



Arbeitsblatt: Statistik mit R und RStudio

Übungsdatensatz: **Uebungsdatensatz_R_RStudio.csv / .xlsx**

Ziel	Du wendest die wichtigsten Arbeitsschritte einer einfachen Datenanalyse in R an: importieren, prüfen, auswerten, visualisieren und interpretieren.
Datei	Verwende den Datensatz Uebungsdatensatz_R_RStudio.csv. Das Trennzeichen ist das Semikolon ;
Arbeitsauftrag	Bearbeite die Aufgaben der Reihe nach. Notiere zu jeder Aufgabe den passenden R-Code und formuliere kurze Aussagen zu deinen Resultaten.

Aufgaben

Aufgabe 1: Daten importieren und ersten Überblick gewinnen

- ▶ Lies den Datensatz korrekt in R ein.
- ▶ Zeige die ersten sechs Zeilen.
- ▶ Prüfe die Struktur des Datensatzes und lasse eine Zusammenfassung ausgeben.

Hinweis: Nutze `read.csv()`, `head()`, `str()` und `summary()`.

Aufgabe 2: Fehlende Werte erkennen

- ▶ Ermittle, in welchen Variablen fehlende Werte vorkommen.
- ▶ Zeige alle Zeilen an, in denen mindestens ein Wert fehlt.
- ▶ Notiere kurz, warum fehlende Werte für Auswertungen wichtig sind.

Hinweis: Nutze `is.na()`, `colSums()` und `complete.cases()`.

Aufgabe 3: Datentypen anpassen

- ▶ Wandle die Variablen Geschlecht, Gruppe, Vorwissen, Abgabe_puenktlich und Bestanden in geeignete Faktorvariablen um.
- ▶ Prüfe anschliessend die Struktur des Datensatzes erneut.

Hinweis: Vorwissen kann sinnvoll geordnet werden: keins, wenig, mittel, gut.

Aufgabe 4: Deskriptive Statistik berechnen

- ▶ Berechne für Alter den Mittelwert, Median, die Standardabweichung und den Wertebereich.
- ▶ Berechne dieselben Kennzahlen für Note.
- ▶ Ermittle die Häufigkeiten für Geschlecht, Gruppe und Bestanden.

Hinweis: Achte bei Note auf fehlende Werte.

Aufgabe 5: Gruppen vergleichen

- ▶ Berechne die Durchschnittsnote pro Gruppe.
- ▶ Berechne die Durchschnittsnote pro Geschlecht.
- ▶ Berechne die durchschnittliche Lernzeit pro Vorwissen.

Hinweis: Nutze z. B. `aggregate()`.



Aufgabe 6: Diagramme erstellen

- ▶ Erstelle ein Balkendiagramm zu Bestanden.
- ▶ Erstelle ein Histogramm für die Variable Note.
- ▶ Erstelle einen Boxplot für Note nach Gruppe.
- ▶ Erstelle ein Streudiagramm für Lernzeit_Std_Woche und Note.

Hinweis: Formuliere zu jedem Diagramm einen kurzen Beobachtungssatz.

Aufgabe 7: Zusammenhänge prüfen

- ▶ Berechne die Korrelation zwischen Lernzeit_Std_Woche und Note.
- ▶ Berechne die Korrelation zwischen Motivation_1_10 und Note.
- ▶ Berechne die Korrelation zwischen Fehlstunden und Note.
- ▶ Ordne jeden Zusammenhang sprachlich ein: positiv, negativ, schwach oder mittel.

Hinweis: Nutze `cor(..., use = "complete.obs")`.

Aufgabe 8: Einfache statistische Tests anwenden

- ▶ Prüfe mit einem t-Test, ob sich die Noten von Gruppe A und Gruppe B signifikant unterscheiden.
- ▶ Prüfe mit einem Chi-Quadrat-Test, ob Gruppe und Bestanden zusammenhängen.
- ▶ Formuliere zu beiden Tests je einen kurzen Entscheidungssatz.

Hinweis: Entscheide jeweils mithilfe des p-Werts.

Aufgabe 9: Lineare Regression als Ausblick

- ▶ Berechne ein lineares Modell mit Note als Zielvariable und Lernzeit_Std_Woche als Einflussvariable.
- ▶ Notiere die Regressionsgleichung in Worten.
- ▶ Beschreibe kurz, was das Vorzeichen des Koeffizienten bedeutet.

Hinweis: Nutze `lm()` und `summary()`.

Aufgabe 10: Transferaufgabe

- ▶ Erstelle eine neue Variable Leistungsgruppe. Verwende die Kategorien "hoch" ab Note 5.0 und "normal" unter Note 5.0.
- ▶ Berechne die Häufigkeiten dieser neuen Variable.
- ▶ Formuliere zwei kurze Erkenntnisse aus dem Datensatz.

Hinweis: Nutze `ifelse()` für die neue Variable.