

Aufgabe 1

16.05.2022

Erstellen Sie für die Beantwortung dieser Aufgabe ein R-Skript, welches Sie unter dem Namen “Aufgabe1.R” abspeichern. Das Skript soll den R Code enthalten, der zum Lösen der einzelnen Teilaufgaben notwendig ist. Laden Sie das fertige Skript bis zur **Deadline (30.05.2022, 14:00)** im Moodle Ordner “04_R-Basics und Arbeiten mit Data Frames” hoch. Nach der Korrektur der Abgaben erhalten Sie als individuelles Feedback Ihre erreichte Punktzahl. Zusätzlich wird dann eine Musterlösung mit möglichen Lösungswegen, sowie eine Sammlung häufiger Fehler veröffentlicht.

Achtung! Das Aufgabenblatt hat 2 Seiten.

Aufgaben

Wesentlicher Bestandteil eines R-Skriptes ist der Header. Schreiben Sie zu Beginn des R-Skriptes einen Header der mindestens folgende Informationen enthält:

- Dateiname des Skriptes
- Input
- Output

Eine Vorlage für den Header finden Sie in der PDF Datei „Übung 9“ im Moodle Ordner „04_R-Basics und Arbeiten mit Data Frames“. Diesen können Sie kopieren und an ihr jetziges Skript anpassen. (1 P)

Sie haben ein Experiment durchgeführt, in dem 20 Versuchspersonen mitgemacht haben. Erstellen Sie aus den Daten des Experiments Vektoren in R:

1. Einen Vektor **Alter** der die folgenden Alter der Versuchspersonen enthält: 21, 18, 24, 19, 20, 22, 27, 20, 19, 23, 18, 19, 23, 20, 18, 29, 20, 24, 21, 20. (1P)
2. Einen Vektor **Geschlecht** der die Geschlechter der Versuchspersonen enthält. Die ersten 10 Versuchspersonen waren männlich, die 10 danach weiblich. Verwenden Sie zum Erstellen dieses Vektors die Funktion **rep()**. (1P)
3. Berechnen Sie den Mittelwert und die Standardabweichung des Alters Ihrer Versuchspersonen. (1P)
4. Ist die 5. Versuchsperson älter als die 11. Versuchsperson? Überprüfen Sie dies durch Indizieren der Werte und einen geeigneten Vergleichsoperator. (1P)
5. Wie viele Versuchspersonen sind unter 22 Jahre alt? (1P)

Für die folgenden Aufgaben lesen Sie zunächst den Datensatz **Erstsemesterbefragung_WiSe19-WiSe21.csv** in einen Data Frame **dat** ein und löschen Sie alle Zeilen mit fehlenden Werten. Einen Beispielcode hierfür finden Sie im Moodle-Ordner “04_R-Basics und Arbeiten mit Data Frames” in den Lösungsvorschlägen für Übung 11. Als Hilfsmittel mit Informationen zum Datensatz können sie außerdem auch das Dokument “Beschreibung_Erstsemesterbefragung_SoSe2022” verwenden.

Folgende Aufgaben sollen Sie mit Hilfe des Datensatzes beantworten:

6. Wie viele Teilnehmer/innen waren männlich (**m**), wie viele waren weiblich (**w**)? Tipp: Schauen Sie sich hierfür die Häufigkeitsverteilung von **Geschlecht** an. (1P)

7. Lassen Sie sich die Nutzungshäufigkeiten für Smartphones und Tablet PCs anzeigen. Was waren die jeweils am häufigsten angegebenen Werte für die Nutzungshäufigkeiten dieser Geräte für das Studium? (1P)
8. Erstellen Sie im Data Frame `dat` eine neue Spalte namens `Geschwindigkeit`, in der Sie für jede/n Teilnehmer/in das Verhältnis von `Entfernung` und `Fahrtdauer` einspeichern.
Tipp: Sie können die `Entfernung` und `Fahrtdauer` Spalten mit dem `$` Operator als Vektoren indizieren und dann `Entfernung` durch `Fahrtdauer` teilen. Das Ergebnis dieser Vektorenrechnung können Sie dann als `Geschwindigkeit` Spalte dem Data Frame `dat` hinzufügen. (1P)
9. Berechnen Sie den Mittelwert der `Geschwindigkeit` Spalte getrennt für Frauen und Männer. (1P)