

In der Übung soll der Datensatz „statistikerumfrage“ von der Vorlesungs-Homepage verwendet werden. Es handelt sich dabei um einen bearbeiteten Ausschnitt der Daten, die im Rahmen der Veranstaltung „Deskriptive Statistik“ im Wintersemester 2014/15 und 2015/16 an der LMU München erhoben wurden. Der Datensatz enthält die folgenden Variablen :

<code>geschlecht:</code>	Geschlecht der Person: m=männlich, w=weiblich
<code>alltagwt_st:</code>	Durchschnittliche Zeit an einem Werktag, die mit dem Studium verbracht wird (in Stunden)
<code>alltagsa_st:</code>	Durchschnittliche Zeit an einem Samstag, die mit dem Studium verbracht wird (in Stunden)
<code>alltagso_st:</code>	Durchschnittliche Zeit an einem Sonntag, die mit dem Studium verbracht wird (in Stunden)
<code>alltag_st:</code>	Durchschnittliche Zeit pro Woche, die mit dem Studium verbracht wird (in Stunden)
<code>alltag_st_kat:</code>	Durchschnittliche Zeit pro Woche, die mit dem Studium verbracht wird (kategorisiert) (1=wenig (0-24 Stunden pro Woche), 2=mittel (25-40 Stunden pro Woche), 3=viel (mehr als 40 Stunden pro Woche))
<code>jahr:</code>	Jahr, in dem der Studierende an der Umfrage teilgenommen hat

### Vorbereitung:

- Legen Sie auf dem Z-Laufwerk (im Cip-Pool) einen Ordner für das Übungsblatt an. Legen Sie alle Dateien, mit denen Sie arbeiten, in diesem Ordner ab.
- Auf der Homepage finden Sie den Datensatz „statistikerumfrage“ in verschiedenen Formaten. Speichern Sie diese Datensätze in Ihrem Ordner.

Die Aufgaben sollten von jedem Studierenden alleine oder in Gruppen bis zum 12. bzw. 13. Dezember bearbeitet werden. Am 12. und 13. Dezember findet im Cip-Pool des Statistik-Instituts eine Fragestunde statt, bei der es die Möglichkeit gibt, Fragen zu den Programmen und Aufgaben zu stellen. Bitte halten Sie dafür Ihre Lösung bereit. Wir weisen ausdrücklich darauf hin, dass **keine Lösung** zu diesem Blatt zur Verfügung gestellt werden wird.

Lösen Sie die folgenden Aufgaben in allen drei vorgestellten statistischen Softwarepaketen **R**, **SPSS** und **SAS**. Verwenden Sie dazu die vorgegebenen Datensätze. Dies ist eine Vorbereitung auf die Hausübung. Bitte nutzen Sie diese Möglichkeit, um sich frühzeitig in die Programme einzuarbeiten und Fragen dazu im Vorfeld zu klären.

### **Aufgabe 1:**

Versuchen Sie, alle bereitgestellten Datensätze (1-6) in jedes Statistikprogramm einzulesen. Schauen Sie sich die Daten vorher gegebenenfalls in einem geeigneten Editor an, um das Spaltentrennzeichen und das Dezimaltrennzeichen zu ermitteln. Welche Datensätze lassen sich fehlerfrei in das jeweilige Statistikprogramm einlesen? Welche Formatierung ist mit welchem Statistikprogramm nicht kompatibel?

Arbeiten Sie ab jetzt mit einem Datensatz Ihrer Wahl weiter, welcher fehlerfrei in das jeweilige Statistikprogramm eingelesen werden konnte.

### **Aufgabe 2:**

Erstellen Sie eine Häufigkeitstabelle von `alltag_st_kat`. Berechnen Sie sowohl die absoluten als auch die relativen Häufigkeiten.

Wie viele Personen verbringen mehr als 40 Stunden pro Woche mit dem Studium?

Wieviel Prozent aller Personen sind das?

### **Aufgabe 3:**

Stellen Sie die absoluten Häufigkeiten von `alltag_st_kat` in einem Balkendiagramm dar.

### **Aufgabe 4:**

Erstellen Sie eine Kreuztabelle zwischen `geschlecht` und `alltag_st_kat`. (Die absoluten Häufigkeiten sind ausreichend.)

Welche Kombination kommt am seltensten vor?

### **Aufgabe 5:**

Stellen Sie die Variablen `geschlecht` und `alltag_st_kat` gemeinsam in einem Balkendiagramm dar (wahlweise gestapelte oder gruppierte Balken).

### **Aufgabe 6:**

Berechnen Sie zur Variable `alltag_st` den Mittelwert, die Standardabweichung und die 5-Punkte-Zusammenfassung.

Was ist die maximale Zeit, die pro Woche mit dem Studium verbracht wird? Was ist die maximale Zeit der 25 Prozent, die sich die geringste Zeit dem Studium widmen?

### **Aufgabe 7:**

Erstellen Sie einen Boxplot der Variable `alltag_st`.

### **Aufgabe 8:**

Erstellen Sie den Boxplot der Variable `alltag_st` getrennt nach Geschlecht.

Welches Geschlecht verbringt mehr Zeit mit dem Studium?

### **Aufgabe 9:**

Erstellen Sie ein Histogramm, das die Häufigkeitsverteilung der Variable `alltag_st` aufzeigt.