,129.00	5,071.00	4,991.50	5,180.50	5,034.00	4,471.00	4,785.00	5,276.50	5,561.00
rer.max	1.17	1.15	1.17	1.12	1.11	1.12	1.14	1.15	1.17	1.16	1.08	1.10
ve.max	201.50	182.50	176.00	176.50	187.50	164.00	182.00	206.00	196.00	194.00	192.00	190.50
bf.max	54.80	49.60	48.35	48.95	69.55	43.35	48.35	47.30	63.85	53.05	80.10	68.80
hr.max	194.00	194.00	194.00	191.00	188.00	187.00	189.00	183.00	180.00	184.00	187.00	181.00
la.max	15.43	12.97	14.77	15.78	16.16	13.93	15.37	9.25	11.10	12.62	8.97	9.07
borg.max	19.00	18.00	18.00	19.00	19.00	18.00	19.00	17.00	16.00	18.00	19.00	20.00

```

df %>%
  select(id, timepoint, weight, w.max:borg.max) %>%

  mutate(rel.vo2max = vo2.max / weight) %>%
  mutate(rel.wmax = w.max / weight) %>%

  pivot_longer(names_to = "names",

```

```

      values_to = "tall",
      cols = w.max:rel.wmax) %>%

group_by(id, names) %>%

summarise(m = mean(tall),
          s = sd(tall)) %>%

mutate(m_s = paste0(signif(m, 3), " (", signif(s, 2), ")")) %>%

ungroup() %>%

select(-m, -s,) %>%

pivot_wider(names_from = "id",
            values_from = m_s) %>%

gt() %>%

cols_label(names = " ")

```

Table 1.2: Table sorted by id. All numbers are mean (sd)

	id13	id14	id15	id16
bf.max	50.9 (3.4)	52.6 (12)	54.7 (8.4)	74.4 (8)
borg.max	18.3 (0.58)	18.8 (0.5)	17 (1)	19.5 (0.71)
hr.max	194 (0)	189 (1.7)	182 (2.1)	184 (4.2)
la.max	14.4 (1.3)	15.3 (0.97)	11 (1.7)	9.02 (0.071)
rel.vo2max	51.3 (0.88)	65.7 (1.1)	39.8 (2.6)	60.2 (1.2)
rel.wmax	3.87 (0.062)	5.51 (0.1)	2.85 (0.12)	4.63 (0.065)
rer.max	1.16 (0.013)	1.13 (0.01)	1.16 (0.01)	1.09 (0.011)
vco2.max	5360 (78)	5090 (81)	4760 (280)	5420 (200)
ve.max	187 (13)	178 (10)	199 (6.4)	191 (1.1)
vo2.max	4590 (48)	4520 (59)	4100 (270)	4960 (130)
w.max	347 (6.9)	380 (5.7)	293 (12)	382 (3.1)

```

ggplot(data = df,
       aes(weight,
           w.max,
           color = id,

```



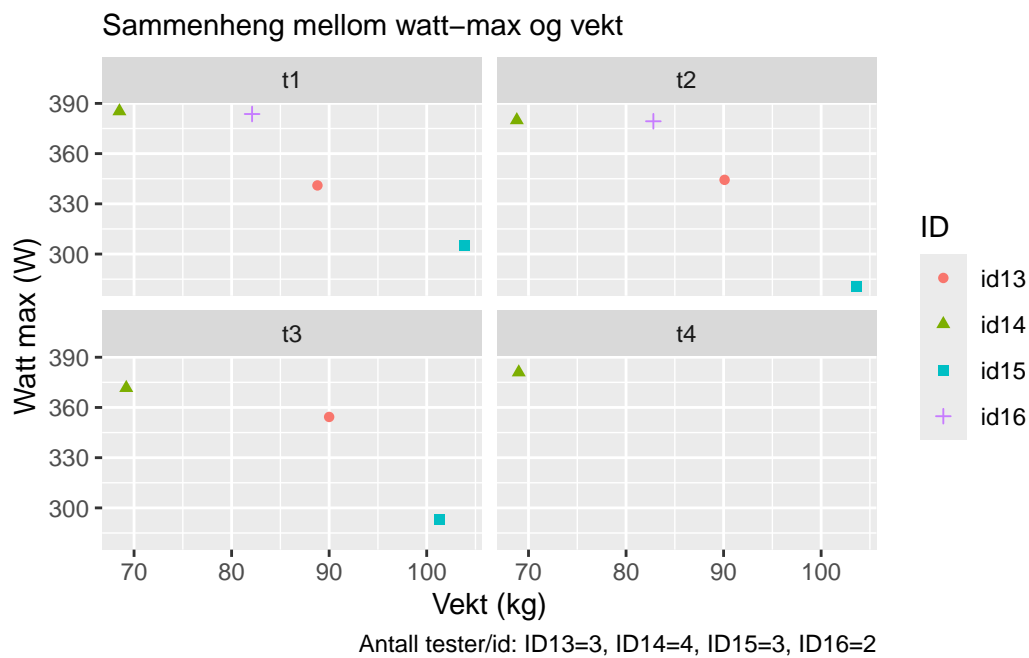
```

    shape = id)) +
  geom_point() +

  facet_wrap(~timepoint) +

  labs(x = "Vekt (kg)",
       y = "Watt max (W)",
       color = "ID",
       shape = "ID",
       subtitle = "Sammenheng mellom watt-max og vekt",
       caption = "Antall tester/id: ID13=3, ID14=4, ID15=3, ID16=2")

```



1.3.1 Typical error

```

df |>
  select(id, timepoint, la.max) |>

  pivot_wider(names_from = timepoint,
              values_from = la.max) |>

```

```
mutate(diff = t2 - t1) |>

summarise(m = mean(c(t1, t2), na.rm = TRUE),
          s = sd(diff, na.rm = TRUE),
          te = s / sqrt(2),
          cv = 100 * (te/m)) |>

print()
```

```
# A tibble: 1 x 4
      m      s    te    cv
  <dbl> <dbl> <dbl> <dbl>
1  12.3  1.79  1.27  10.3
```

1.3.2 referanse Hopkins

(Hopkins 2000)

1.3.3 Data i tekst

to add data in text we can use data fra environment%

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?

¹Avoid 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
```

```
## Exploratory plot of strength data
```

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

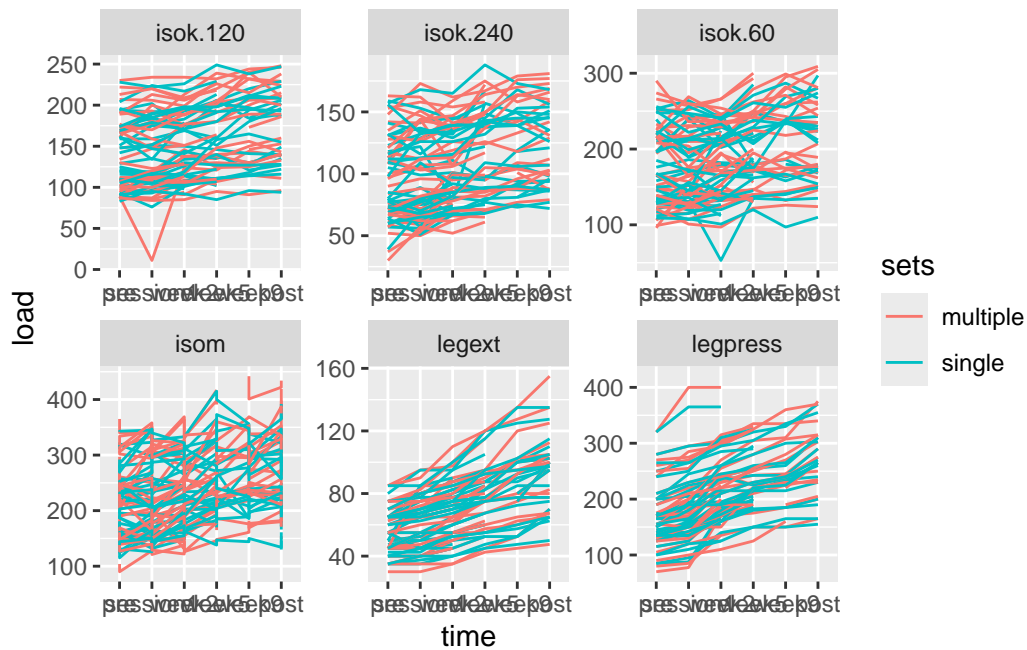
```
# A tibble: 2,856 x 8
```

	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
```

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

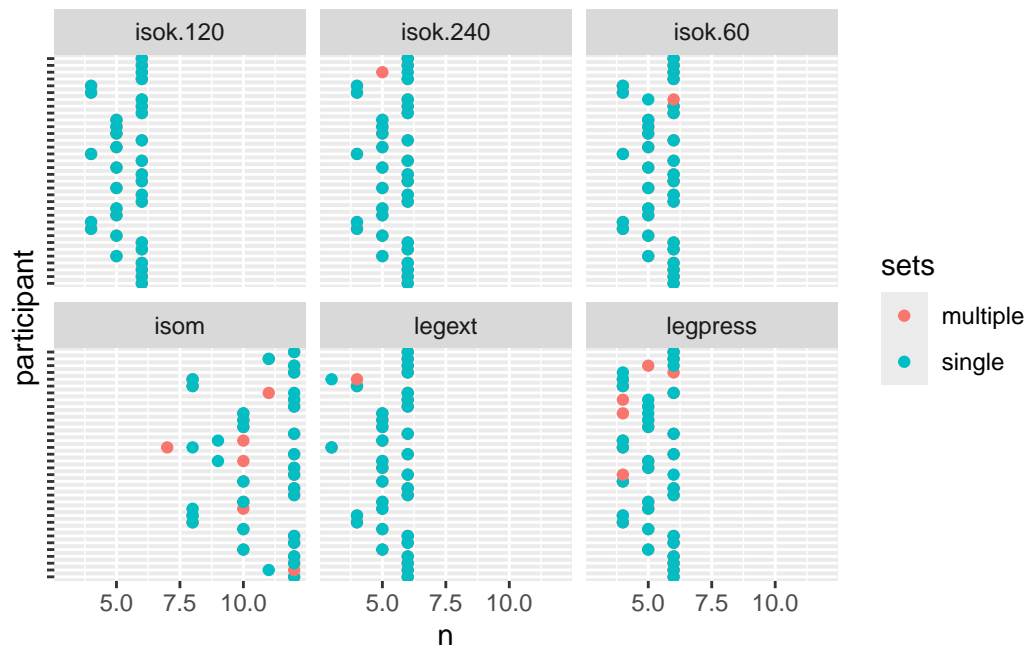
Warning: Removed 5 rows containing missing values or values outside the scale range (``geom_line()``).



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

Hopkins, W G. 2000. "Measures of Reliability in Sports Medicine and Science." *Sports Med.* 30 (1): 1–15.