Notes e-book

Generic chapters (all blocks, needs to be done now)

Data analysis methods theory (all blocks, some needs to be done now, some can wait e.g. PLS)

MST chapters (block 3, missing)

F&MCR chapters (block 4, missing)

TFIH chapters (block 1, needs to be done now)

General formatting:

Whenever in narrative mentioning packages put them in bold: \*\*devtools\*\*

Index (work in progress):

1. Introduction to the book
2. (#, 01) Introduction to R – generic DONE
   1. (##) How to get started - understanding R (and RStudio) DONE
      1. (###, 01) Organise and save scripts DONE
   2. (##) How to import data DONE (there is no text, so easy 😊)
      1. (###) Import data from R-package DONE
      2. (###) Importing a csv file DONE
      3. (###) Importing an Excel file/sheet DONE
      4. (###) Clipboard import DONE
      5. (###) Looking at the imported elements DONE
      6. (###) Numbers and factors - changing categorisation DONE
   3. (##, 01) How to merge two datasets
      1. (###, 01) Import and merge in R (MANGLER: do we need to write all codes again or simply refer/link to other sections?) Julius siger: Jeg har fjernet import-delen, for det giver ikke mening at det er der to gange i et kapitel. Der er heller ikke nogen grund til at henvise til det, det giver fin mening i sig selv.
         1. (####, 01) Edit using Tidyverse DONE
         2. (####, 01) Merging datasets DONE
            1. (#####, 01) Adding survey to buffet
            2. (#####, 01) Adding buffet to survey (MANGLER: some explanation and JULIUS is asking a question)
   4. (##, 01) How to save the data (semi- DONE – check if the explanation is sufficient now
   5. (##, 01) How to export data / results DONE
   6. (##, 01) Ready for analysis DONE
3. (#, 02) Libraries (semi- DONE – check if text is sufficient
4. (#, 03) Descriptive statistics
   1. ## Descriptives for a continuous variable DONE
      1. ### Mean/ median DONE
      2. ### Variance DONE
      3. ### Standard deviation DONE
      4. ### Calculations DONE
   2. ## Distributions of count data (MANGLER: skal vi overveje at lave dette med CATA data i stedet for? jeg er ikke helt vild med legumes datasættet nemlig.... men kan også bare blive stående)
   3. ## Aggregate DONE
   4. ## Tidyverse DONE
5. (#, 05) Plotting data DONE
   1. ## Histograms and boxplots semi- DONE – check if text is sufficient
   2. ## Scatter plots (semi- DONE – check if text is sufficient
   3. ## How to export plots DONE
6. (#, 06) Introduction to PCA and multivariate data (ved ikke om der mangler noget???)
   1. ## Interpreting model output (MANGLER: Eksempel på PCA output? Jeg foreslår PCA på sensorik datasæt lagt ind I Excel filen: BeefSensoryProfile. Jeg har beskrevet datasættet her)
7. (#, 07) Introduction linear and mixed models (ALT MANGLER)
   1. ## What is a linear model?
   2. ## Normal and Mixed models
      1. ### Normal model
      2. ### Mixed model
      3. ### ANOVA
         1. Tukey’s Honest Significant Difference (HSD)
8. Buffet data
   1. Importing and looking at the buffet data (including descriptive stats)
   2. XX
   3. XX
9. Survey data
   1. Linear models
      1. Simple
      2. Multivariable models
   2. PCA?
10. Buffet and survey data
    1. Merging data sets
    2. ANALYSIS???
11. FMCR COURSE HERE
12. (#, 10) CATA data (Check-all-that-apply)
    1. ## Importing and looking at the beer data (MANGLER: checke om data er korrekt)
    2. ## Two versions of the data (MANGLER: punktopstilling. put in picture. Is my interpretation of the functions correct? Forklare de to do outputs. Indsætte kapitel link)
    3. ## Cochran’s Q test
       1. ### Post hoc contrasts (MANGLER: Morten: explain the code in words. Skal man bruge p.value eller p.adjust?)
       2. ### For all attributes in one run (nice to know) (MANGLER: Morten explain the code in words)
    4. ## PCA on CATA data (MANGLER: explain the code in words, BOM: interpretation of the PCA plot – send by email. Bodil tilføjer når det kommer)
13. (#, 11) Hedonic rating scores (DONE)
    1. (##, 10) Plotting liking scores (MANGLER: fortolkning af output, Bom)
    2. (##, 10) Simple mixed models (MANGLER: fortolkning af output, Bom)
       1. (###, 10) Post hoc test (MANGLER: stor note er slettet, det skal udbygges med tiden?)
    3. (##, 10) Multivariable models (DONE)
       1. ### Additive models (DONE)
       2. ###Effect modification and Interactions (DONE)
14. (#, 12) CATA and hedonics (DONE)
    1. ## Individual attributes and liking (DONE)
       1. ### An example with Refreshing (DONE)
       2. ### All attributes (MANGLER: explanation for models, is this a “nice to know”??)
       3. ### A beer centric model (MANGLER: explanation for models)
    2. ## PCA on CATA and Liking (MANGLER: explanation for models)
15. (#, 13) Projective mapping (MANGLER: ALT, men vi ser ikke på det i år 2022)
    1. PCA on projective mapping data
    2. (MFA on projective mapping data)
16. (#, 14) TFIH exercises (MANGLER: ALT, der er skrevet et skelet dog)
17. (#, 15) Preference mapping (først til 2023: PLS on CATA and liking data)

TO DO LIST:

Bodil

* Edit Word file to delete the stuff in the list, which has been fixed
* Add **MST** chapters and main contents (before Xmas)
  + Descr. Stats
  + Linear models (maybe mixed models in the future??)
  + PCA for pattern recognition
* Simplify **FMCR** data for exercises, names of variables e.g.
  + Descr. stats
  + k means + own choices
  + Naming clusters
  + log reg on clusters incl. multiway (?)

Morten

* Add Beef data (from Sensory Science) to datapackage
* Add ryebread data (from Sensory Science) to datapackage (Tjek samtykke).
* Sæt beef m PCA som intro to PCA.
* Unwrap liking long and liking wide in chp 8.
* Change headings in chapter 10 to analysis of CATA + hedonic rating – subchapters ## PCA and ## PLS.
* Add more on mixed model.
* Code for exporting model output as an Excel table (make headline so it is easy to find)
* Check all Rmds for [notes to answer] – a lot, sorry
* Write a “Smart tips and tricks” section (all the short codes, you always forget), e.g. :
  + rm(list = ls())
  + str()
  + find more... 😊
  + ask Julius for more

Bom

* Write through CATA introduction (10\_CATA) - CHECK
* Explain CATA PCA output in e.g. screen cast – Done in Text
* Explain projective mapping PCA output in e.g. screen cast
* Explain mixed models output in e.g. screen cast

Helene

* Intro – Chose what to do.
* Initiere basic plots (outliere and distributions)
  + Histogram
  + Boxplots
  + Point plots med liner til dommer
* Kontakte følgende ift. nye datasæt (se IFRO møde noter):
  + Thomas Bøker Lund, fødevareusikkerhed som hustandssurvey?
  + Sinne Smed, Indkøbsdata via GFK
  + Jørgen Dejgaard, kombinering af 4-5 forskellige databaser.
  + Lotte Holm & Thomas Bøker Lund: Mad og spisning i de nordiske lande. Forløbsdata. Tilbage fra 1990’erne.
  + COOP data, spørger til Aftagerpanelmøde
  + DTU kostundersøgelser
  + Varefakta

**Notes from meeting 3/10-2022 (Hennnnrik, Bodil, Helene, Morten)**

Make introduction chapter with

* Definition of variables (ordinal, conti,…)
* How to choose statistical model (pointing towards the chapters, what is a multivariate problem? What is a univariate problem? )

Make chapter on sensory = “Aroma summer course”

MST course (mostly just descriptives and plots and **adding linear models**)

Food+Meal CR course (plotting /descrp. K.means, logit, profiling)

In the Master in Technology, we will use the material as go-to for self-brush-up on stats.

Can this chapter on PCA be used as general introduction to multivariate / PCA for foreing students? [**Ask Åsmund to review**].

**Other ideas for data sets**

* Analysis of IFRO-kind-data (see meeting notes)
* House-hold data
* Single person
* New data: twitter, recepies, fødevarestyrelsen’s database…
* Marianne Thomsen?
* Dairy data?
* FQMC data? Nils?
* SOL Helene
* COOP dataset?
* App data? Maybe Inge Tetens?
* DTU kostundersøgelser – Helene kan nogle navne.

**Notes from IFRO meeting with Kia 24/10-2022:**

* Kia: Introducing Chi2 test in Excel in Social Science Methods, mentions R will come later in block 3 (us, MST course). Bør forstå bivariate sammenhænge efter hendes kursus.
* Kia kontakter Arne Hemmingsen ift hans R bog.
* Thomas data, fødevareusikkerhed som hustandssurvey? Helene kontakter.
* Indkøbsdata – Sinne Smed lektor <https://ifro.ku.dk/medarbejdere/?pure=da/persons/227352> ? Data ejes af firma. Tutorial datasæt kan måske frikøbes. Specialer. POCS. Ellers Sigrid Denver (???). Hedder GFK data. Helene kontakter.
* Jørgen Dejgaard data? Kombinering af 4-5 forskellige databaser. Helene kontakter.
* Lotte Holm & Thomas: Mad og spisning i de nordiske lande. Forløbs data. Tilbage fra 1990’erne. Helene kontakter.

Notes for exerecises

Ex1 – plots and descriptive hedonisk and cata (count data)

Ex XX – linear model … use the consumer thing from Ex1

Bodils egnes noter

Mixed models: Graphical user interface, text, application, Word

Description automatically generated

PCA on survey answers: 

CATA: Graphical user interface, text, application, email

Description automatically generated

Jeg mangler følgende i bogen:

* Mixed models på fx liking, hvor der ikke er gentagelse af dag, men fx smages 5 forskellige prøver per dag per forbruger. Kan være jeg bare skal skrive mere introtekst på til kapitel 7?
* PCA skal måske ligge som et separat punkt? Og ikke under CATA, for mig er det lidt rodet at der er tale om opdeling i både sensoriske metoder og statistiske metoder. Tænker statistikken er ”først”. Men hvis det giver mere mening for de studerende at det er opdelt per kursus, så kunne man evt. bare gentage info om fx PCA hvor det er relevant? Ellers et introkapitel hvor der står emner og så links til stat metoder som et overblik, og bogen så er er opdelt efter stat-metoder? Hvis det er CATA der er overskriften så skal vi også have en overskrift for ”Liking” mm.
* Med tiden skal PLS med
* Med tiden skal Introduction to R laves til Mortens format – med mindre der i FDA findes noget lignende

Intro til metode, opdeles i type af datasæt i overskrifter.

TFIH: CATA + Liking