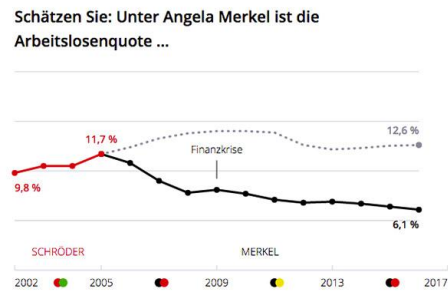
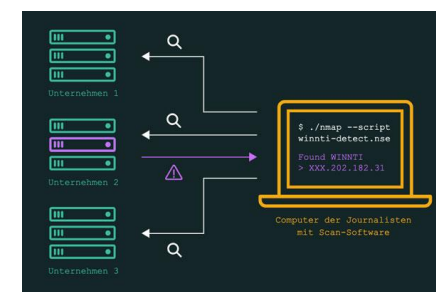
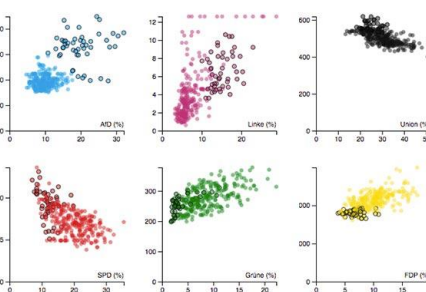
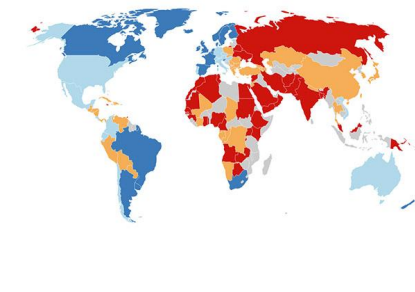
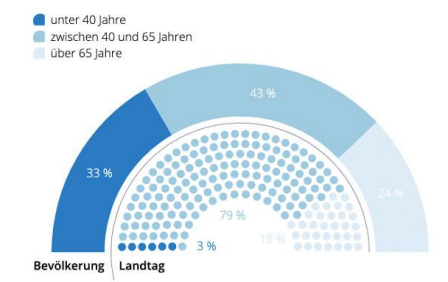
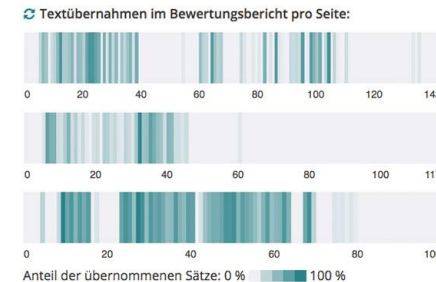
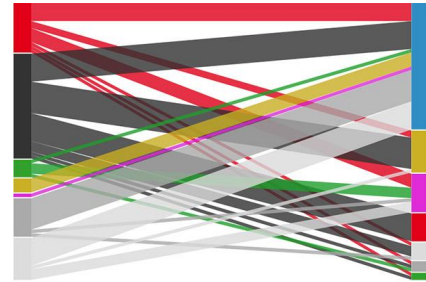
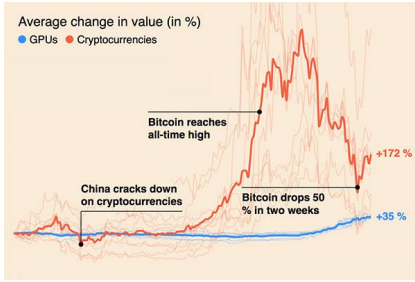
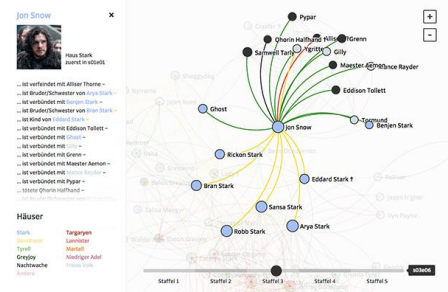
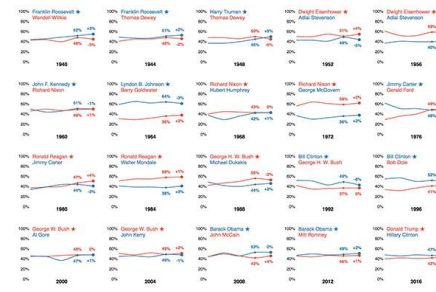
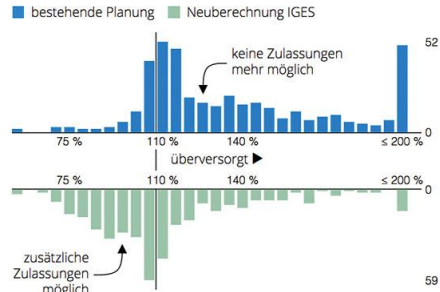
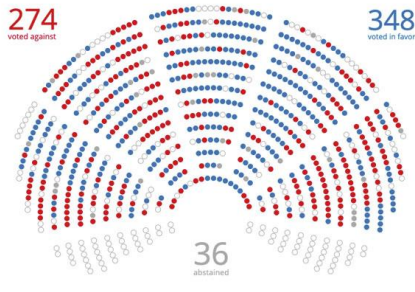


Datenvisualisierung mit R

Steffen Kühne
ifp Bootcamp 2019

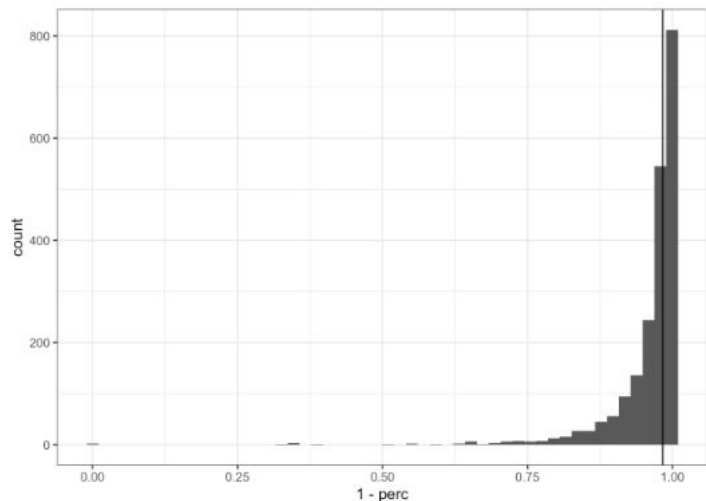


Datenvisualisierung



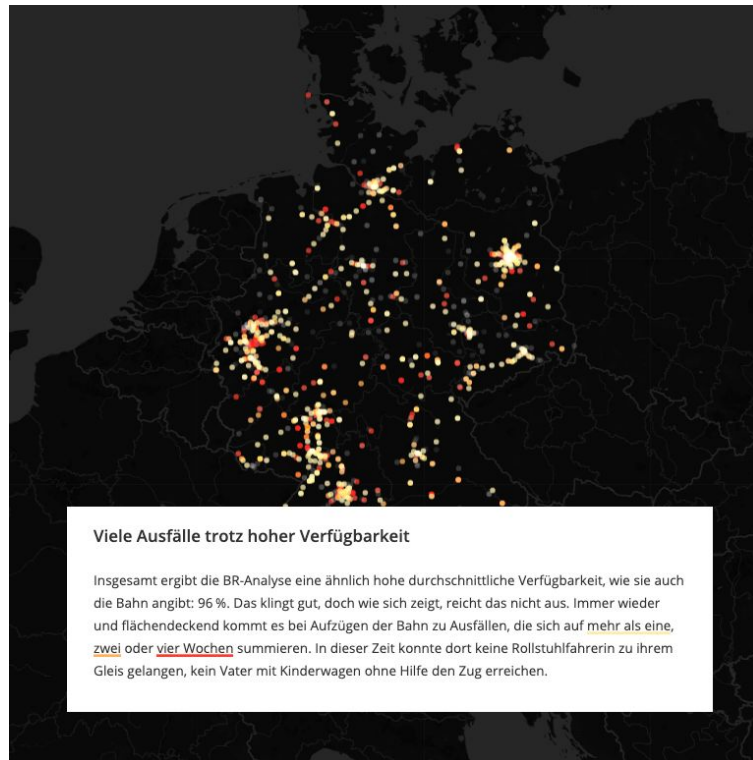
Datenvisualisierung

Bei der anderen Hälfte der Aufzüge schwankt die durchschnittliche maximale Verfügbarkeit pro Aufzug.



Immerhin erreichten 75% der Aufzüge ($n = 1545$), also **3 von 4 Aufzügen** eine maximale Verfügbarkeit von **95.4%** oder mehr.

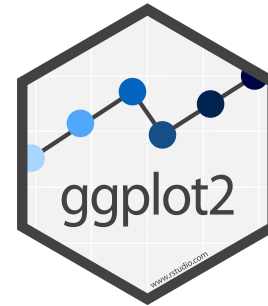
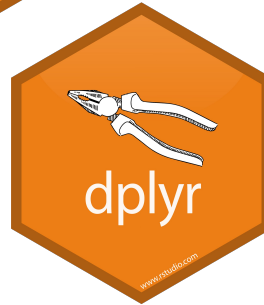
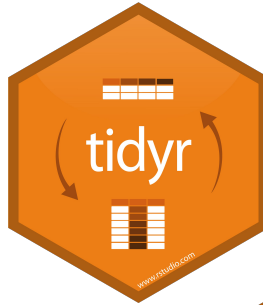
VS.



Vorteile von R

- Schnellen Einblick in die Daten bekommen
- Fehler und Ausreißer frühzeitig erkennen
- Einfach in bestehende Workflows zu integrieren
 - SVG-Export und Weiterverarbeitung in Illustrator
- Standardisierung von Grafiken durch Designvorlagen
 - [How the BBC works with graphics in R](#)
- Potential für Automatisierung von Grafiken

Workflow (ideal)



Scraping

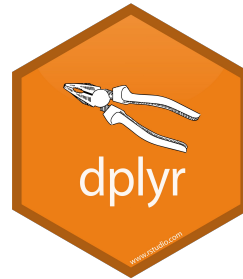
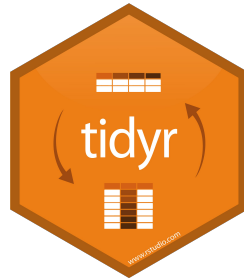
Cleaning & Analyse

Visualisierung

Workflow (realistisch)



Scraping



Cleaning & Analyse

Datawrapper

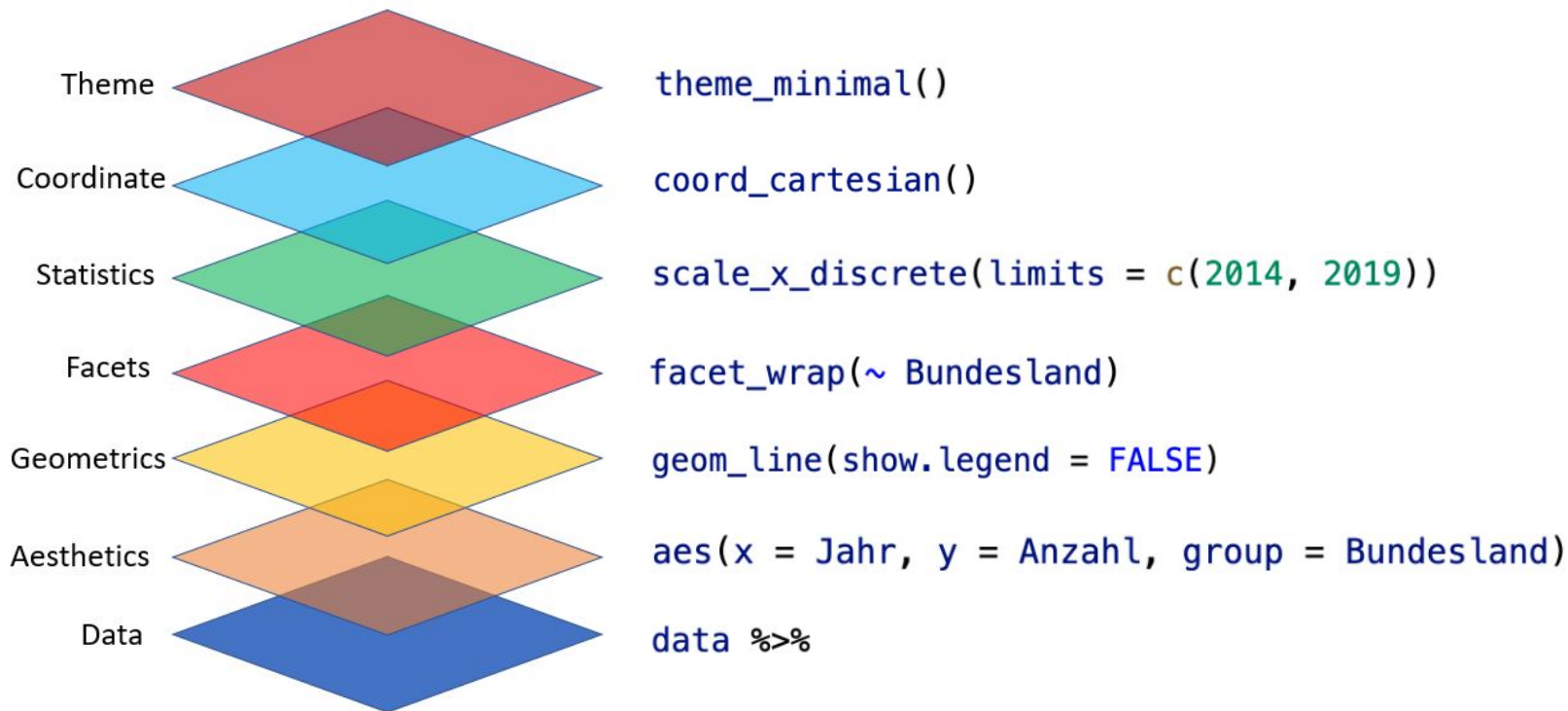


Visualisierung

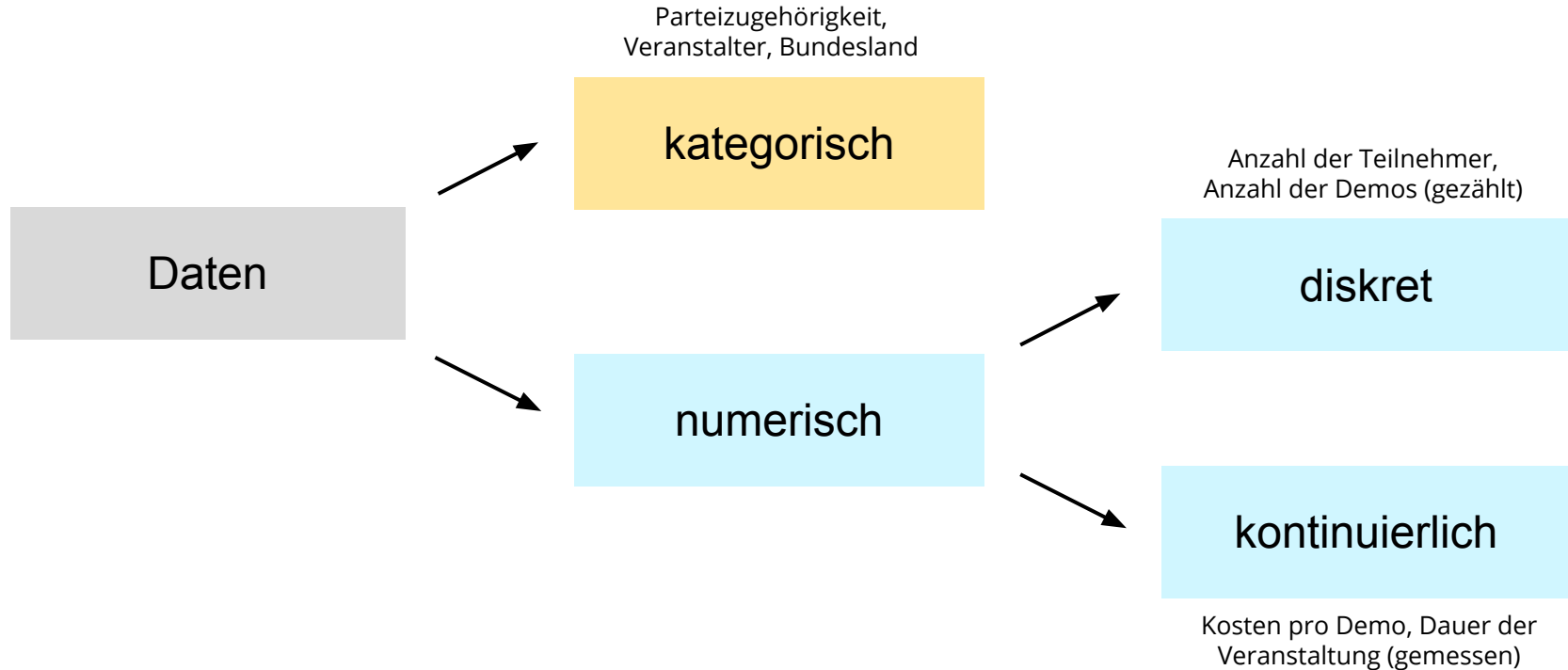
ggplot2 Layer

```
data %>%  
ggplot(aes(x = Jahr, y = Anzahl, group = Bundesland)) +  
  geom_line(show.legend = FALSE) +  
  facet_wrap(~ Bundesland) +  
  scale_x_discrete(limits = c(2014, 2019)) +  
  coord_cartesian() +  
  theme_minimal()
```


ggplot2 Layer

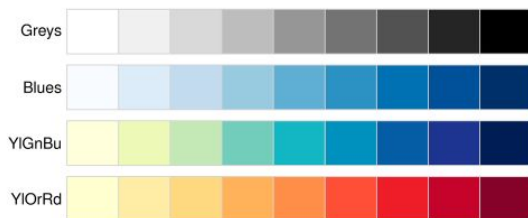


Skalen

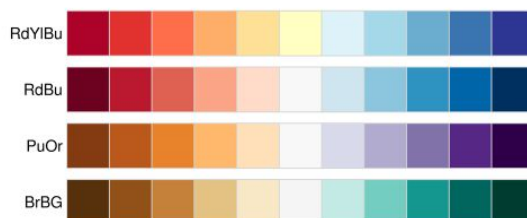


Farbskalen

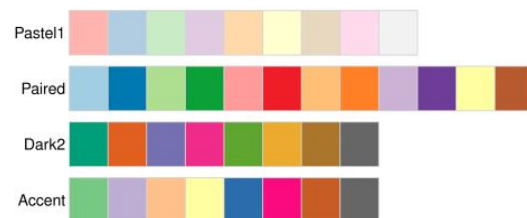
Sequential palettes



Diverging palettes



Categorical palettes



Sequentiell

- wenig → viel
- 0 % → 100 %
- 21 → 721

Divergierend

- negativ → neutral
→ positiv
- -100 % → 0 % → 100 %
- -25 → 0 → 75

Kategorisch

- A, B oder C
- CDU, SPD, Grüne, FDP, Linke

Quelle: [Better Figures](#)

Vielen Dank!

s.kuehne@onlinehome.de

stekhn.de

twitter.com/stekhn

linkedin.com/in/stekhn/

xing.com/profile/Steffen_Kuehne11