

Testverfahren

Im Rahmen Ihres Testtermins werden wir Sie bitten, eine kleine Arbeitsprobe anzufertigen. Die Aufgabe ist (hoffentlich) sehr überschaubar. Dennoch möchten wir Ihnen Gelegenheit geben, sich entsprechend vorzubereiten, daher stellen wir die Aufgabe bereits jetzt dar, die wir Ihnen am Testtag vorlegen werden. Bitte bereiten Sie sich so gut wie möglich vor!

Bitte sehen Sie davon ab, uns inhaltliche Rückfragen zu der Programmierung zu senden. Das Testverfahren läuft standardisiert ab. Jede*r Teilnehmer*in soll unter den gleichen Ausgangsbedingungen getestet werden. Daher können wir keine Person bei der Testvorbereitung unsererseits unterstützen.

Aufgabenstellung:

Die folgende Aufgabenstellung werden Sie am Testtag ohne mitgeführte Hilfsmittel und bei laufender Webcam und freigeschaltetem Bildschirm bearbeiten:

Ziel ist, einen Datensatz in R einzulesen und diesen (automatisiert) auszuwerten. Der Datensatz enthält drei Variablen: „Studiengang“, „Qualität“ und „Verbleib“. Den entsprechenden Datensatz legen wir dieser Mail bei. Er wird Ihnen aber natürlich auch zum Vorstellungstermin vorgelegt. Zwei Variablen des Datensatzes („Qualität“ und „Verbleib“) sollen in Form einer grafischen Abbildung dargestellt werden. Diese Abbildungen sollen in PDFs gespeichert werden. Dabei soll die Variable „Studiengang“ des Datensatzes als Input für eine Schleife dienen, so dass je Studiengang nur die Fälle dieses Studiengangs ausgewählt und eine entsprechende Auswertung (PDF) mit zwei Abbildungen gespeichert wird.

Detaillierte Beschreibung:

Fachsprachliches: Einen Datensatz mag man sich als Tabelle vorstellen, die aus Zeilen und Spalten besteht. Menschen, die häufig mit Datensätzen umgehen, werden die Spalten oftmals als (Datensatz)-Variablen und die Zeilen als Fälle bezeichnen.

- Bitte öffnen Sie R-Studio und legen Sie ein Skript an.
- Fügen Sie als Kommentar: „Autor:“ und Ihren Namen in das Skript ein.
- Speichern Sie das Skript unter Ihrem Namen, bevor Sie mit der Bearbeitung fortfahren.

Inhalt des Scripts:

-Einlesen der csv-Datei „Test“ im Ordner „Arbeitsprobe“ auf Ihrem Desktop. (Dies ist Inhalt des Skripts – also bitte nicht den Fehler machen, die Datei über die R-Studio-Oberfläche zu öffnen!)

-Programmierung einer Schleife:

Die Werte der Datensatzvariablen Studiengang dienen als Input der Schleife.

Die Werte sind um Doppler zu bereinigen, so dass ein Set entsteht in dem jeder empirisch vorliegende Studiengang nur einmal enthalten ist. Es handelt sich um insgesamt $n=3$ Studiengänge, somit sollte die Schleife drei Durchläufe haben.

Je Schleifendurchlauf wird ein subset gebildet, dass nur die Fälle enthält, die dem gerade im Schleifendurchlauf aufgerufenen Studiengang entsprechen (Etwa: Nur Fälle, die BA BWL entsprechen)

Je Studiengang werden zwei barplots erzeugt, die jeweilig die einfachen Anzahlen darstellen (default-Ausgabe-Werte). Die barplots werden für die Datensatzvariablen „Qualität“ und „Verbleib“ erzeugt.

Diese beiden barplots werden jeweilig (also je Schleifendurchlauf) in einer PDF gespeichert.

Bearbeiten Sie die Aufgabe, so weit, wie Sie kommen.

Bitte speichern Sie das Skript auf dem Desktop ab und informieren Sie uns, sobald Sie die Aufgabe beendet haben.

Tipps:

Grundsätzlich: R-Studio verfügt über eine gute Hilfsfunktion: im Fenster unten rechts: Reiter „Help“

Die Funktionen `read.csv`, `unique`, `barplot`, `table`, `paste`, `as.character` und `pdf` könnten für die Lösung der Aufgabe hilfreich sein.

Es gibt sehr viele Wege, die Aufgabe zu lösen. Wir interessieren uns im anschließenden Gespräch dafür, warum Sie einen bestimmten Lösungsweg gewählt haben.