

Literatuuroverzicht: Module 6

Introduction to Data Processing & Programming in R

Dit document geeft een overzicht van de verplichte leerstof (dit is de stof die op het tentamen getoetst kan worden) en eventuele extra verdiepende stof voor Module 6: 'Introduction to Data Processing & Programming in R'. De extra verdiepende stof is altijd openlijk toegankelijk (open access) en is bedoeld als aanvulling om de verplichte stof breder te kunnen begrijpen en toepassen. De extra verdiepende stof is geen leerstof voor het tentamen.

Het is aan te raden om de verplichte leerstof voorafgaand aan het hoorcollege in ieder geval door te nemen om een goed beeld te krijgen van de onderwerpen die aan bod komen. Zo wordt het makkelijker om goede en gerichte feedback te krijgen van de docent wanneer u vragen heeft over de collegeslides en/of de verplichte leerstof. Het is niet nodig om de leerstof al voor het college diep te bestuderen – al kan dat geen enkel kwaad. In het hoorcollege zal een duidelijke focus voor de leerstof worden aangebracht. De collegeslides en opdrachten in de werkcolleges zijn altijd leerstof en zijn daarom niet in het onderstaande schema opgenomen.

Lijst van gebruikte afkortingen

- R4DS = Het boek “*R for Data Science*”, 2nd edition:
 - <https://r4ds.hadley.nz/>
 - [LET OP: u heeft de 2e editie nodig – gebruikt u alstublieft de juiste link!]
- ISLR = Het boek “*An Introduction to Statistical Learning with R*”:
 - <https://www.statlearning.com>
- ADVR = Het boek “*Advanced R*”, 2nd edition:
 - <https://adv-r.hadley.nz>
- FIMD = Het boek “*Flexible Imputation of Missing Data*”:
 - <https://stefvanbuuren.name/fimd/>

Datum	Onderwerp	Verplichte leerstof	Extra verdiepende stof
2 juni	Intro to R & RStudio	R4DS: Hoofdstuk 2, 4.1, 4.2, 6, 8, 27.2.1 en 28 ISLR: Hoofdstuk 2.3	RStudio Cheat Sheet – Rstudio (https://rstudio.github.io/cheatsheets/rstudio-ide.pdf) Dit Posit recept
3 juni	Data Management	R4DS: Hoofdstuk 3.1 – 3.4, 4.3, 4.5, 7, 16, 20	Data Import Cheat Sheet – RStudio (https://posit.co/resources/cheatsheets/)

4 juni	Data Manipulatie	R4DS: Hoofdstuk 3, 5, 12, 13, 14, 15, 19; 27.1 – 27.3	Get started – dplyr (https://dplyr.tidyverse.org/articles/dplyr.html) dplyr versus base R – (https://dplyr.tidyverse.org/articles/base.html)
5 juni	Statistische analyse	ISLR: Hoofdstuk 2.1, 3.6.1 en 3.6.2	Intermediate Statistics with R – Hoofdstuk 5 en 6 (https://greenwood-stat.github.io/GreenwoodBookHTML)
9 juni	Functioneel Programmeren	ADVR: Hoofdstuk 5, 6.2, 9 R4DS: Hoofdstuk 27.4, 27.5	Tidy Modeling with R – Hoofdstuk 2 (https://www.tmw.org/tidyverse) furrr – (https://furrr.futureverse.org) purrr versus base R – f(https://purrr.tidyverse.org/articles/base.html)
10 juni	Geavanceerde data technieken	FIMD: Hoofdstuk 1, 2 R4DS: Hoofdstuk 18	Tidy Modeling with R (https://www.tmw.org/) Ch. 2, 3
12 juni	Visualisatie van data	R4DS: Hoofdstuk 1, 4.4, 9, 10, 11	Fundamentals of Data Visualization – Hoofdstuk 29 (https://clauswilke.com/dataviz/telling-a-story.html)
13 juni	Ontwikkelen en programmeren in R	Alleen collegeslides	Opname van de workshop R package maken (https://www.dropbox.com/scl/fi/hscfsrh30slk23k7z6d1d/wk6_recording.mp4?rlkey=yhy7wt8ldfmsybbb3000z1hl0&dl=0). Download voor de volledige opname.

Totaaloverzicht leerstof:

R4DS: Hoofdstuk 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19, 20, 27.1 – 27.5, 28

ISLR: Hoofdstuk 2.1, 2.3, 3.6.1, 3.6.2

ADVR: Hoofdstuk 5, 6.2, 9

FIMD: Hoofdstuk 1, 2

Collegeslides van alle hoorcolleges

Werkcollege opdrachten van alle werkcolleges