

TUTORIUM

Datenauswertung

EINFÜHRUNG UND GRUNDLAGEN VON R

AGENDA

- Vorstellungsrunde
- Organisatorisches
- Überblick: Statistik
- Einführung: R, R Studio, R Markdown
- Wiederholung: Skalenniveaus, Variablen, Hypothesen

VORSTELLUNGSRUNDE

- Wie heißt du?
- Wie alt bist du?
- Wo kommst du her?
- Was hast du vor dem Studium gemacht?
- Zweitfach?
- Welches Tier wärst du gerne?

SEMESTERPLAN

Datum	Thema
10.04.2019	Organisatorisches, Einführung: R-Studio, R-Markdown
17.04.2019	Dateneingabe
24.04.2019	Tutorium fällt aus (Ostern)
01.05.2019	Datenmanagement PLAN B
08.05.2019	Darstellung kategorialer Variablen (Faktoren)
15.05.2019	Darstellung metrischer Variablen (Vektoren)
22.05.2019	Schätzen und Testen
29.05.2019	Einfluss einer kategorialen auf eine kategoriale Variable PLAN B
05.06.2019	Einfluss einer kategorialen auf eine metrische Variable
12.06.2019	<i>Tutorium fällt aus (Pfingsten)</i>
19.06.2019	Einfluss einer metrischen auf eine metrische Variable PLAN B
26.06.2019	Einfluss mehrerer metrischen Variablen auf eine metrische Variable
03.07.2019	Klausurvorbereitung
08.07.2019	Klausur

ERWARTUNGEN AN MICH

- Beantwortung von Fragen
- Wiederholung der wichtigsten Inhalte der Vorlesung
- Hilfe bei der Auswertung von Daten
 - Rechnen (per Hand und mit R Studio)
 - Darstellung und Interpretation von Ergebnissen
- Ermutigung zum kritischen Hinterfragen von Forschung(-smethoden)
- Hilfsbereitschaft außerhalb des Tutoriums

ERWARTUNGEN AN EUCH

- Vor- und Nachbereitung des Vorlesungsstoffs (!)
- Regelmäßige Anwesenheit
- Keine Angst vor „dummen Fragen“ (es gibt keine!) oder falschen Antworten
- Einhaltung konventioneller Kommunikationsregeln

LEARNWEB, KONTAKT

–Learnweb:

- Kursname: DUD-2019_1
- Dozent: Volker Gehrau
- Passwort: dud_2019
- Meine Unterlagen: Tutorien MII5 und MII7 – Vitus

–Kontakt:

- E-Mail: vitus.schaefftlein@uni-muenster.de
- facebook: Vitus Schäfftlein

METHODENMODUL – LEISTUNGEN

– Studienleistungen:

- Anwesenheit in Vorlesung und Tutorium
- Befragung
 - Zeitraum: 11.04.2019 - 14.04.2019
 - Eingabe der Fragebögen: 17.04.2019 → Fragebögen mitbringen!
- Erstellen eines Forschungsberichts in R Studio + Markdown

– Prüfungsleistung: Klausur am 08.07.2019

→ Ein-Fach-Bachelor bekommen 10LP für das Modul,
Zwei-Fach-Bachelor bekommen 7LP.

FORSCHUNGSBERICHT

Thema	Aufgabe 1-Fach	Aufgabe 2-Fach
R-Studio, R-Markdown	Befragung	Befragung
Dateneingabe, Forschungsab.	Bericht Methode	
Datenmanagement	Bericht Stichprobe	
Häufigkeiten	Bericht Häufigkeiten	Bericht Häufigkeiten
Verteilungen	Bericht Verteilungen	Bericht Verteilungen
Schätzungen		
Kreuztabellen	Bericht Kreuztabellen	Bericht Kreuztabellen
Mittelwertvergleiche	Bericht Mittelwertvergleiche	Bericht Mittelwertvergleiche
Korrelationen	Bericht Korrelationen	
Regressionen	Bericht Regressionen	

STATISTIK - UNTERSCHIEDUNGEN

–Nach Anzahl der Variablen

- **Univariat:** Verteilung einer einzelnen Variable
- **Bivariat:** Zusammenhang/Unterschied von zwei Variablen
- **Multivariat:** Zusammenhang/Unterschied von mehr als zwei Variablen

–Nach Vorgehensweise

- **Deskriptiv:** Beschreiben der Stichprobe durch Kennwerte
- **Inferenziell:**
 - Hochrechnen der Stichprobe auf die Grundgesamtheit
 - Hypothesentests und Prognosen

R UND R STUDIO



VORTEILE VON R

Die Programmiersprache R

- Ist Open Source und deshalb
 - Kostenlos
 - Kann jeder sie verändern, wie er/sie will → Pakete
 - Wird sie ständig weiterentwickelt
 - Hat sie eine aktive Community
- Bietet viel mehr als andere Statistiksoftware, bspw.
 - Kompatibilität mit anderen Programmiersprachen (z. B. Python, SQL) und Programmen (z. B. LaTeX, Stan)
 - Data Mining
 - Interaktive Graphiken

R STUDIO

R Studio

- Ist eine Benutzeroberfläche für R
- Macht R übersichtlicher und effizienter
- Besteht aus vier Bereichen:

Bereich	Position	Funktion
Script-Editor	Oben links	Ermöglicht es, Abfolgen von Befehlen zu speichern → erstellt Scripte
Konsole	Unten links	Führt in die Konsole eingegebene Befehle aus
Environment/History/...	Oben rechts	Gibt Überblick über Datensätze und bereits eingegebene Befehle
Files/Plots/...	Unten rechts	Gibt Übersicht über Pakete, zeigt Graphiken bzw. Markdown-Dokumente an, erklärt Befehle

R BASICS I

– Befehle

- Sind R-syntaxgemäße Zeichenkombinationen
- Verändern bereits existierende Daten oder erstellen neue
- Beispiele: `sqrt(Objekt)`, `mean(Objekt)`, `plot(Objekt)`

– Objekte

- Speichern die Ergebnisse von Befehlen aus R
- Werden nach dem Muster `Objekt <- Befehl` erstellt
- Werden mit der Funktion `rm(Objekt)` gelöscht
- Beispiele: Variablen, Datensätze, Graphiken, Mittelwerte

R BASICS II – WICHTIGE OBJEKTE

– Vektoren

- Werden benutzt, um *metrische Variablen* darzustellen
- Befehl in R: `c(Zahl1, Zahl2, ... Zahln)`

– Faktoren

- Werden benutzt, um *kategoriale Variablen* darzustellen
- Befehl in R: `factor(c(„Wort1“, „Wort2“, ... „Wortn“))`

– Dataframes

- Werden benutzt, um Datensätze zu darzustellen, sodass
 - Jede Spalte eine Variable repräsentiert
 - Jede Zeile einen Fall (i. d. R.: eine Person) repräsentiert
- Befehl in R: `data.frame(Variable1, Variable2, ... Variablen)`

R – HILFE

–In R Studio: Help → Cheatsheets

–Online

- [Learnweb](#)
- [Lynda](#)
- [YouTube](#)

–Literatur:

- Luhmann, M. (2015). *R für Einsteiger. Einführung in die Statistiksoftware für die Sozialwissenschaften*. Beltz: Basel.
- Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. Sage: Los Angeles [u. a.].

MARKDOWN

Markdown

- Ist eine Programmiersprache
- Hat zwei grundlegende Funktionen:
 - Text-Formatierung (einfache Syntax)
 - Konvertierung von formatiertem Text nach HTML
- Wurde entwickelt, um das Arbeiten mit HTML zu erleichtern

R MARKDOWN

R Markdown

- Ist ein Paket, das Markdown in R implementiert
- Kann Scripte erstellen, die zwischen Text- und Codeebene unterscheiden
- Ermöglicht es dadurch, R Scripte in wichtige Formate (Word, HTML, PDF) zu konvertieren → „**knitten**“
- kann mit dem Paket „rticles“ verschiedenste Formatvorgaben einhalten (beispielsweise von Journals)

R MARKDOWN – INSTALLIEREN

```
Console Terminal x
~/
> install.packages("markdown")
Installing package into 'C:/Users/vitus/Documents/R/win-library/3.5'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.5/markdown_0.9.zip'
Content type 'application/zip' length 189536 bytes (185 KB)
downloaded 185 KB

package 'markdown' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
  C:\Users\vitus\AppData\Local\Temp\Rtmpm26vKR\downloaded_packages
> library(markdown)
Warning message:
Paket 'markdown' wurde unter R Version 3.5.3 erstellt
> |
```

ERKLÄRUNGEN ZU MARKDOWN

- Umlaute werden nicht angezeigt: File → Reopen with Encoding → UTF-8
- Dateinamen der Markdown-Scripts rot: zeigt an, dass die Datei geändert wurde
- Rauten vor *install.packages()* in Markdown-Scripts:
 - Verhindern, dass der Code ausgeführt wird
 - Sollten nur entfernt werden, wenn das Paket noch nicht installiert ist

R – STANDARDS SETZEN

- R Studio als Standardprogramm für Markdown-Scripts (Endung: .rmd) setzen
- Vorschau für Markdown-Dokumente einstellen: Zahnrad → Preview in Viewer Pane
- Auf Groß- und Kleinschreibung achten
- Alle Dateien eines Projekts (bspw. Scripts, Datensatz) in einen Ordner speichern
- Diesen Ordner als Arbeitsverzeichnis festlegen

ARBEITSVERZEICHNIS SETZEN

```
Console Terminal x
~/R/ ↵
> load("daten_2019.RData")
Error in readChar(con, 5L, useBytes = TRUE) : cannot open the connection
In addition: warning message:
In readChar(con, 5L, useBytes = TRUE) :
  cannot open compressed file 'daten_2019.RData', probable reason 'No such file or directory'
> |
```

Läuft nicht → die Datei „daten_2019.Rdata“ befindet sich nicht im angegebenen Ordner

```
Console Terminal x
C:/Users/vitus/Desktop/Statistik/ ↵
> setwd("C:/Users/vitus/Desktop/Statistik")
> load("daten_2019.RData")
>
```

Läuft → Arbeitsverzeichnis richtig gesetzt

FRAGEN?

SELBSTSTUDIUM

- Interviews führen (11.04.2019 – 19.04.2019)
- Kritisch mit dem Fragebogen auseinandersetzen
- R Markdown installieren und ausprobieren
- Standards setzen
- Statistik – Unterscheidungen
- Wiederholung: Skalenniveaus, Variablen, Hypothesen

BIS NÄCHSTE WOCHEN!