

Einführung in die Datenanalyse mit R - Modularer Aufbau von R

Jan-Philipp Kolb

08 Mai, 2019

Wo man Routinen findet

- Viele Funktionen sind in Basis-R enthalten.
- Viele spezifische Funktionen sind in zusätzliche Bibliotheken integriert.
- R kann modular durch sogenannte Pakete oder Bibliotheken erweitert werden.
- Die wichtigsten Pakete, die auf CRAN gehostet werden (14117 at Mi Mai 08)
- Weitere Pakete findet man z.B. unter **bioconductor**

Übersicht R-Pakete

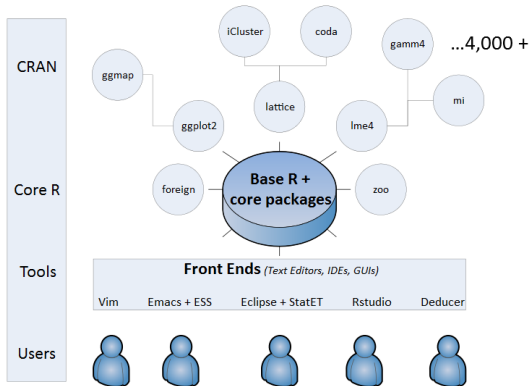


Figure 1: Overview R-packages

Installation von Paketen

- Die Anführungszeichen um den Paketnamen herum sind für den Befehl `install.packages` notwendig.
- Sie sind optional für den Befehl `library`.
- Man kann auch `require` anstelle von `library` verwenden.

```
install.packages("lme4")
```

```
library(lme4)
```

Installation von Paketen mit RStudio

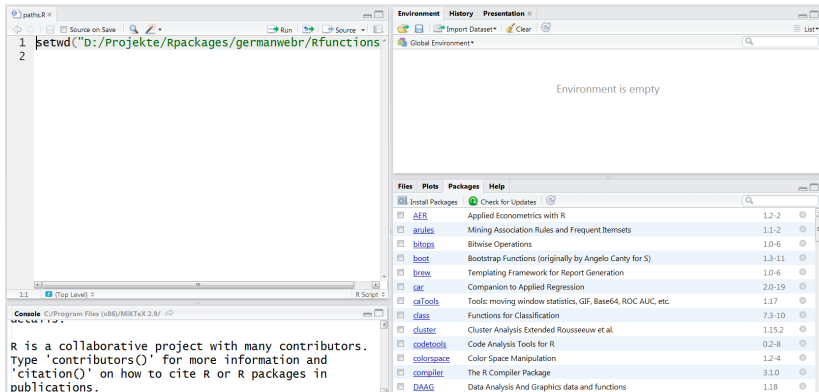


Figure 2: Package installation with Rstudio

Bestehende Pakete und Installation

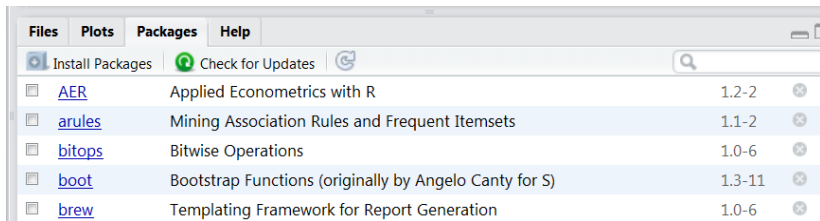


Figure 3: Existing packages

Übersicht über viele nützliche Pakete:

- Luhmann - **Table with many useful packages**

Weitere interessante Pakete:

- Paket für Import/Export - `foreign`
- **sampling-Paket für die Stichprobenziehung**
- `xtable` Paket zur Integration von LaTeX in R (**xtable Galerie**)
- `dummies` - **Paket zur Erstellung von Dummies**
- **Paket `mvtnorm` um eine multivariate Normalverteilung zu erhalten.**
- **Paket `maptools` um Karten zu erzeugen**

Pakete aus verschiedenen Quellen installieren

Pakete vom CRAN Server installieren

```
install.packages("lme4")
```

Pakete vom Bioconductor Server installieren

```
source("https://bioconductor.org/biocLite.R")  
biocLite(c("GenomicFeatures", "AnnotationDbi"))
```

Pakete von Github installieren

```
install.packages("devtools")  
library(devtools)  
  
install_github("hadley/ggplot2")
```


Wie bekomme ich einen Überblick?

- Entdecke Pakete, die kürzlich auf den **CRAN** Server hochgeladen wurden
- Nutze eine Shiny Web-App, die **Pakete anzeigt, die kürzlich von CRAN** heruntergeladen wurden.
- Werfe einen Blick auf eine **Quick-Liste nützlicher Pakete**
-, oder auf eine Liste mit den **besten Paketen für die Datenverarbeitung und -analyse**,
-, oder schaue unter **die 50 meistgenutzten Pakete**

CRAN Task Views

- Bezüglich mancher Themen gibt es einen Überblick über alle wichtigen Pakete - (**CRAN Task Views**)
- Momentan gibt es 35 Task Views.
- Alle Pakete einer Task-View können mit folgendem Befehl installiert werden: **command:**

```
install.packages("ctv")  
library("ctv")  
install.views("Bayesian")
```

CRAN Task Views

[Bayesian](#)

[ChemPhys](#)

[ClinicalTrials](#)

[Cluster](#)

[DifferentialEquations](#)

[Distributions](#)

[Econometrics](#)

[Environmetrics](#)

[ExperimentalDesign](#)

Bayesian Inference

Chemometrics and Computational Physics

Clinical Trial Design, Monitoring, and Analysis

Cluster Analysis & Finite Mixture Models

Differential Equations

Probability Distributions

Econometrics

Analysis of Ecological and Environmental Data

Design of Experiments (DoE) & Analysis of Experimental Data

A1C Übung - zusätzliche Pakete

Geht auf <https://cran.r-project.org/> und sucht nach Paketen...

- die sich für die deskriptive Datenanalyse eignen.
- mit denen man fremde Datensätze einlesen kann (z.B. SPSS data)
- mit denen man Lasso Regressionen rechnen kann
- mit denen man große Datenmengen bearbeiten kann