

Dynamische Dokumente mit Markdown und R

Sebastian Jeworutzki

26. Oktober 2016

- Markdown ist eine einfache *Auszeichnungssprache* mit der gegliederte und formatierte Dokumente erstellt werden können
- Dazu wird ein Dokument in „Markdown“ geschrieben und anschließend in HTML, Word, ODT oder LaTeX/PDF umgewandelt
- Mit speziellen Vorlagen können Berichte, Präsentationen oder ganze Bücher geschrieben werden

- In Markdown-Dokumente kann R-Code eingebunden und ausgeführt werden, um bspw. Tabellen oder Grafiken dynamisch zu erzeugen
- Oder umgekehrt formuliert: Markdown kann genutzt werden, um die in R durchgeführten Analysen innerhalb des Syntaxdatei zu dokumentieren

Warum sollte man R und Markdown nutzen?

- Dokumente mit R-Code sind reproduzierbar:
 - Berechnungen und Präsentation der Ergebnisse sind zusammen gespeichert
 - Ursprung von Tabellen und Grafiken immer klar
- Die Zusammenarbeit mit anderen wird dadurch deutlich vereinfacht
- Es lassen sich dynamische Dokumente erzeugen, die immer den aktuellen Datenbestand nutzen.
- Immer wiederkehrende Auswertungen lassen sich (teil-)automatisieren, z.B. Datensatzbeschreibungen, Tabellenbände, ...
- R Markdown kann auch genutzt werden, um die Vorbereitung von Grafiken und Tabellen für die Textverarbeitung der Wahl zu nutzen

- R
- RStudio (besonders empfohlen für Windows-Nutzer)
- R-Pakete: `knitr`, `rmarkdown`
- Um PDFs zu erzeugen wird eine LaTeX Distribution benötigt (TexLive oder MikTeX)
- Linux- und MacOS-Nutzer, die kein RStudio nutzen benötigen noch Pandoc
- Das VIM R-Plugin unterstützt `knitr` und `rmarkdown`

- Ein Beispiel mit RStudio
- Wie sieht ein Markdown-Dokument aus?
- Grundlagen zu Gliederung und Auszeichnung
- Word-Vorlagen anpassen
- Code Chunks und Knitr-Optionen
- PDF-Dokumente mit LaTeX

- Formatierungen: Handbuch zu [Pandoc Markdown](#)
- knitr: [Chunk Optionen](#)
- YAML: Optionen für die verschiedenen Dokumentarten
<http://rmarkdown.rstudio.com/formats.html> und
<http://pandoc.org/MANUAL.html#templates>

- Layoutaufgaben lassen sich nicht immer gleichermaßen in allen Formaten lösen
- Komplexe Tabellen funktionieren nur mit HTML oder LaTeX – mit Word derzeit nicht
- Oftmals ist schwierig sowohl Word-Dokumente als auch LaTeX zu erzeugen. Es gibt Pakete, die dabei helfen (bspw. ztable oder xtable) – teilweise sind aber dennoch manuelle Überarbeitungen notwendig
- Meine Empfehlung: wenn es geht LaTeX nutzen
- Teilweise kann der Umweg von HTML nach Word einfacher sein

Erweiterungen

- **bookdown**
 - Unterstützt Nummerierungen und Querverweise für Absätze, Gleichungen etc.
 - Alternative für Querverweise: **kfigr**
- **citr** und **knitcitations** für Zitate
- Komplexe Tabellen für HTML und LaTeX
 - **xtable** Traditionell genutztes Paket
 - **htmlTable** Tabellen mit zusammengefassten Zellen usw.
 - **ztable** Tabellen mit zusammengefassten Zellen usw.
 - Kann in in Rstudio auch für copy and paste von Tabellen genutzt werden
 - klappt aber nicht immer mit Word - OpenOffice ist da besser!
 - **sjPlot** löst das Zusammenfassen von Zellen besonders effektiv, copy and paste wie bei ztable möglich

- [prettydoc](#) Schöne Vorlagen
- [rticles](#) Vorlagen für einige (wissenschaftliche) Zeitschriften
- [tufte](#) PDF und HTML Vorlage mit Tufte-Design
- [tufterhandout](#) typografisch schlichteres Tufte-Design

- `rmdformats` Vorlagen für HTML Dokumente
- `markdowntemplates` Vorlagen für HTML Dokumente
- `flexdashboard`
 - Dashboards mit R
 - Mit OpenCPU oder shiny auch dynamisch generierbar

- Neben den von Rstudio integrierten Vorlagen gibt es weitere Präsentationsvorlagen
- **rmdshower**
 - Steuerbar über Tastaturkommandos
 - 2 Spalten Layouts
 - Gut als PDF druckbar
- **revealjs**
 - Ermöglicht das Schrittweise einblenden von Aufzählungen
 - Animierte Folienübergänge
 - Verschiede Möglichkeiten zur Foliengestaltung (z.B. Hintergrundbild oder -video)
 - Erweiterungen: Vortragsnotizen im separaten Fenster, Zoom mit Alt + klick

- Neben R können auch andere Sprachen verwendet werden
- Alle unterstützten Formate gibt es auf der [knitr-Homepage](#) python, c, fortran, sql, awk, ruby, haskell, bash, perl, dot, tikz, SAS, coffeescript, Scala, Python, and Bash

- LaTeX-Dokumente mit knitr
- **ReporteR**
 - Kein direktes Arbeiten am Dokumenttext möglich. Elemente des Dokumentes werden über R Befehle erzeugt
 - Erzeugt Powerpoint und Word sowie HTML
 - In Powerpoint mit **editierbaren Vektorgrafiken**
 - Sehr flexible und einfache **Tabellengestaltung** (inkl. verbundener Zellen)