

Angewandte Datenverarbeitung und Visualisierung (WiSe23/24)

WiSe23/24

Daniela Palleschi

Vorle0ung 07/09/23

Inhaltsverzeichnis

Preface	3
Course Description	3
Wieso R?	3
Resources	3
I. Overview	4
Syllabus	5
Kursübersicht	7
II. Course mats	8
1. Intro to R	9
1.1. What is R?	9
1.2. Why R?	9
1.3. When to use R	9
1.4. How to use R	9
1.4.1. RStudio	9
1.4.2. RProjects	9
1.4.3. Rmarkdown and Quarto	9
1.4.4. Posit Cloud	9
III. Reports	10
2. Bericht 1	11
3. Bericht 2	12
4. Bericht 3:	13
References	14

Preface

```
rbbt::bbt_update_bib("index.qmd")
```

Wrote 1 reference to './references.bib'

This is the course website for “Angewandte Datenverarbeitung und Visualisierung” at the Humboldt-Universität zu Berlin, Institut der deutsche Sprache und Linguistik. (winter_statistics_2019?)

Course Description

Dies ist ein Einführungskurs in das Denken, Arbeiten und Kommunizieren mit / über sprachliche Daten. Der Kurs fokussiert sich auf praktische Anwendungen und die Vermittlung übertragbarer Fähigkeiten. In RStudio machen sich die Teilnehmenden sich mit der Programmiersprache R vertraut und entwickeln Fähigkeiten zur Erstellung und Vermittlung zusammenfassender Statistiken für den akademischen und beruflichen Kontext. Die Teilnehmenden lernen, Rohdaten zu laden und zu manipulieren, Tabellen mit deskriptiven Statistiken zu erstellen und die Daten angemessen visuell darzustellen. Am Ende des Kurses werden die Teilnehmenden ein besseres Verständnis dafür haben, wie man mit Daten umgeht und die Fähigkeiten besitzen, Ergebnisse klar zu kommunizieren. Studierende, die keinen eigenen Laptop zum Unterricht mitbringen können, setzen sich bitte so früh wie möglich mit der Dozentin in Verbindung, damit ein alternativer Laptop organisiert werden kann. Der Kurs wird auf Deutsch gehalten.

Wieso R?

Resources

Most of our materials are based on the book “R for Data Science” by Hadley Wickham (2nd Edition). Where possible I’ve replaced the data used in this book with linguistic datasets so that you can get an idea of how linguists might use R.

Teil I.

Overview

Syllabus

Wrote 1 reference to './references.bib'

The main aim of this course is to develop the knowledge and skills required to implement an “Exploratory Data Analysis (EDA)”. EDA is not a formal process with specific rules, but is rather “a state of mind” (Wickham et al., o. J., Ch. 11). The knowledge required to carry out an EDA is simply understanding your data and how to explore its structure to get an understanding for its distribution and patterns. The skills required to carry out an EDA are specific to the language used to conduct your EDA, which in our case is R.

v Reading from "WiSe23/24 - BA Datenverarbeitung syllabus".

v Range 'Sheet1'.

Warning: HTML tags found, and they will be removed.

* Set `options(gt.html_tag_check = FALSE)` to disable this check.

HTML tags found, and they will be removed.

* Set `options(gt.html_tag_check = FALSE)` to disable this check.

HTML tags found, and they will be removed.

* Set `options(gt.html_tag_check = FALSE)` to disable this check.

HTML tags found, and they will be removed.

* Set `options(gt.html_tag_check = FALSE)` to disable this check.

HTML tags found, and they will be removed.

* Set `options(gt.html_tag_check = FALSE)` to disable this check.

Woche	Datum	Thema	Vorbereitung
1	2023-10-18	Intro to R	[R4DS - Ch 1 (Introduction)
2	2023-10-25	Data Viz 1: Distributions and comparing groups	R4DS - Ch. 1 (Introduction)
3	2023-11-01	Wrangling 1: Transforming data	R4DS - Chp 6
4	2023-11-08	Dynamic reports with Quarto	ggplot2-book - Chp 5
5	2023-11-15	Regular Expressions	R4DS - Chp 21
6	2023-11-22	Bericht 1	fdv - Chp 14
7	2023-11-29	Descriptive Stats	fdv - Chp 29

8	2023-12-06	Reading in data	Six part series on W.E.B. D
9	2023-12-13	Wrangling 2: Tidying data	fdv - Chp 10
10	2023-12-20	Data Viz 2: Visual summaries	
Vorlesungsfrei	2023-12-27		
Vorlesungsfrei	2024-01-03		
11	2024-01-10	Bericht 2	The case against diverging s
12	2024-01-17	Writing functions	sf vgnette 1 + sf vgnette 2
13	2024-01-24	Inferential statistics	fdv - Chp 19 fdv - Chp 4
14	2024-01-31	Data Viz 3: visualising your model	gganimate vignette
15	2024-02-07	Bericht 3	
16	2024-02-14	Open session: Q&A	

Kursübersicht

Angewandte Datenverarbeitung und Visualisierung

Wrote 0 references to './references.bib'

Teil II.

Course mats

1. Intro to R

1.1. What is R?

We can do simple math.

```
5 + 5
```

Or print strings.

```
print("Hello world!")
```

Or do more complicated math using functions.

1.2. Why R?

1.3. When to use R

1.4. How to use R

1.4.1. RStudio

1.4.2. RProjects

1.4.3. Rmarkdown and Quarto

1.4.4. Posit Cloud

Teil III.

Reports

2. Bericht 1

3. Bericht 2

4. Bericht 3:

References

Wickham, H., Çetinkaya-Rundel, M., & Grolemund, G. (o. J.). *R for Data Science* (2. Aufl.).
<https://r4ds.hadley.nz/>