



Angewandte Statistik

Julian Hinz — Universität Bielefeld

Session 1

Organisatorisches, Roadmap und R

Team

Julian Hinz

Lars Abt

Norwin Heine

Aaron Lohmann

Was ist *angewandte* Statistik?

Statistik

- ... entwickelt Methoden zur Analyse von Daten
- ... und unterstützt bei deren Anwendung auf echte Daten

Was ist angewandte Statistik?

Mathematische Statistik

- befasst sich primär mit **Methodenentwicklung**
- im Vordergrund stehen die mathematischen Eigenschaften
- oft abstrakt, praktischer Nutzen nicht immer sofort klar (Mathe halt...)

Was ist angewandte Statistik?

Angewandte Statistik

- soll mit geeigneten Methoden aus **echten Daten** Informationen gewinnen
- braucht Methodenkompetenz sowie Verständnis des inhaltlichen Problems
- in der Regel stark interdisziplinäre Zusammenarbeit

Was zeichnet angewandte Statistik aus?

- Bedeutung der angewandten Statistik wächst im Datenzeitalter rasant
- hohes Maß an Interdisziplinarität
 - mit allen Vor- und Nachteilen: hoher Bedarf an effektiver Kommunikation
- oft aufwändige Datenaufbereitung und viel Programmierarbeit
- komplexe Daten, welche es eben nicht erlauben, Standardwerkzeuge anzuwenden

**Angewandte Statistik klingt
langweilig?**

Data Science.

Slides



<https://statistik.julianhinz.com>

Plan

1. Organisatorisches
2. Roadmap
3. R

Vorlesung

- **Dozent:** Julian Hinz (julian.hinz@uni-bielefeld.de)
- zwei Gastvorlesungen: Roland Langrock und Aaron Lohmann
- Termin: Dienstags, 12:15–13:45 Uhr in H7
- Theorie und Praxis: Computer mitbringen, wir werden programmieren!

Tutorien

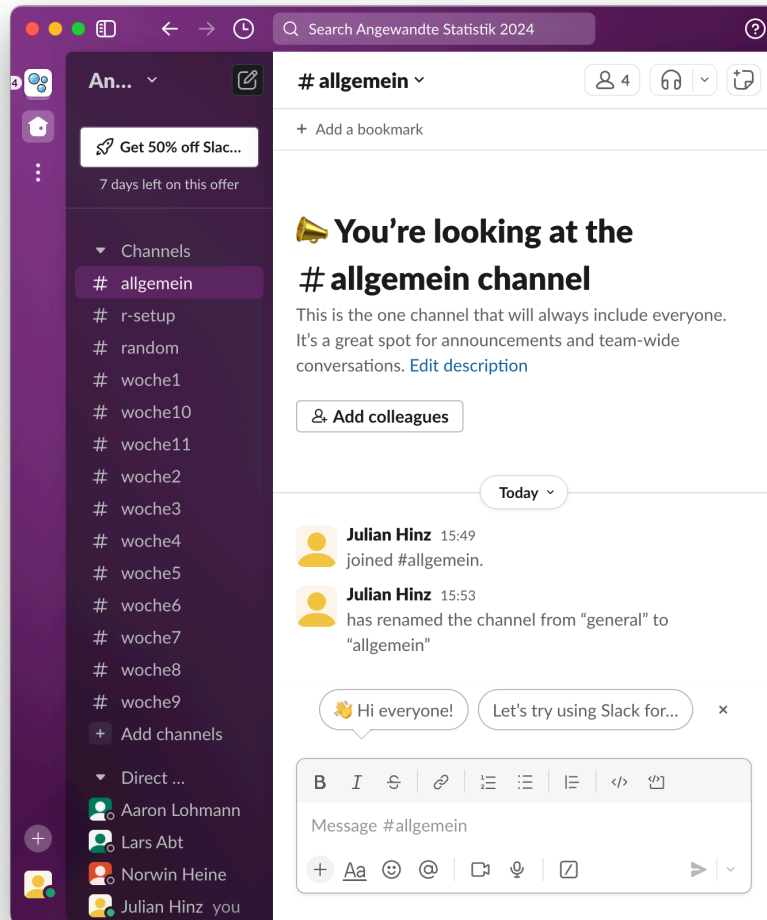
- **Tutoren:** Norwin Heine & Lars Abt
- 4 Tutorien pro Woche:
 - Mittwochs 14-16h
 - Donnerstags: 12-14h, 14-16h, 16-18h
- Platzvergabe morgen, also: **HEUTE** anmelden im ekvv!
- 3 Tutorien in den eKVV-Stundenplan nehmen, inkl. Prioritäten
- Ziel: Optimal auf die Projektaufgaben vorzubereiten
- Tutorien beginnen nächste Woche (17./18. April)

Platzvergabe morgen, also:
JETZT anmelden im ekvv!

Material

- Kurs Website: <https://statistik.julianhinz.com>
- Peter Bruce, Andrew Bruce & Peter Gedeck (2021), *Praktische Statistik für Data Scientists*, O'Reilly
- Fahrmeir, Kneib, Lang (2009), *Regression*, Springer

Kommunikation außerhalb der Vorlesung



Slack-Forum für Fragen insb. zu den Aufgabenzetteln:

- <http://statistik2024.slack.com>
- Registrierung mit ...@uni-bielefeld.de E-mail-Adresse
- kann direkt im Browser oder per App genutzt werden
- bitte Vorname & Nachname als Nutzernamen angeben

Prüfungsleistung

Ersttermin im Sommersemester:

- Portfolioprüfung über Öko/CoMet/AngStat
- je 30 Punkte pro Veranstaltung
- Öko & CoMet liegt bereits hinter Ihnen
- die 30 Punkte in Angewandte Statistik werden über **Projektaufgaben** vergeben

Zweitertermin im Sommersemester:

- normale Klausur (vermutlich im September) über alle drei Vorlesungen
- kein Nachteil was den Schwierigkeitsgrad/die Bewertung betrifft

Projektaufgaben

Formalia:

- **zwei Aufgabenzettel** mit je 3-5 Aufgaben, 15 Punkte pro Zettel
- jeweils über einen Zeitraum von einer Woche zu bearbeiten
- Zettel 1: Mitte Mai, Zettel 2: Ende Juni
- **nur Einzelabgaben**

Inhaltlich:

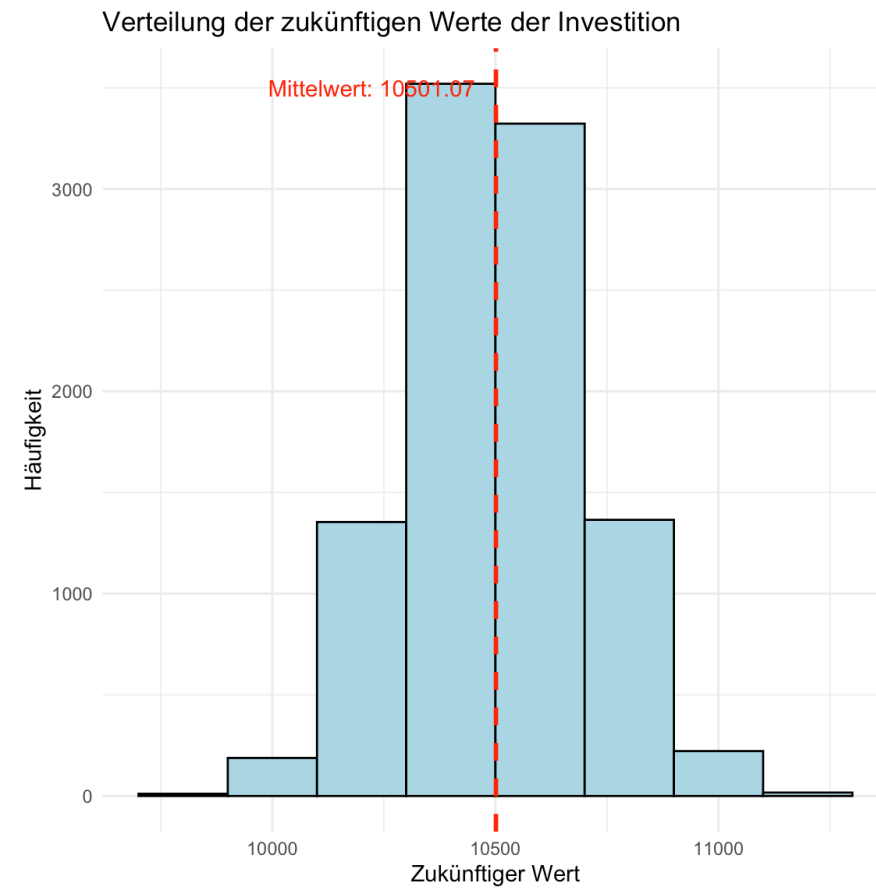
- jede*r von Ihnen bekommt **eigene Datensätze** zur Analyse
- Methoden die in Vorlesung & Tutorien geübt werden
- **Nutzung von R** ist essenziell

Roadmap

Session 1

heute

- R
- Simulationen



angewandte-statistik - RStudio

code1.R

Source on Save Run Source

Environment History Connections Tutorial

Import Dataset 126 MiB List

R Global Environment

Environment is empty

Files Plots Packages Help Viewer Presentation

Home > angewandte-statistik

Name	Size	Modified
..		
angewandte-statistik.Rproj	205 B	Apr 8, 2022
code		
input		
output		
temp		

1:1 (Top Level) Copilot: No completions available. R Script

Console Terminal Background Jobs

R 4.2.2 · ~/angewandte-statistik/

```
R version 4.2.2 (2022-10-31) -- "Innocent and Trusting"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: aarch64-apple-darwin20 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

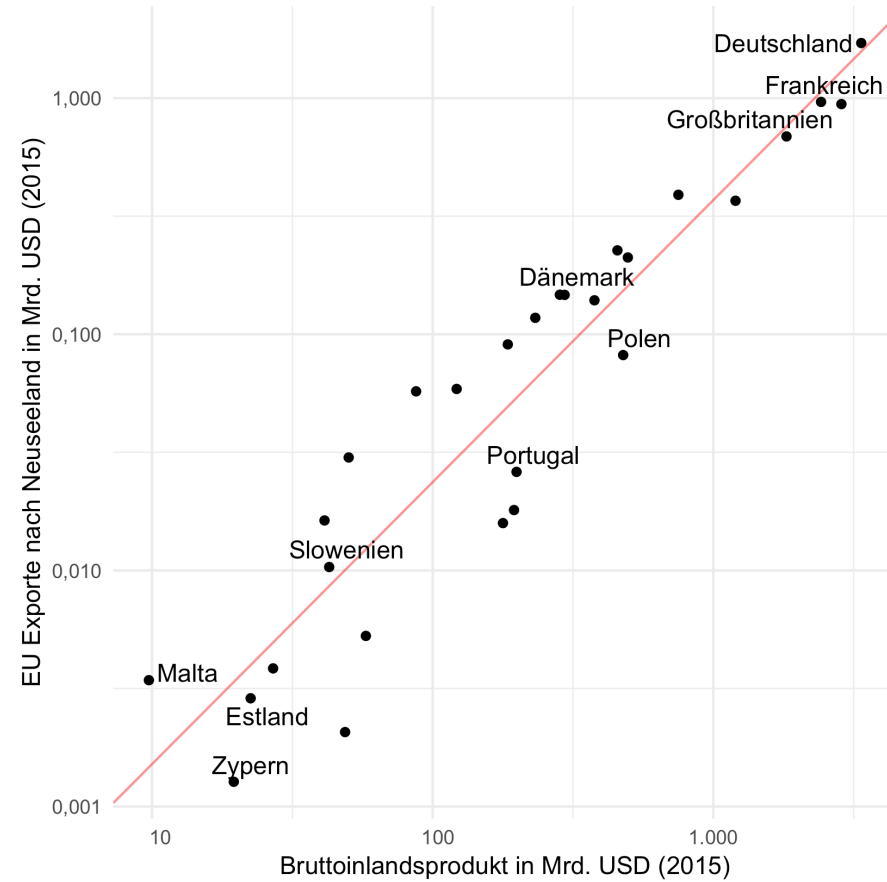
Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
```

Session 2

16. April

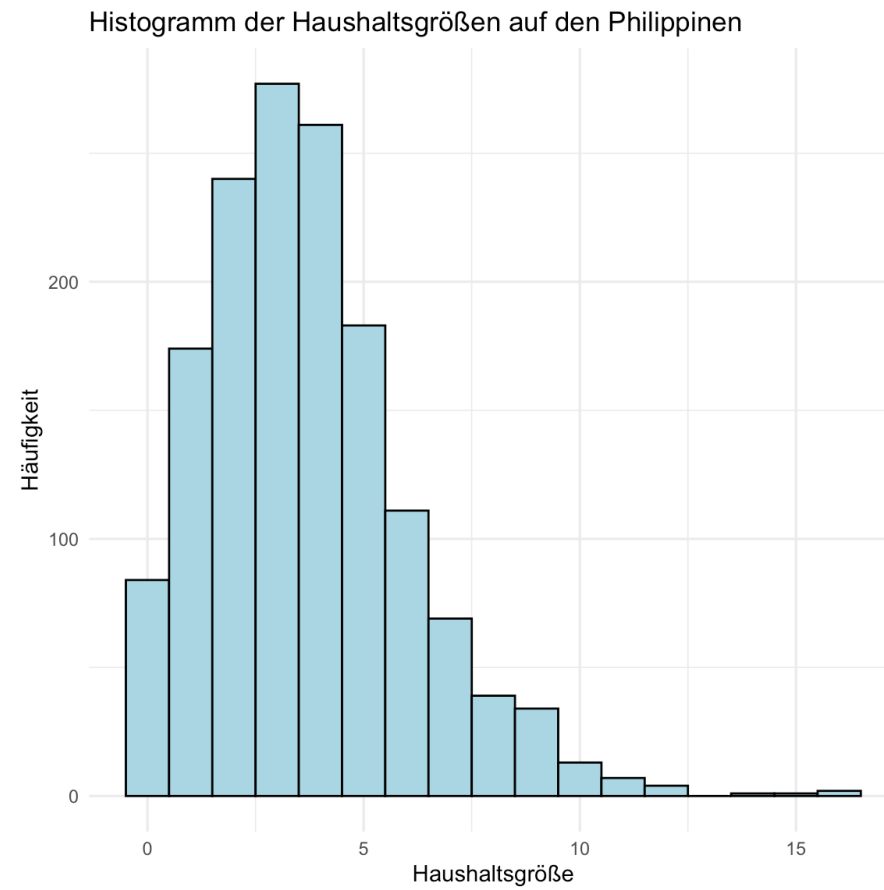
- Parametrische Regression
- Wiederholung Lineare Regression



Session 3

23. April

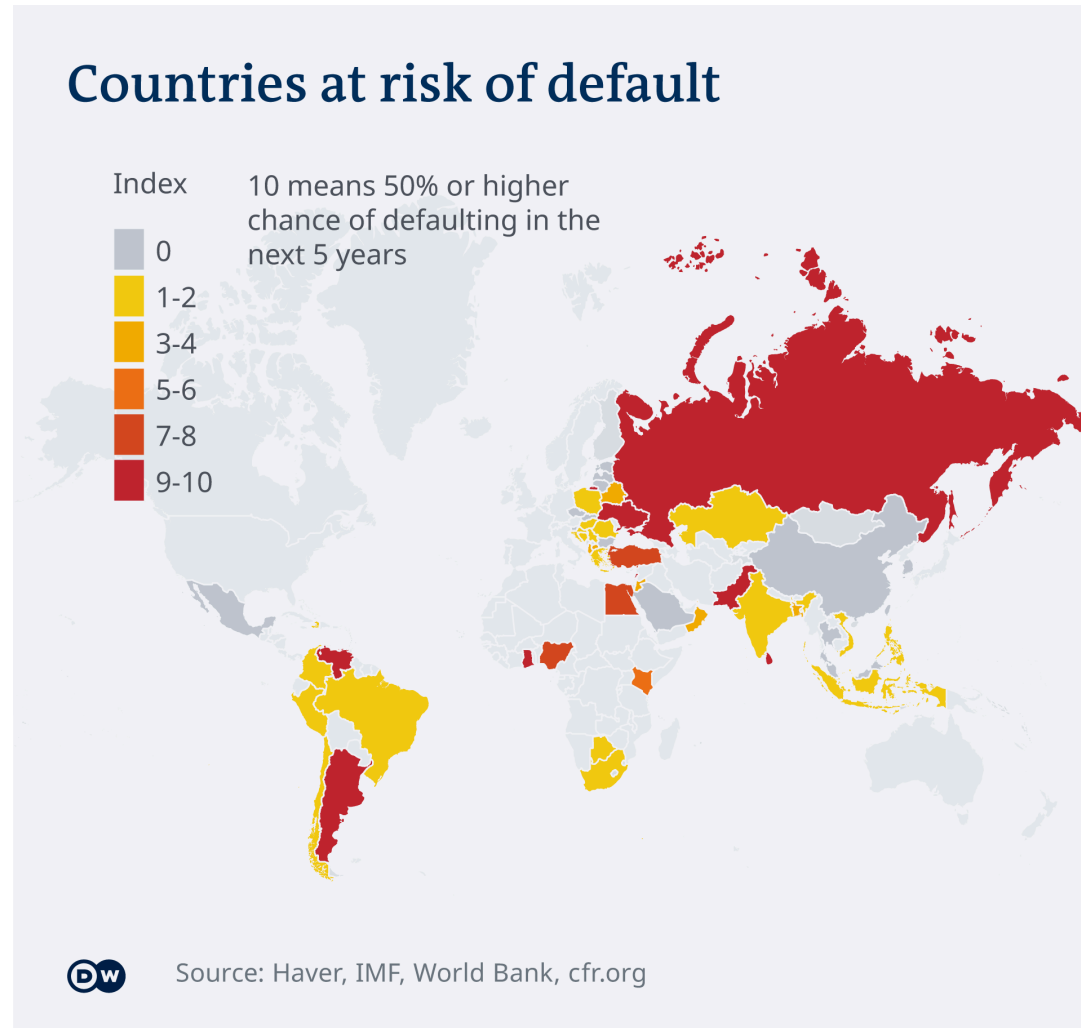
- Generalisierte Lineare Modelle
- Poisson Regression



Session 4

30. April

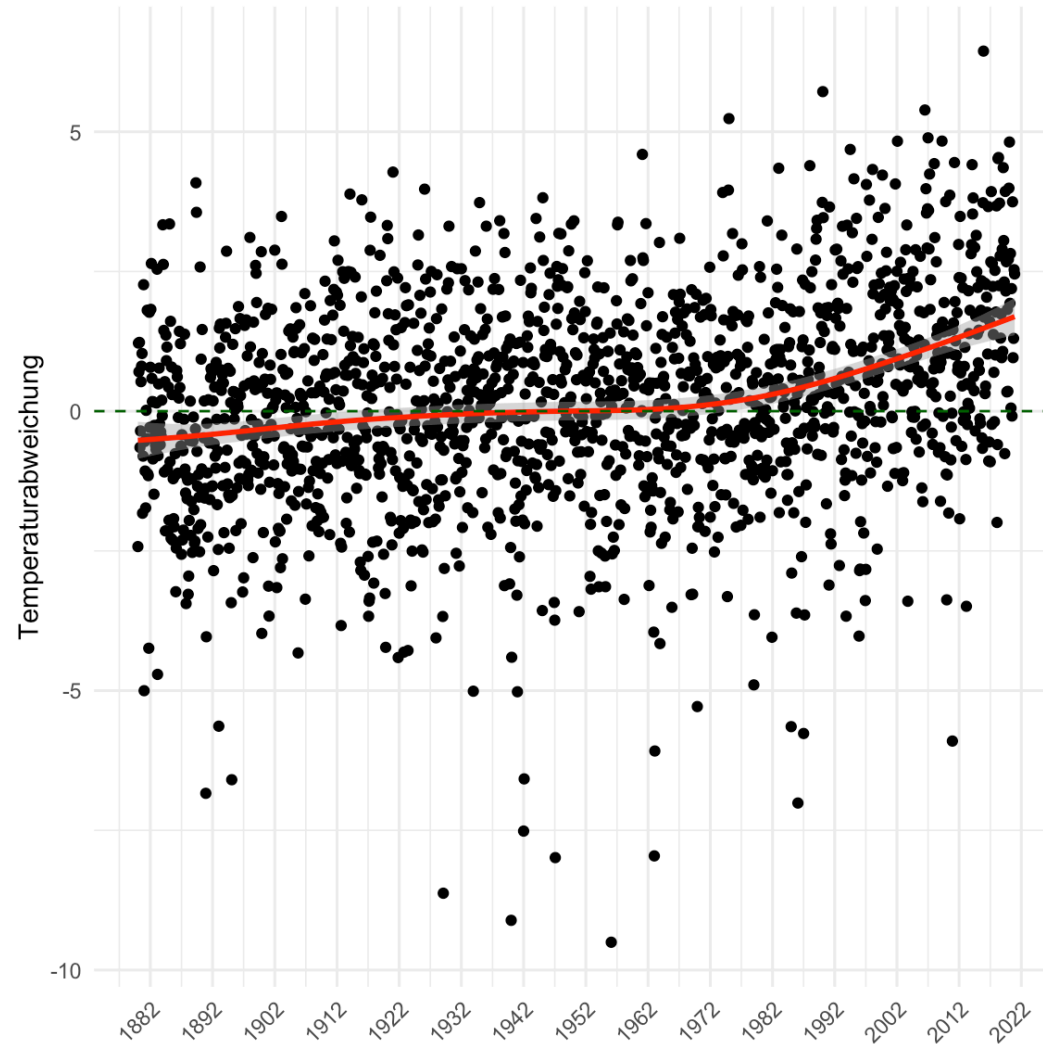
- Logistische Regression



Session 5

7. Mai — Roland Langrock

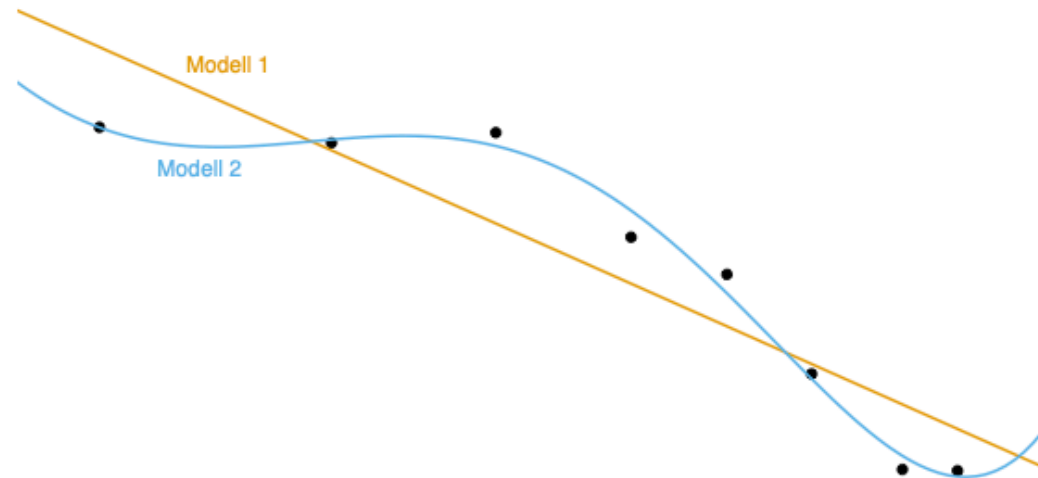
- Nichtparametrische Regression



Session 6

14. Mai

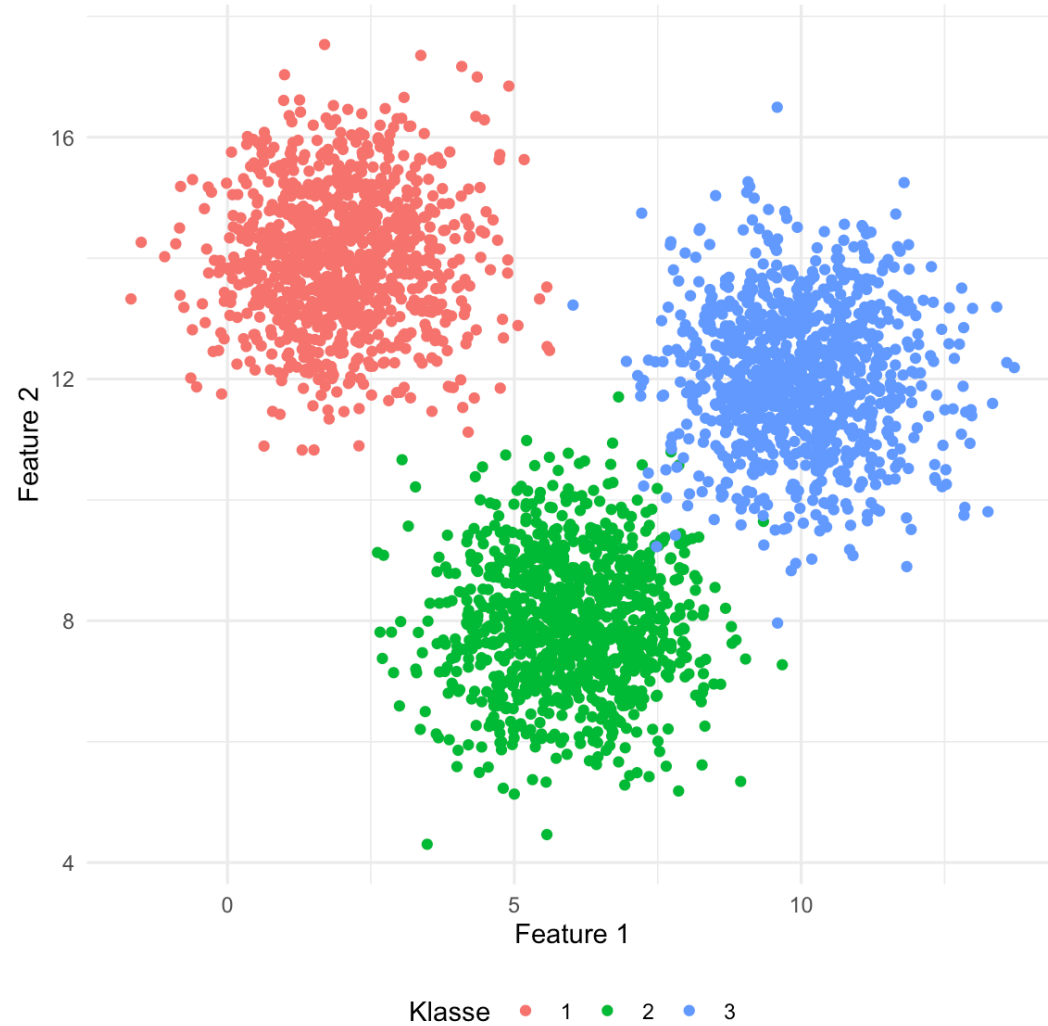
- Modellwahl



Session 7

21. Mai

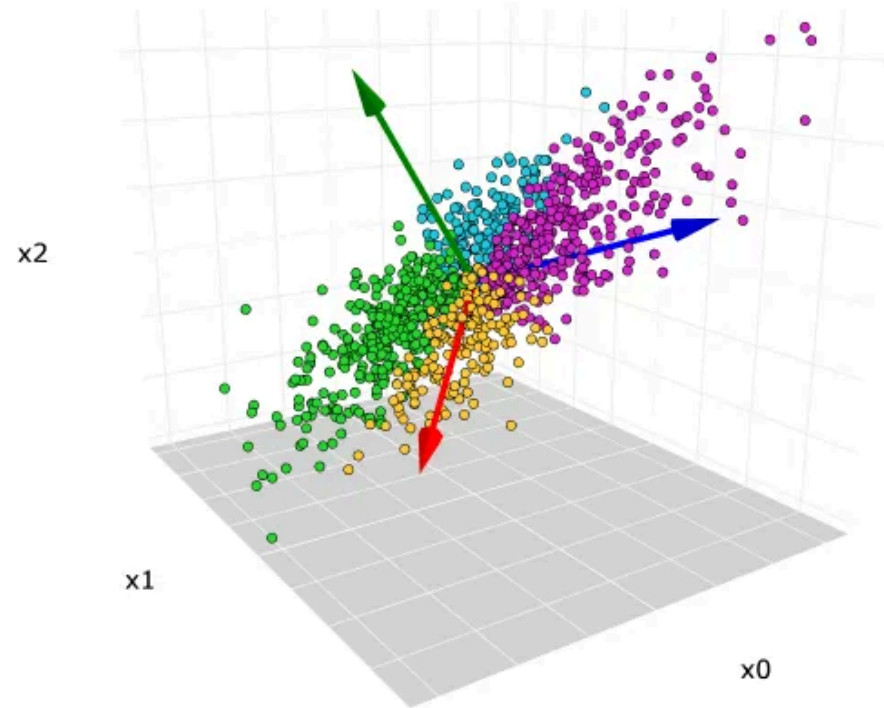
- Klassifikation



Session 8

28. Mai

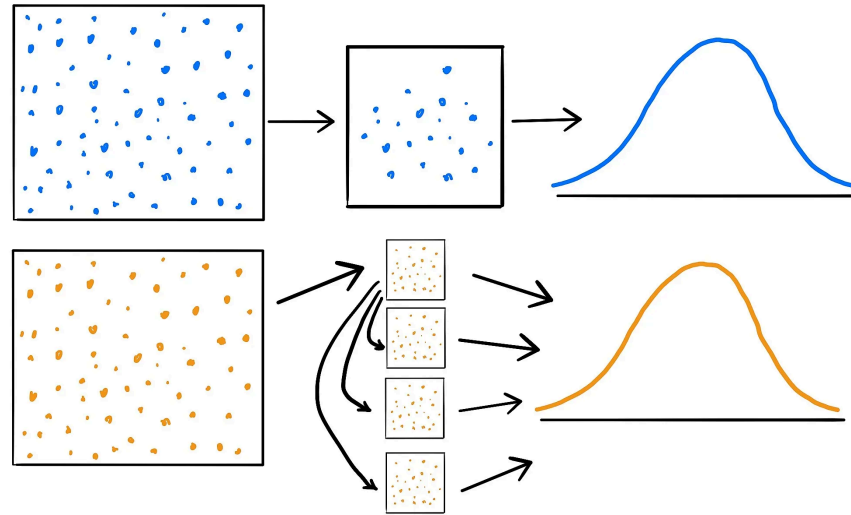
- Dimensionsreduktion



Session 9

4. Juni

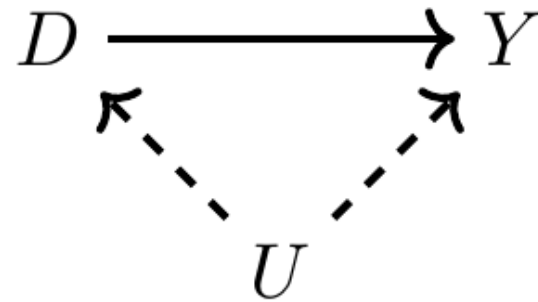
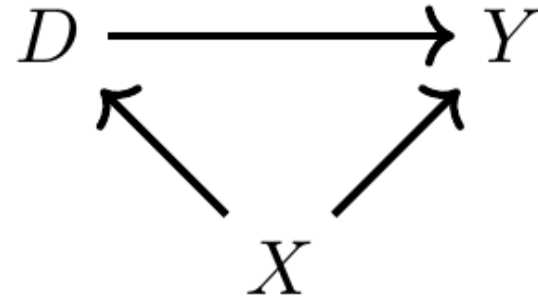
- Bootstrapping



Session 10

11. Juni — Aaron Lohmann

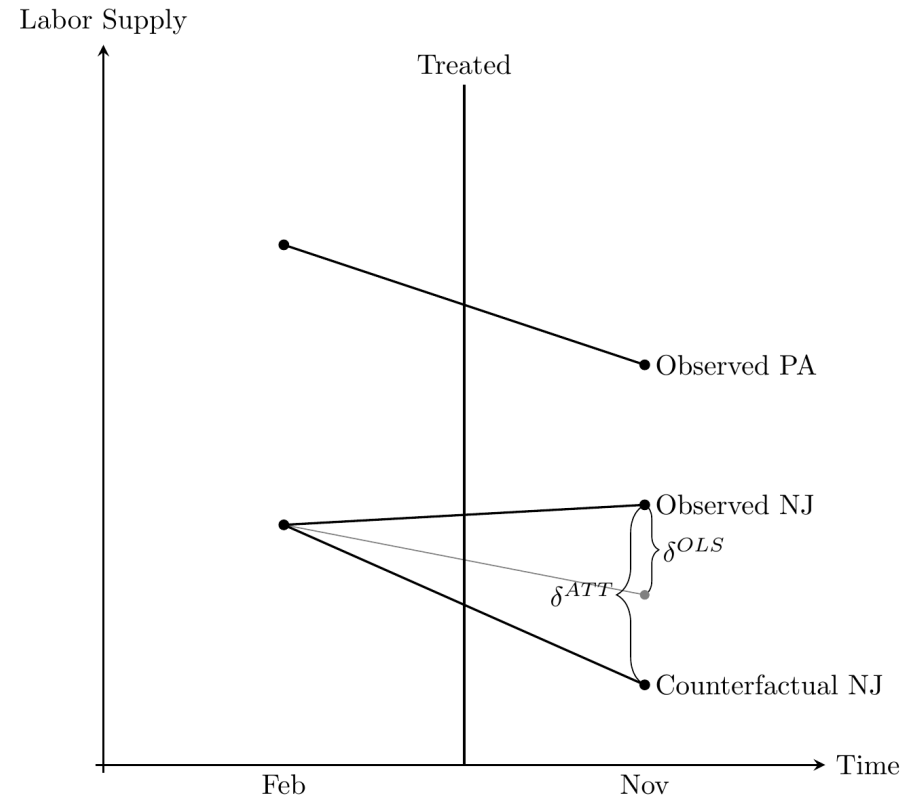
- Kausale Inferenz



Session 11

18. Juni

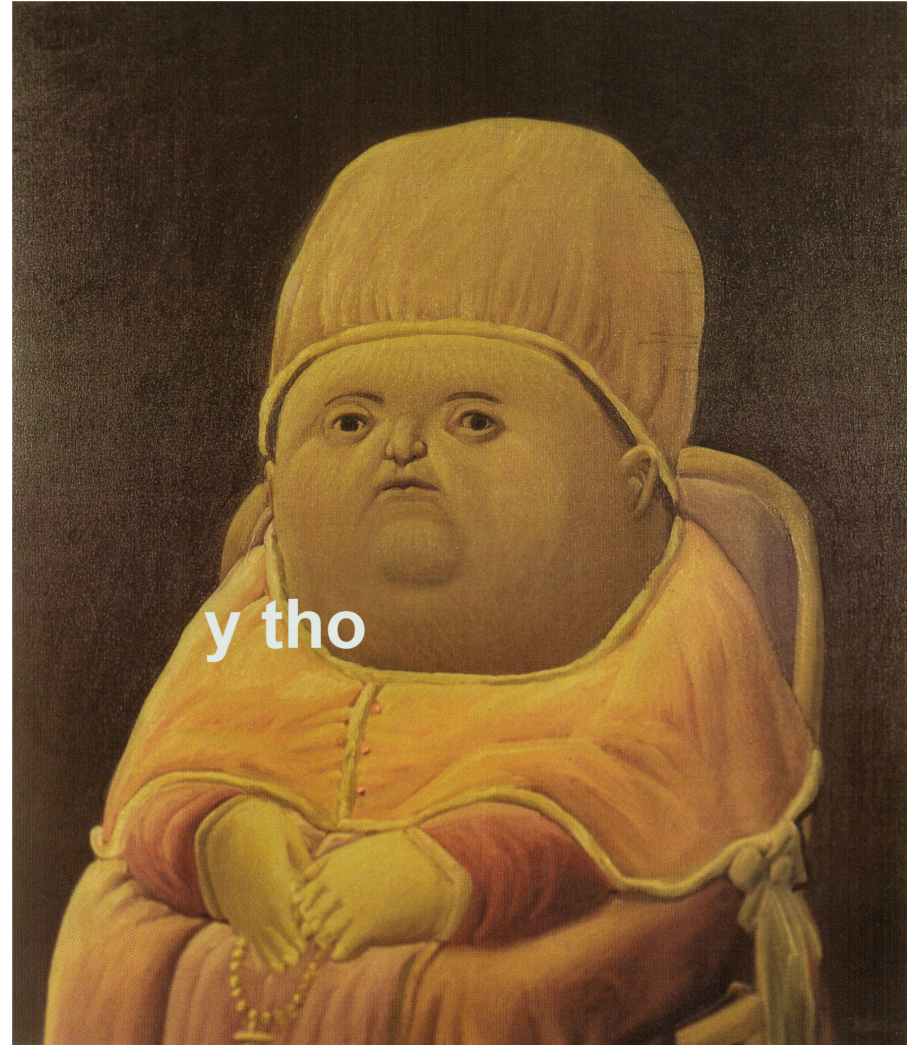
- Difference-in-Differences



Session 12

25. Juni

- Q & A

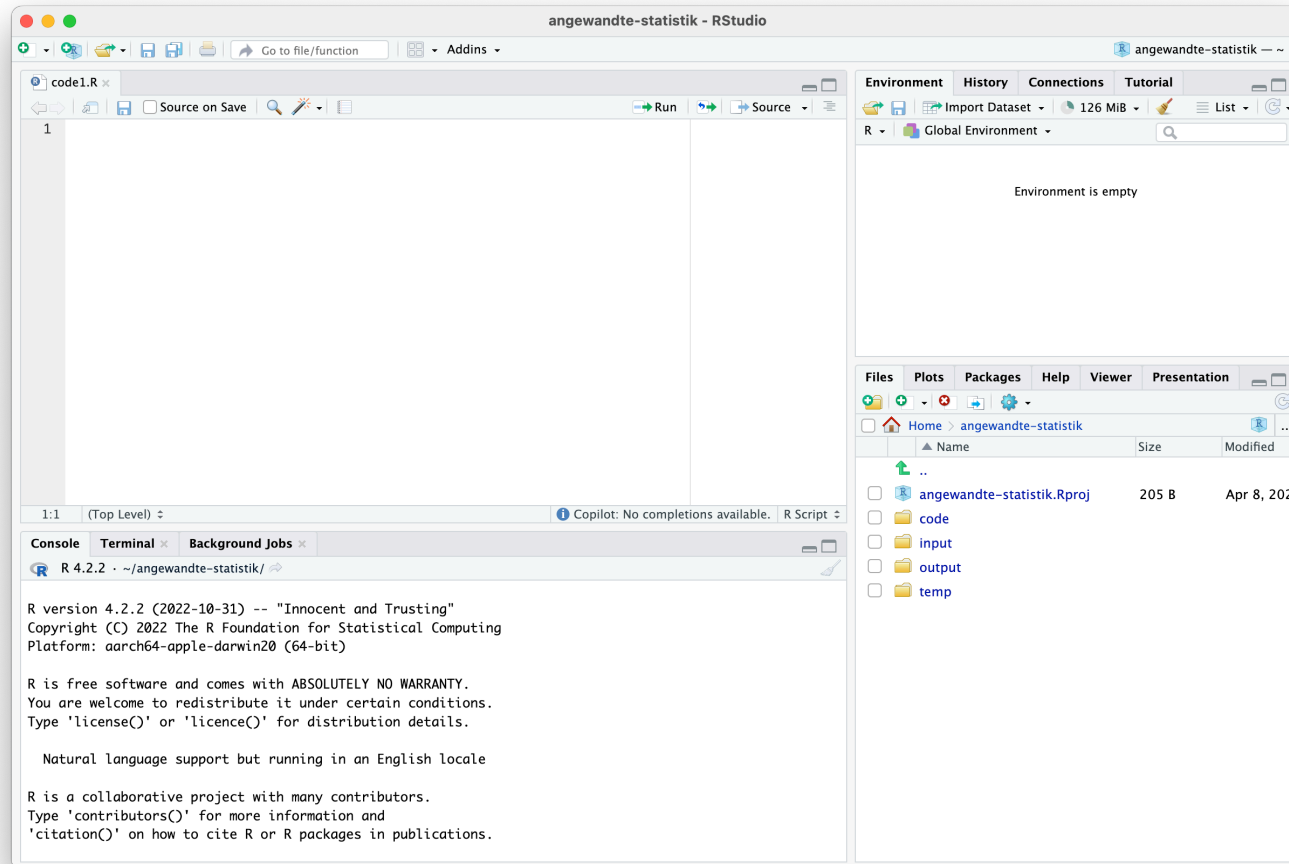


R

R

- Angewandte Statistik ist ohne Programmierung undenkbar
- Statistiksoftware R: sehr mächtig, kostenlos, und open source → ideal!
- wir wissen: R Kenntnisse sehr heterogen über Hörer*innenschaft verteilt
→ Tutorien entsprechend gestaltet

Installation von R und RStudio



Installation von R und RStudio

Schritt 1: R installieren

- Mac (Apple Silicon): <https://cran.rstudio.com/bin/macosx/big-sur-arm64/base/R-4.3.3-arm64.pkg>
- Mac (Intel): https://cran.rstudio.com/bin/macosx/big-sur-x86_64/base/R-4.3.3-x86_64.pkg
- Windows: <https://cran.rstudio.com/bin/windows/base/R-4.3.3-win.exe>
- Linux: `sudo apt install r-base r-base-dev`

Installation von R und RStudio

Schritt 2: RStudio installieren

- Mac: <https://download1.rstudio.org/electron/macOS/RStudio-2023.12.1-402.dmg>
- Windows: <https://download1.rstudio.org/electron/windows/RStudio-2023.12.1-402.exe>
- Linux: ...

Fragen? → Slack.