

Data Science for Business (M-IBS 9)

Syllabus

Prof. Dr. Stephan huber

Winter term 2023/24

Abstract

Regardless of position and discipline, evidence-based decision-making is a key competence in business, especially in an increasingly data-driven environment. Managers need to be able to situate practical problems in scientific contexts, generate reliable arguments from scientific discourse, critically evaluate research findings, and eventually interpret and present data they or the company have gathered themselves. This module furthers qualitative and quantitative research by enabling them to apply said methods in a data-driven business environment.

Learning Outcomes

After successful completion of the module, students are able to:

- demonstrate their proficiency in using computer programs to solve business problems,
- identify, process, visualize, and analyze data sets using the programming language R,
- explain various methods of data science, such as data mining, regression analysis, or clustering,
- assess and apply data science methods to solve problems in practical and professional contexts as well as in theoretical and scientific contexts,
- visualize solutions and communicate the challenges and outcomes of a data analysis process in written and oral form.

Lecturer

Prof. Dr. Stephan Huber

- *Email:* Stephan.Huber@hs-fresenius.de
- *Office:* Building 4b OG1 Room 1
- *Office hours:* Thursday 1-2pm and [upon request](#)

Lecture Times

Meeting	Day	Date	Time	Classroom
1	Friday	Sep 15	09:10 - 13:05	HS 4c OG1 / SR5+6
2	Friday	Sep 22	09:10 - 13:05	HS 4c OG1 / SR5+6
3	Friday	Sep 29	09:10 - 13:05	HS 4c OG1 / SR5+6
4	Friday	Oct 6	09:10 - 13:05	HS 4c OG1 / SR5+6
5	Friday	Oct 13	09:10 - 13:05	HS 4c OG1 / SR5+6
6	Friday	Oct 20	09:10 - 13:05	HS 4c OG1 / SR5+6
7	Friday	Oct 27	09:10 - 13:05	HS 4c OG1 / SR5+6

Module Content

Scope and Nature of Data Science

- What defines Data Science (Artificial Intelligence, Machine Learning, Big Data, Data Mining, Data Analytics, Statistics, Information Technology, Computer Science)
- The Importance of Data Science in Businesses
- Data Science Process
- Automated Decision-Making in Businesses
- Overview of Data Science Software Tools

The Programming Language R

- Installation (R, RStudio)
- Data Management and Writing Scripts
- Rmarkdown
- Git and Github

Data Visualization

- Descriptive Statistics
- Data Mining
- Empirical Models (Regression, Clustering, Difference-in-Difference)

Case Studies: Application of Data Science Methods

- Data Mining
- Regression Analysis
- Clustering and Similarity Analysis

Weekly Learning Plan

Meeting	Content	Readings
1	Introduction, project description, R installation, swirl, PC basics	Huber (2023a, pp. ch.1, 7.2)

Meeting	Content	Readings
1	R basics, R scripts, assignment operator, data import, swirl	Huber (2023a, ch. 2-3); Wickham & Grolemund (2023, ch. 3)
2	Data visualizations, descriptive statistics	Huber (2023a, ch. 4); Wickham & Grolemund (2023, ch. 2)
3	Data management, tidyverse, dplyr	Huber (2023a, ch. 5); Wickham & Grolemund (2023, ch. 4), Neth (2023, ch. 3)
4	Quantitative analysis (causal inference, data mining, regression analysis)	Huber (2023b, ch. 1-3)
5	Exercises (data management, exploratory data analysis)	Huber (2023a, ch. 5); Neth (2023, ch. 4)
6	R Markdown, Git, GitHub	Huber (2023a, ch. 6)
7	Student presentations	

Workload

125h Overall

35h physical synchronous contact hours

28h guided private study hours

62h private study hours

Learning Material:

Can be found on [ILIAS](#)

Assessment Methods and Criteria

Students complete this module with a case study (10-15 pages). In a presentation and subsequent discussion, lasting for 10-15 minutes per student, students show that they are capable of describing the status of their research, their approach, findings and results. The presentation and subsequent discussion take place during the lecture period; the exact date is set by the lecturer. Group work is permitted. The maximum group size is 5 students. In case of group work, it must be possible to clearly define and assess each student's individual performance on the basis of specified sections, page numbers, or other objective criteria. Determination of module grade: The case study contributes 65% to the module grade, the presentation and discussion contributes 35%. A passing grade in this module is achieved when the overall grade is greater than or equal to 4.0.

The project is explained in a separate document.

Course Policy

- Punctuality is expected. Please arrive on time for class.
- Refrain from taking photos or recordings during the class.
- Always take your laptop with you to class.
- Feel free to ask questions during the class or reach out via email (stephan.huber@hs-fresenius.de) after the class. Your inquiries are welcome and encouraged.

Syllabus Disclaimer

Please note that the syllabus is subject to change. While I will make every effort to inform you in advance of any modifications, there may be instances where adjustments are necessary to enhance the learning experience or address unforeseen circumstances. It is your responsibility to stay updated on any changes, so I recommend regularly checking the course announcements and your email for notifications regarding syllabus updates. Your flexibility and understanding in such situations are greatly appreciated.

Literature

- Huber, S. (2023a). *How to use R for data science: Lecture notes*. Accessed September 11, 2023. Retrieved from <https://hubchev.github.io/ds/>
- Huber, S. (2023b). *Quantitative methods: Lecture notes*. Accessed September 11, 2023. Retrieved from <https://hubchev.github.io/qm/>
- Neth, H. (2023). *ds4psy: Data science for psychologists*. Konstanz, Germany: Social Psychology; Decision Sciences, University of Konstanz. <https://doi.org/10.5281/zenodo.7229812>
- Wickham, H., & Grolemund, G. (2023). *R for data science (2e)*. Accessed January 30, 2023. Retrieved from <https://r4ds.hadley.nz/>