

# Auswertung musikhistorischer Daten

Matthias Richter

8. und 9. September 2022

Einführung R

Einführung mvdb

# Einführung R

Programmiersprache spezialisiert auf Statistik  
mit R arbeiten heißt:

- ▶ Daten in Variablen ablegen  
`var <- read_csv("data.csv")`
- ▶ Variablen bearbeiten  
`data <- prepare(raw)`
- ▶ Variablen ausgeben  
`View(data)`
- ▶ Daten visualisieren  
`ggplot(data) + geom_point()`
- ▶ Sequenzen aus Arbeitsschritten in Skripts ablegen und ausführen

# R

- ▶ entwickelt von R-Foundation
- ▶ Open Source
- ▶ durch *packages* erweiterbar

# RStudio

- ▶ Integrierte Entwicklungsumgebung für R
- ▶ Editor für Skripts
- ▶ Debugger für Variablenbelegungen etc.
- ▶ Explorer für Grafiken etc.
- ▶ R-Konsole zum Ausführen
- ▶ entwickelt von RStudio PBC

entwickelt eine Reihe von R-packages, bspw.

- ▶ Tidyverse
- ▶ knitr
- ▶ RMarkdown
- ▶ Shiny

# Tidyverse

- ▶ Skripten in Vanilla-R kann schmerzhaft sein  
`data[data$col1=="3",]$col2 <- "4"`
- ▶ Tidyverse vereinfacht Handling enorm  
`data %<>% mutate(col2 = ifelse(col1 == "3", "4", NA))`



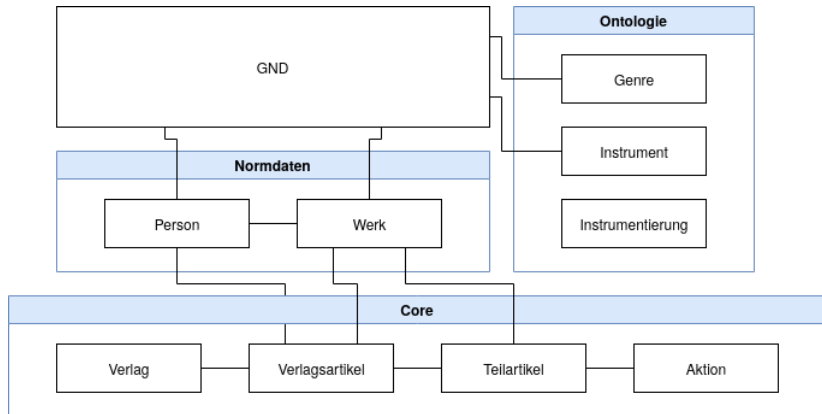
# Installation RStudio

# RStudio Shortcuts

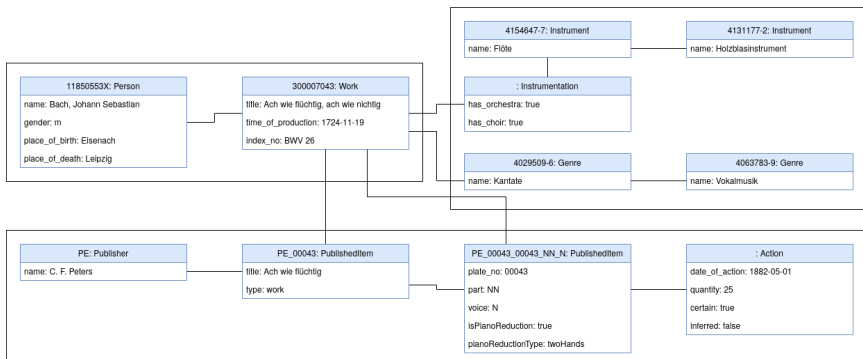
- ▶ Strg + Enter: Zeile ausführen
- ▶ Strg + Z: Rückgängig
- ▶ Strg + Shift + Z: Wiederherstellen
- ▶ Alt + runter/hoch: Zeile bewegen
- ▶ Alt + Shift + runter/hoch: Zeile duplizieren

# Musikverlagsdatenbank

# Datenmodell



# Beispiel



# Architektur

