

# Sportinformatik

## WS 2019/2020 Übung 8

Abgabetermin: Mittwoch, 18. Dezember 2019, 8 Uhr

### Aufgabe 1: Football Data

10 Punkte

Auf <http://www.football-data.co.uk/germanym.php> gibt es die Spiele der Deutschen Bundesliga der letzten Jahre.

1. Schreiben Sie ein R-Skript, das die oben angegebene Seite durchsucht und alle dort verfügbaren CSV-Dateien in einen Unterordner herunterlädt (ohne explizite Direktlinks im Code). 5 Punkte
2. Schreiben Sie ein R-Skript, das alle heruntergeladenen CSV-Dateien einliest (ohne explizite Dateinamen im Code) und ein Data Frame mit den Daten erzeugt, das die folgenden Informationen erhält:  
Date, HomeTeam, AwayTeam, FTHG (Full Time Home Team Goals),  
FTAG (Full Time Away Team Goals), League 4 Punkte
3. Speichern Sie das Data Frame in einer Datei. 1 Punkte

---

### Aufgabe 2: Football Data 2

7 Punkte

Weitere Spieldaten deutscher Mannschaften sind in der Datei `football.db` (SQLite) verfügbar (Quelle).

1. Schreiben Sie ein R-Skript, das die Spieldaten (Tabelle "games") einliest, die IDs der Mannschaften durch den Namen ersetzt, Liga und Saison verknüpft und ein Data Frame mit den folgenden Informationen erzeugt:  
Date, HomeTeam, AwayTeam, FTHG (Full Time Home Team Goals),  
FTAG (Full Time Away Team Goals), League, Season  
(Beispiele für SQL, R analog z.B. mit `merge` oder `vlookup`) 6 Punkte
2. Speichern Sie das Data Frame in einer Datei. 1 Punkte

### **Aufgabe 3: Ranglisten im Fußball - 3-Punkte Ranking** **13 Punkte**

1. Schreiben Sie eine Funktion, die ein Data Frame mit den Spalten `Date`, `HomeTeam`, `AwayTeam`, `FTHG` (Full Time Home Team Goals), `FTAG` (Full Time Away Team Goals) übergeben bekommt und für alle Spiele im Zeitraum für alle Mannschaften das klassische 3-Punkte Ranking berechnet, dass heutzutage in der Bundesliga verwendet wird (Sonderregeln wie Torverhältnis können, müssen aber nicht beachtet werden.). Die Funktion soll sowohl den Punktstand am Ende des Zeitraums, als auch ein Data Frame mit den Punktenständen der Mannschaften während des gesamten Zeitraums zurückgeben. 6 Punkte
2. Nutzen Sie die ihre Funktion aus 3.1 um für die kompletten Datensätze aus Aufgabe 1 und Aufgabe 2 das Ranking zu bestimmen. Schreiben Sie eine Funktion, die den finalen Punktstand absteigend sortiert und übersichtlich in eine Datei schreibt. Erzeugen Sie jeweils die Dateien mit dem finalen Punktstand. 2 Punkte
3. Schreiben Sie eine Funktion, die Veränderungen im Ranking für die besten X Mannschaften visualisiert (Platzierung auf der y-Achse, Datum auf der x-Achse). Erzeugen Sie für beide Datensätze jeweils einen Plot für die 10 besten Mannschaften. 3 Punkte
4. Berechnen Sie nun für jeden Datensatz das Bundesliga Ranking für die Saison 2018/2019 und erzeugen Sie die Tabellen für das Saisonende. Vergleichen Sie ihr Ergebnis mit folgender Tabelle: 2 Punkte  
<http://www.sportal.de/fussball/bundesliga/tabelle/tabelle-saison-2018-2019>  
Stimmen die Ergebnisse überein?
5. Nutzen Sie ihre Funktion aus 3.1 um den Verlauf des Rankings der 10 besten Mannschaften der Saison 2018/2019 zu visualisieren. 2 Punkte

---

**Gesamtpunktzahl:**

**30 Punkte**