

Datenvisualisierung mit R

Übungsaufgabe vom 08.05.2025

Diese Datei enthält den Lösungsweg für die am Ende der letzten Sitzung gestellte Aufgabe. Sie finden unten den kommentierten Code und dann die Variante ohne Kommentare.

Zunächst fügen wir das Package `scales` hinzu, um die y-Achse in Prozentwerte umzuwandeln.

```
if(!require("pacman")) {install.packages("pacman");library(pacman)}
p_load(here, rio, scales, tidyverse)

fraktionen <- import(here("data", "bt_fraktionen.xlsx"))
```

Lösungsweg

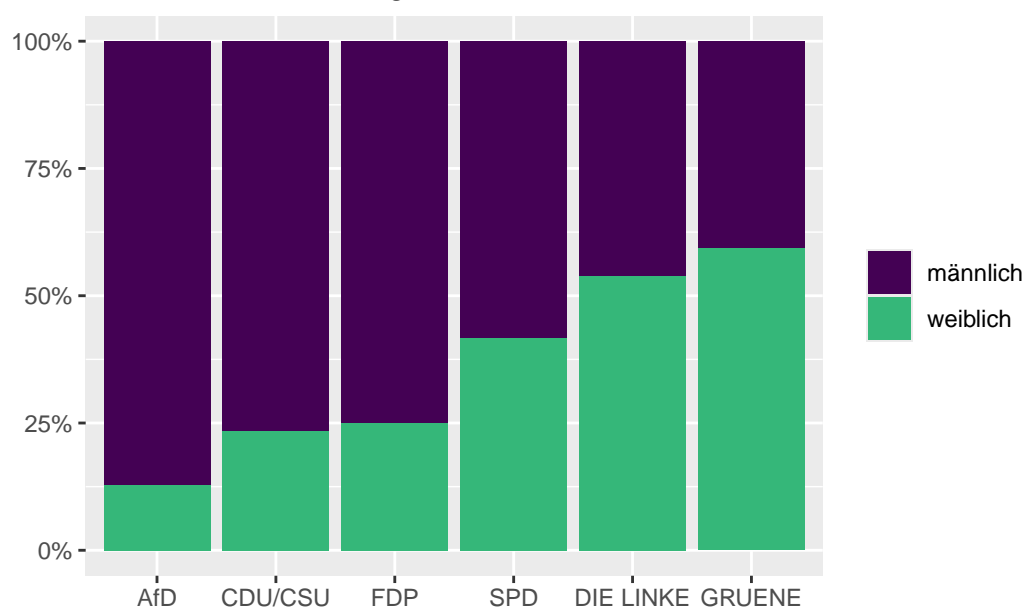
```
fraktionen %>%
  # Startet eine Abfolge von Arbeitsschritten mit dem Pipe-Operator (%>%),
  # der die Daten Schritt für Schritt transformiert.
  mutate(alle = m + w,
    # Erstellt eine neue Spalte `alle`,
    # die die Summe der männlichen (m) und weiblichen (w)
    # Abgeordneten enthält.
    weiblich = w / alle,
    # Erstellt eine neue Spalte `weiblich`,
    # die den Anteil der weiblichen Abgeordneten an der
    # Gesamtanzahl berechnet.
    männlich = m / alle) %>%
    # Erstellt eine neue Spalte `männlich`,
    # die den Anteil der männlichen Abgeordneten an der
    # Gesamtanzahl berechnet.
  select(Partei, weiblich, männlich) %>%
```

```

# wählt aus dem Datensatz nur die Spalten `Partei`,
# `weiblich` und `männlich` aus.
mutate(Partei = fct_reorder(Partei, weiblich)) %>%
# Ordnet die Kategorien in der Spalte `Partei` basierend auf dem
# Anteil der weiblichen Abgeordneten (aufsteigend).
pivot_longer(-c("Partei"),
  # Transformiert die Daten von einem "wide format" (breit)
  # in das "long format".
  # Alle Spalten außer `Partei` (angegeben mit `-c("Partei")`)
  # werden in die neuen Spalten aufgeteilt.
  names_to = "geschlecht",
  # Die ursprünglichen Spaltennamen (`weiblich` und `männlich`)
  # werden in die neue Spalte `geschlecht` geschrieben.
  values_to = "anteil") %>%
  # Die Werte der ursprünglichen Spalten werden in die
  # neue Spalte `anteil` geschrieben.
ggplot(aes(x = Partei, y = anteil, fill = geschlecht))+
# Erstellt ein ggplot-Objekt mit der ästhetischen Zuordnung:
# - `x = Partei`: Die Parteien werden auf der x-Achse dargestellt.
# - `y = anteil`: Der Anteil (weiblich/männlich) wird
# auf der y-Achse dargestellt.
# - `fill = geschlecht`: Farben werden basierend auf dem
# Geschlecht zugeordnet.
geom_col()+
# die geometrische Form ist ein Balkendiagramm.
scale_y_continuous(labels = percent)+ # hierfür benötigen wir scales package.
# Konfiguriert die y-Achse, sodass die Werte als Prozentsätze
# angezeigt werden.
scale_fill_manual(values = c("#440154", "#35b779"))+
# Legt die die Füllfarben fest:
# - `#440154` für eine Kategorie (männlich).
# - `#35b779` für die andere Kategorie (weiblich).
labs(x = NULL,
  # Entfernt die Beschriftung der x-Achse.
  y = NULL,
  # Entfernt die Beschriftung der y-Achse.
  fill = NULL,
  # Entfernt die Legendentitel (beschreibt die Füllfarben).
  title = "Anteil weiblicher Abgeordneter")

```

Anteil weiblicher Abgeordneter



Lösungsweg ohne Annotierung

```
fraktionen %>%
  mutate(alle = m + w,
         weiblich = w / alle,
         männlich = m / alle) %>%
  select(Partei, weiblich, männlich) %>%
  mutate(Partei = fct_reorder(Partei, weiblich)) %>%
  pivot_longer(-c("Partei"),
               names_to = "geschlecht",
               values_to = "anteil") %>%
  ggplot(aes(x = Partei, y = anteil, fill = geschlecht)) +
  geom_col() +
  scale_y_continuous(labels = percent) +
  scale_fill_manual(values = c("#440154", "#35b779")) +
  labs(x = NULL,
       y = NULL,
       fill = NULL,
       title = "Anteil weiblicher Abgeordneter")
```