Data sets in the textbook
1. feb 10:00 – 12:00
This zip file includes all the data sets mentioned in the course textbook. Start with the README file to see how the files are matched into the textbook.

- AMSAdata.zip

Lecture 1: Introduction
1. feb 10:00 – 12:00
We will get started with reviewing some concepts in univariate context and warm-up to the multivariate cases. Materials are covered in §1.3 and §1.5, §2.5 and §2.6 of the textbook.

- Lecture1-2023.pdf

- T11-6.dat
- L1T116.Rd

Exercise Set 1: R basics
1. feb 08:00 – 10. feb 14:00
1. Understand the R-commands colMeans(), cov() and cor().
2. Review basic R commands within the introduction file (Bashir_R_intro.pdf).

During this course, students are expected to prepare for the exercises **BEFORE** the TE-session.

- Bashir_R_intro.pdf

Lecture 2 : sample mean vector and sample covariance matrix
8. feb 10:00 – 12:00
In this lecture, we will go through the sample mean vector and sample covariance matrix etc in details with an example of the Bull data provided in Table 1.10 of Exercise 1.26.

- L2-2023.pdf

- T1-10.dat
- Bull.R

Exercise Set 2a
9. feb 08:00 – 17. feb 09:00
Textbook:

Exercise 1.4, 1.5, 1.7 and 1.27.

Prepare them before the 02/03 TE-class.

-
Lecture 3
22. feb 10:00 – 12:00
We will discuss the 'random vectors and matrices'. See detailed content in slides.

- L3-2023.pdf

Exercise Set 2b
22. feb 10:00 – 2. mar 16:00
Textbook:
1. Example 2.12 and Example 2.13
2. Solve exercises 2.23 and 2.41
Prepare them before 02/03 TE-class.

-
Lecture 4
1. mar 10:00 – 12:00
In this lecture, we will see why the population mean vector can be properly estimated from sample mean vector. If we have time left, we start the topic of multivariate normal distribution as well.

- L3b-2023-final.pdf

Lecture 5
8. mar 10:00 – 12:00
MVN Part 1: We will review some statistic functions first (PMF, PDF, CDF, quantile function). Then we look into normal/Gaussion distribution in the univariate case and see how to use the QQ plot to access whether some given 'univariate data is normally distributed'. We are going to use April22.R as the R-script to go through the procedure with data sets provided in t4-1.dat and t4-5.dat. Details see slides 'MVN1-2023.pdf'

- April22.R
- t4-1.dat
- t4-5.dat
- MVN1-2023-final.pdf

Exercise Set 3a
8. mar 10:00 – 12:00
Access normality for radiation data (door open case) as given in t4-5.dat. Using methods introduced in the lecture on 08/03. The solution should include the following:

Q-Q plot for original data + hypothesis test;
create Fig 4.12 but for the door open data;
find a proper power transformation using Box-Cox method.
Q-Q plot for the transformed data (with step 3) + hypothesis test
-
Lecture 6
15. mar 08:00 – 10:00
Access normality for the MVN p>=2 cases.

- FindCrikChi2Update.R

- t4-1.dat
- t4-5.dat
- MVN2-2023final.pdf

Exercise 3b
15. mar 08:00 – 23. mar 09:00
Check MVN of the stiffness data provided in Table 4.3. This includes:
1. Check normality for each of the four attributes
2. Check normality for each pair of the two attributes (out of four)
3. Check normality for the entire four attributes
4. if time permit, try out Box-Cox transformation.
Get it prepared before 23/03 TE
-
Lecture 7
22. mar 10:00 – 12:00
We will complete the MVN part in this lecture, including the last part of access normality and the central limit theorem.
If we have time left, we start the classification topic as well.

- MVN-complete.pdf
- matrixMVN.R

Tempo of lectures survey
22. mar 10:00 – 12:00
See how students feel about the tempo of the course

- Tempo of the lectures in general

Lecture 8
12. apr 10:00 – 12:00
We will first finish the last 8 slides in the MVN. Then we move on to the classification topic.

- MVN-complete-1.pdf
- T11-8.dat
- T11-8Script.R
- classification-1.pdf

Exercise 4
12. apr 12:00 – 21. apr 10:00
Exercises should be prepared before TEs in Week 16. See screenshots taken from the international version (to avoid problems with different versions of textbook)

- Screenshot 2023-04-12 at 12.32.39.png

- Screenshot 2023-04-12 at 12.38.37.png
- Screenshot 2023-04-12 at 12.38.58.png
- Screenshot 2023-04-12 at 12.39.07.png
- Screenshot 2023-04-12 at 12.39.28.png

Lecture 9
19. apr 10:00 – 12:00
We continue with the classification (Slide #14) topic and complete the classification for 2 groups.
-
Lecture 10
26. apr 10:00 – 12:00
We will use data from Table t11-6.txt (admission data) to see how to handle classification with at least 3 groups. R-script is admission2022.R

- t11-6.txt
- admission2022.R
- hemophiliaEVA.R
- t11-5.txt
- irisaper.R
- ROCR.pdf
- classification-1&2&3.pdf

Exercise 5
26. apr 10:00 – 12:00
Choose a data set (as provided in the poster presentation in itslearning) or use your own data set to complete a classification analysis step-by-step.
-
Lecture 11
3. maj 10:00 – 12:00
We will finish the classification topic and then have a look at the PCA.

- PCA.pdf
- PCA.R

MCQ exam 2022
24. maj 10:00 – 12:00
This is a subset of MCQ questions used in 2022. One could use it for the exam preparation. Solutions will be discussed in the last TE-class.

- MCQ2022.pdf
- Examdata3.tsv

- Examdata4.tsv
- Examdata5a.tsv
- Examdata5b.tsv