

Hausaufgabe IV: Datentransformation, Themes und Tabellen

1. Vorbereitendes und Datentransformation

- a. Erstellen Sie ein neues Verzeichnis mit dem Namen `GDSR_5`, definieren Sie dieses Verzeichnis als Ihr Arbeitsverzeichnis und speichern dort ein neues R-Skript mit dem Namen `Hausaufgabe4_5.R`.
- b. Laden Sie die Pakete `tidyverse`, `readxl`, `gt` und `gtExtras`.
- c. Laden Sie den Datensatz `Starwars.xlsx` von StudOn oder GitHub herunter und lesen Sie ihn in RStudio ein.
- d. Entfernen Sie alle Beobachtungen mit fehlenden Informationen.
- e. Generieren Sie eine neue Variable mit dem Namen `Human` basierend auf der Variable `species` als Faktorvariable mit den Ausprägungen *Human* und *Non-human*.
- f. Das Excel-File ist ein Auszug aus dem Tidyverse-Datensatz `starwars`, d.h. Sie können nach Laden des Tidyverse Pakets über `?starwars` eine Datensatzbeschreibung erhalten.
- g. Berechnen Sie das Alter der Charaktere im Jahr 0 (Jahr der Schlacht von Yavin, d.h. im Jahr der Zerstörung des 1. Todessterns → siehe Datenbeschreibung!).

2. Abbildungen und Themes

- a. Erstellen Sie einen Scatterplot mit dem Gewicht der Charaktere auf der x-Achse und der Größe der Charaktere auf der y-Achse. Unterscheiden Sie im Scatterplot farblich zwischen `Human` und `Non-human`. Fügen Sie eine Regressionsgerade (lineares Modell) hinzu. Beschriften Sie die Achsen und fügen Sie einen Titel und Untertitel hinzu.
- b. Erstellen Sie ein Säulendiagramm mit den Mittelwerten für das Gewicht der Charaktere. Unterscheiden Sie nach `Human` und `Non-human`. Achten Sie auch hier auf eine vollständige Beschriftung Ihrer Abbildung.

- c. Definieren Sie ein eigenes vollständiges Theme (Schriftarten, Farbe etc.) und weisen Sie es beiden Abbildungen zu.

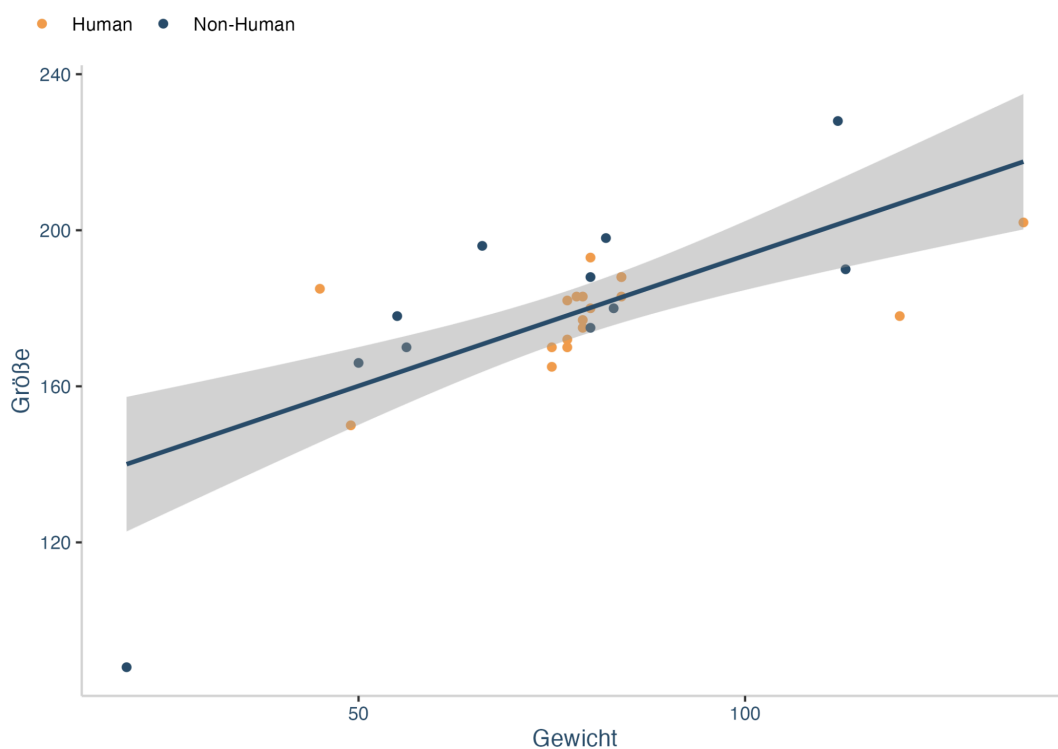
3. Tabelle

- Sortieren Sie die Daten absteigend nach der Größe und wählen Sie die 10 größten Charaktere aus.
- Erstellen Sie eine Tabelle aus den Variablen `name`, `height`, `mass`, `hair_color`, `species` und `homeworld`.
- Fügen Sie ein inline-Balkendiagramm für die Größe hinzu.
- Achten Sie auch hier auf eine aussagekräftige Beschriftung.
- Formatieren Sie die Tabelle analog des oben von Ihnen definierten Themes für ggplot. D.h. gleiche Schriftarten und Farben etc.

Bitte entwickeln Sie sich ein eigenes Theme. **Ziel ist ein einheitliches Format für Abbildungen und Tabellen.** Zur Inspiration: Eine Lösung könnte z.B. so aussehen:

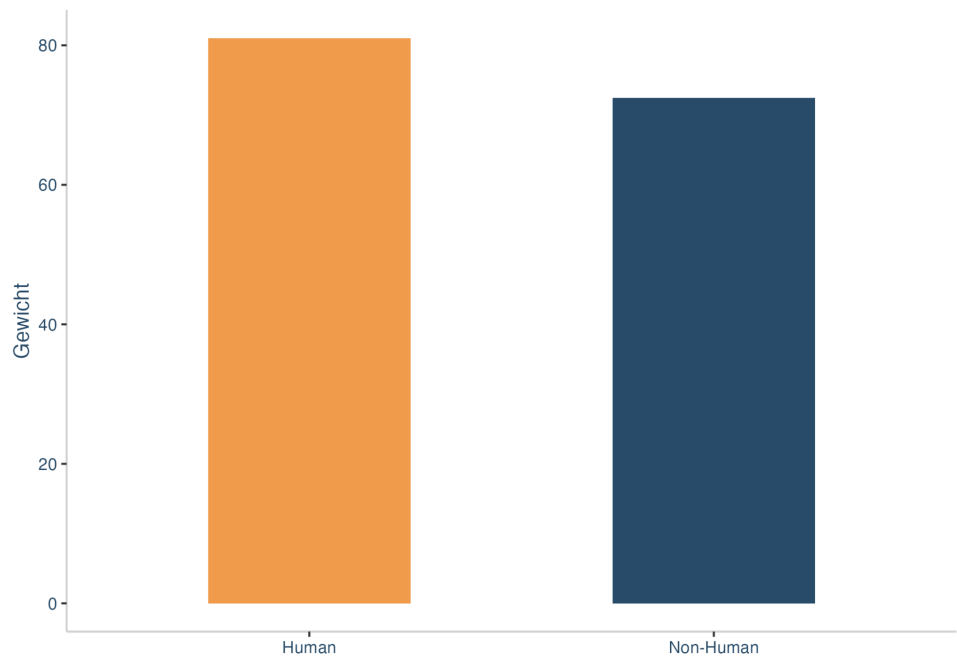
Der Zusammenhang zwischen Größe und Gewicht im Star Wars Universum

Daten aus dem Starwars-Datensatz



Gewichtsunterschied im Star Wars Universum zw. Humans und Non-Human

Daten aus dem Starwars-Datensatz



Die 10 größten Star Wars Charaktere

Daten aus dem Starwars-Datensatz

Name	Größe	Gewicht	Haarfarbe	Spezie	Heimatplanet
Chewbacca	228	112.0	brown	Wookiee	Kashyyyk
Ki-Adi-Mundi	198	82.0	white	Cerean	Cerea
Dooku	193	80.0	white	Human	Serenno
Bossk	190	113.0	none	Trandoshan	Trandosha
Mace Windu	188	84.0	none	Human	Haruun Kal
Jango Fett	183	79.0	black	Human	Concord Dawn
Obi-Wan Kenobi	182	77.0	auburn, white	Human	Stewjon
Darth Maul	175	80.0	none	Zabrak	Dathomir
Palpatine	170	75.0	grey	Human	Naboo
Luminara Unduli	170	56.2	black	Mirialan	Mirial

Note: Eigene Darstellung.