

## Lösung zu Kapitel 4: Beispiel 1

Erstellen Sie einen Data Frame namens `min.dat` und geben Sie Folgendes ein:

a	ges	gr	gew
21	m	181	69
35	w	173	58
829	m	171	75
2	e	166	69

R

```
> min.dat <- edit(data.frame())
```

Hier erscheint der Dateneditor für den Dataframe.

## Lösung zu Kapitel 4: Beispiel 2

Benennen Sie die Variablen in `Geschlecht`, `Alter`, `Grösse` und `Gewicht` um.

R

```
> colnames(min.dat) <- c("Alter", "Geschlecht", "Grösse", "Gewicht")
```

## Lösung zu Kapitel 4: Beispiel 3

Erstellen Sie eine Tabelle der Variable `Geschlecht` und prüfen Sie, ob Eingabefehler vorliegen. Der Wert `e` entstand durch einen Tippfehler, wir nehmen an, dass im Fragebogen *weiblich* stand. Ändern Sie diesen Wert auf `w`.

R

```
> table(min.dat$Geschlecht)
```

```
e m w  
1 2 1
```

R

```
> min.dat$Geschlecht[4] <- "w"
```

## Lösung zu Kapitel 4: Beispiel 4

Prüfen Sie die Anzahl der Levels von `Geschlecht` und entfernen Sie Levels, die nicht verwendet werden.

R

```
> levels(min.dat$Geschlecht)
```

```
[1] "e" "m" "w"
```

R

```
> min.dat$Geschlecht <- factor(min.dat$Geschlecht, labels = c("m",  
+ "w"))
```

## Lösung zu Kapitel 4: Beispiel 5

Benennen Sie die verbleibenden Levels von `Geschlecht` in *weiblich* für *w* und *männlich* für *m* um.

R

```
> min.dat$Geschlecht <- factor(min.dat$Geschlecht, labels = c("m",  
+ "w"))
```

## Lösung zu Kapitel 4: Beispiel 6

Kontrollieren Sie das Minimum und Maximum des Alters mit `range()`. Offensichtlich hat es grobe Eingabefehler gegeben, die Sie nun folgendermaßen ersetzen: Erstellen Sie eine Variable `auswahl`, die das Ergebnis der logischen Abfrage für *Alter geringer als 20 oder größer als 80* enthält. Wir kennen die wahren Werte nicht, also nutzen Sie die Variable `auswahl`, um die Beobachtungen, deren Altersangaben unter 20 oder über 80 liegen, *NA* zu setzen.

R

```
> range(min.dat$Alter)
```

```
[1] 2 829
```

R

```
> auswahl <- (min.dat$Alter < 20) | (min.dat$Alter > 80)
> min.dat$Alter[auswahl == T] <- NA
```

## Lösung zu Kapitel 4: Beispiel 7

Berechnen Sie den Body-Mass-Index (BMI) aller Personen und weisen Sie das Ergebnis der Variable **BMI** in Ihrem Data Frame zu. Geben Sie die Variable aus und runden Sie sie dabei auf eine Nachkommastelle.

R

```
> min.dat$BMI <- min.dat$Gewicht/(min.dat$Grösse/100)^2
> round(min.dat$BMI, 1)
```

```
[1] 21.1 19.4 25.6 25.0
```

## Lösung zu Kapitel 4: Beispiel 8

Verwenden Sie die Funktion `attach()`, um alle Variablen direkt aufrufen zu können.

R

```
> attach(min.dat)
> Alter
```

```
[1] 21 35 NA NA
```

R

```
> Grösse
```

```
[1] 181 173 171 166
```

R

```
> Gewicht
```

```
[1] 69 58 75 69
```

R

```
> Geschlecht
```

```
[1] m w m w  
Levels: m w
```

R

```
> BMI
```

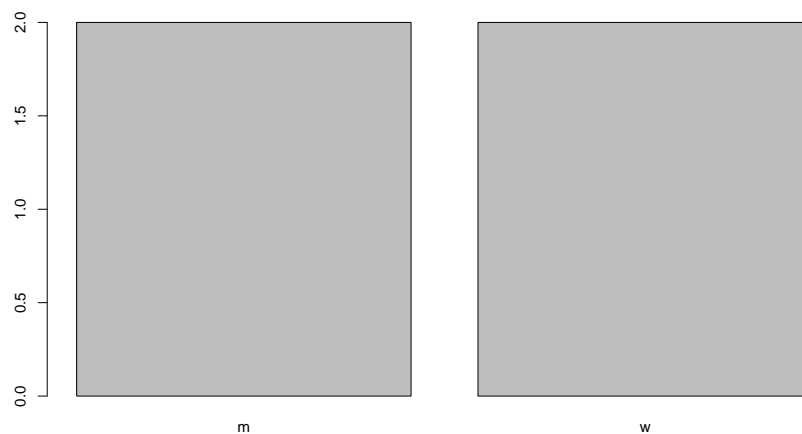
```
[1] 21.06163 19.37920 25.64892 25.03992
```

## Lösung zu [Kapitel 4: Beispiel 9](#)

Erstellen Sie ein Balkendiagramm für die Variable `Geschlecht`.

R

```
> plot(Geschlecht)
```



## Lösung zu Kapitel 4: Beispiel 10

Rekodieren Sie die Variable `Grösse` folgendermaßen: Alle Werte unter 170 nennen Sie `klein` und alle größer oder gleich 170 nennen Sie `gross`. Erstellen Sie hiermit eine neue Variable `Grösse_rec` in Ihrem Data Frame und tabulieren Sie sie.

**R**

```
> min.dat$Grösse_rec <- ifelse(Grösse < 170, "klein", "gross")
> table(min.dat$Grösse_rec)
```

```
gross klein
      3      1
```