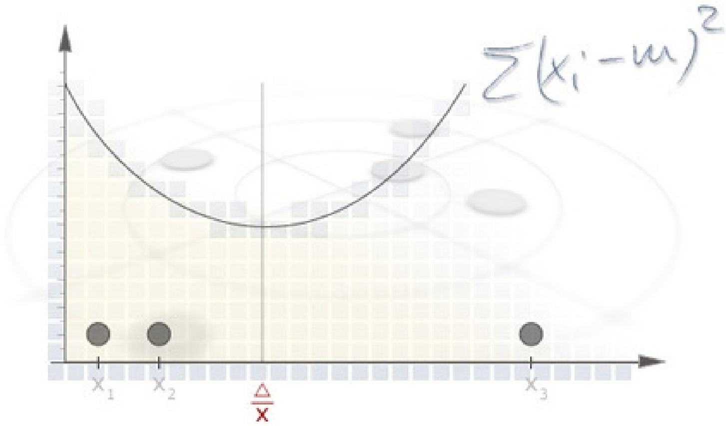
ARM - Arithmetisches Mittel



Überblick und Lernziele

Zunächst werden wir Ihnen Maßzahlen und Lagemaße als allgemeines Konzept vorstellen, dann beschäftigen wir uns ausführlich mit dem arithmetischen Mittel.

In dieser Lerneinheit betrachten wir das arithmetische Mittel.

Lernziele

Nach dem Durcharbeiten dieser Lerneinheit sollten Sie

- die Bedeutung und die Berechnung des arithmetischen Mittels erklären können,
- das gewichtete arithmetische Mittel berechnen können,
- wichtige Eigenschaften des arithmetischen Mittels: die Schwerpunkt-, die Minimum- und die Ersatzwerteigenschaft erklären können,
- das Verhalten des arithmetischen Mittels bei linearen Transformationen erklären können sowie
- das arithmetische Mittel gepoolter - also zusammengefasster Datensätze berechnen können,
- das arithmetische Mittel bei klassierten Daten berechnen können,
- das Verhalten des arithmetischen Mittels bei Ausreißern kennen.

In diesem Kapitel werden die folgenden Begriffe eingeführt:

- arithmetisches Mittel,
- gewichtetes arithmetisches Mittel,
- Schwerpunkt-, Minimum- und Ersatzwerteigenschaft

Gliederung der Lerneinheit

1. Einleitung
2. Maßzahlen univariater Häufigkeitsverteilungen
3. Eigenschaften und Interpretationen des arithmetischen Mittels
4. Arithmetisches Mittel bei klassierten Daten
5. Arithmetisches Mittel und Ausreißer

Zusammenfassung
Wissensüberprüfung
Übungen mit der Statistiksoftware R

Zeitbedarf und Umfang

Für die Durcharbeitung dieser Lerneinheit benötigen Sie ca. 150 Minuten und für die Übungen mit der Statistiksoftware **R** ca. 120 Minuten.

1 Einleitung

Mittelwerte und andere Lagemaße sind das Thema dieser Lerneinheit. Wer kennt das nicht, durchschnittliches Einkommen, Durchschnittsnote, überall treffen wir auf Mittelwerte.

Allerdings wird jetzt richtig gerechnet, nicht nur gezählt und angeordnet. Wieder ist der eigentliche Rechengang nicht aufregend kompliziert. Aber Sie sollten nach dem Durcharbeiten des Stoffes verstehen, weshalb bspw. die Stiftung Warentest den Median und nicht den Mittelwert als mittlere Preisangabe bei Marktübersichten einsetzt.

Also, erst mal den „Formelkram“ erledigen und dann nach dem Motto „Übung macht ...“ ab in die Praxis.

1.1 Maßzahlen univariater Häufigkeitsverteilungen

Eine Häufigkeitsverteilung beschreibt ein Merkmal zahlenmäßig vollständig. Wie wir schon gesehen haben, lassen sich charakteristische Eigenschaften der Lage und der Form von Häufigkeitsverteilungen durch grafische und tabellarische Darstellung erkennen. Oft ergeben sich dabei die folgenden Fragen:

- Wo liegt das Zentrum der Daten?
- Ist die Verteilung symmetrisch oder schief?
- Wie stark streuen Daten um das Zentrum?

Meist stehen diese Fragen auch beim Vergleich von Verteilungen im Vordergrund. Die deskriptive Statistik verwendet neben den Tabellen und Grafiken auch Maßzahlen zur Darstellung von Häufigkeitsverteilungen, wobei man versucht, mit möglichst wenigen Zahlen das Typische einer Verteilung zu charakterisieren.

Um eine Beobachtungsreihe knapp zu charakterisieren, sucht man nach Zahlenwerten, die alle Daten repräsentieren. Diese statistischen Kennwerte (Synonyme: Maßzahlen, Parameter, Statistiken, Indizes, Kennziffern) lassen sich in zwei Gruppen einteilen:

Tab.: Zwei Gruppen von statistischen Kennwerten

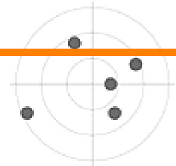
Statistischer Kennwert	Synonym	Erklärung
Lagemaße	Lageparameter, Lokalisationsmaße	Maße für zentrale Lage einer Verteilung
Streuungsmaße	Streuungsparameter, Dispersionsmaße	Maße, die die Variabilität der Verteilung kennzeichnen

Die Lagemaße und die Streuungsmaße gehören zur Charakterisierung univariater Häufigkeitsverteilungen.

1.2 Lagemaße univariater Häufigkeitsverteilungen

Die Lage der Daten eines Datensatzes ist eine der wichtigsten Eigenschaften von Datensätzen. Maßzahlen der Lage beschreiben das „Niveau“ der quantitativen Merkmalsausprägung einer Verteilung (das Zentrum einer Verteilung) oder die Lage des Datensatzes durch einen numerischen Wert.

Zu den Lagemaßen gehören unter anderem:



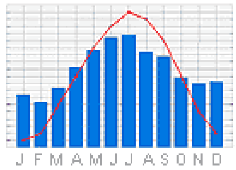
- arithmetischer Mittelwert,
- geometrischer Mittelwert,
- Median (Zentralwert),
- Modus (Dichtemittel),
- Quantile.

Das arithmetische Mittel erfordert eine Kardinalskala und das geometrische Mittel benötigt sogar eine Verhältnisskala. Der Median und Quantile benötigen mindestens eine Ordinalskala und der Modus nur eine Nominalskala.

2 Das arithmetische Mittel

Den arithmetischen Mittelwert einer Datenreihe berechnet man, indem man alle Werte addiert und durch ihre Anzahl dividiert. Mittelwerte geben einen summarischen Eindruck der Daten. Der kann sehr informativ sein.

Wir können uns alle vorstellen, was es heißt, dass das Jahresmittel der Temperatur in Sibirien -1,1 Grad Celsius, im Sahel 25 Grad Celsius beträgt. Mäuse und Elefanten haben unterschiedliche mittlere Gewichte. Ob die mittlere Füllmenge von Keksdosen mit dem angegebenen Inhalt übereinstimmt, führt zu Fragen des Verbraucherschutzes und der Qualitätsprüfung.



Wir können also wichtige Informationen in einer einzigen Zahl zusammenfassen. Das ist die eigentliche Aufgabe statistischer Maßzahlen. In der mathematischen Statistik gehört das arithmetische Mittel zu den am besten untersuchten Kenngrößen. Auch deshalb widmen wir dem arithmetischen Mittel eine ganze Lerneinheit. Aus Einzelbeobachtungen lässt sich der Mittelwert ganz einfach bestimmen:

Beispiel

Arithmetisches Mittel

Das arithmetische Mittel der Zahlen 7, 3, 8, 13, 10 ist:

$$\overline{x} = \frac{7+3+8+13+10}{5} = \frac{41}{5} = 8{,}2$$

Das summarische Vorgehen hat seinen Preis: Ändern wir nur einen Wert der Datenreihe, macht sich das sofort bemerkbar. Ersetzen wir z. B. in obiger Reihe die 10 durch den Wert 100 (könnte ein Tippfehler sein), dann ergibt sich 26,2 anstatt 8,2 als Mittelwert.
Das arithmetische Mittel ist empfindlich gegenüber extremen Werten, sogenannten Ausreißern.

Damit wir uns ganz genau verstehen, hier die Definition.

Definition

Das arithmetische Mittel

Sei X ein kardinales Merkmal. Dann heißt die Maßzahl $\overline{x}$ (sprich „x quer“) arithmetisches Mittel der Werte x_1 , ... x_n.

$$\overline{x} = \frac{1}{n}(x_1 + x_2 + \ldots + x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Synonym

Als Synonym wird häufig der Begriff Durchschnitt für das arithmetische Mittel benutzt. Genau genommen ist das unpräzise da es auch andere Durchschnittswerte gibt, wie das geometrische Mittel, das Sie hier im Abschnitt II Lage kennenlernen werden.

Anmerkung

Die Berechnung des arithmetischen Mittels ist bei ordinalen oder nominalen Merkmalen nicht sinnvoll. Mit Abweichungen von dieser Regel, wie bei der Berechnung des Notendurchschnitts, sollten Sie kritisch umgehen.

2.1 Das arithmetische Mittel - zwei Beispiele

Im ersten Beispiel berechnen wir das arithmetische Mittel der Mietpreise der Berliner Einzimmerwohnungen, die wir schon aus der Lerneinheit QBX kennen.

Beispiel

99 freie Einzimmerwohnungen

Für die Berechnung benutzen wir die bereits bekannte Tabelle aus Lerneinheit QBX:

Geordnete Mietpreise; Angebote der Einzimmerwohnungen mit der Angabe der Mietpreise.

231 247 251 254 256 256 263 277 288 301 302 306 322 333 333 333 340 348 352 354 359 363 364 368 368 369 373 373 374 375 378 380 382 384 384 384 384 388 388 395 395 395 395 405 405 409 413 417 420 423 431 432 433 444 446 448 454 461 473 474 475 477 481 500 500 501 506 506 507 511 512 519 535 538 546 546 552 553 564 581 589 609 609 628 666 681 691 699 727 742 832 916 947 1008 1151 1175 1190 1253 1429

Das arithmetische Mittel daraus:

$$\overline{x} = \frac{1}{99} \cdot (231 + 247 + \ldots + 1253 + 1429) = \frac{1}{99} \cdot 49005$$

$$= 495€$$

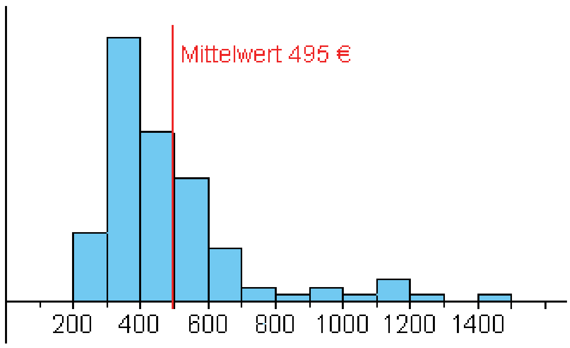


Abb.: Das arithmetische Mittel als Schwerpunkt des Histogramms

Sie sehen, das arithmetische Mittel liegt im Schwerpunkt des Histogramms.

Beispiel

Semesterpartys

Bekanntlich besteht ein Studentenleben nicht nur aus Lernen und Klausuren. Daher haben sich ein paar findige Studierende daran gemacht an ihrer Hochschule die Anzahl der Partys je Fachbereich im Semester zu zählen. Sicher nicht das wichtigste Kriterium bei der Wahl der Hochschule und des Studienfachs aber zumindest eine Überlegung wert.

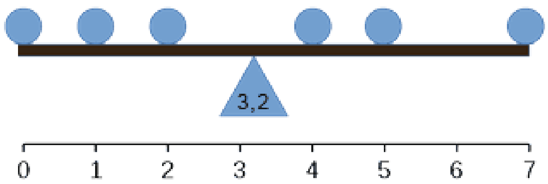
Sie sind zu folgendem Ergebnis gekommen:

Studienfach	Anzahl Semesterpartys
Mathematik – Physik – Chemie	4
Maschinenbau, Veranstaltungstechnik, Verfahrenstechnik	7
Elektrotechnik – Mechatronik – Optometrie	0
Bauingenieur- und Geoinformationswesen	1
Wirtschafts- und Gesellschaftswissenschaften	5
Architektur und Gebäudetechnik	2

Das arithmetische Mittel liegt bei

$$\overline{x} = \frac{1}{6} \cdot (4 + 7 + 0 + 1 + 5 + 2) \approx 3{,}2$$

Das bedeutet, der Schwerpunkt der Daten liegt bei rund 3,2.



2.2 Das arithmetische Mittel (Übungsaufgaben)

Berechnen

Übung ARM-01

Zahlen

Bestimmen Sie das arithmetische Mittel der Zahlen
2, 4, 6, 7, 3, 2, 4, 5, 5, 7, 6, 5, 7, 7, 8, 5, 4, 7, 3, 7, 8, 7, 4, 4, 7, 3, 4, 2, 7, 8.

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-02

Punktzahl

Die Punktzahlen, die ein Studierender bei sieben Klausuren erreichte, waren 87, 78, 97, 67, 91, 92 und 79.

Bestimmen Sie das arithmetische Mittel der Punktzahl.

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-03

Gehälter von acht Angestellten

Die Monatsgehälter von acht Angestellten betrugen 1500 €, 2400 €, 1800 €, 2000 €, 2100 €, 5500 €, 2200 €, 1900 €.

Bestimmen Sie das arithmetische Mittel ihrer Gehälter.

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-04

Gehälter in drei Abteilungen

Ein Betrieb besitzt drei Abteilungen L, M und N. Die Monatsgehälter der Angestellten für diese Abteilungen sind:

Abteilung L: 1600 €, 1850 €, 1550 €, 2200 €, 2100 €.
Abteilung M: 2500 €, 1950 €, 2250 €, 1700 €, 2300 €, 1500 €, 2100 €.
Abteilung N: 2400 €, 2350 €, 1875 €.

Bestimmen Sie die mittleren Monatsgehälter für die Abteilungen L, M, N und das mittlere Monatsgehalt aller Angestellten.

Lösung (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

2.3 Das gewichtete arithmetische Mittel (Beispiele)

Schon in der Grundschule haben wir alle gelernt, dass $5+5+5=3\cdot 5$ gilt. Diese elementare Rechnung können wir verwenden, wenn die gleichen Daten mehrfach in einem Datensatz auftreten.

Beispiel

Zahlen

Für den Datensatz in der Übung ARM-01 aus dem letzten Kapitel kommen die Zahlen 2, 4, 6, 7, 3, 5 und 8 jeweils mit den Häufigkeiten 3, 6, 2, 9, 3, 4 und 3 vor.

Zahlen	Häufigkeit
2	3
4	6
6	2
7	9
3	3
5	4
8	3

In diesem Fall können wir das arithmetische Mittel vereinfacht berechnen:

$$\overline{x} = \frac{(3\cdot 2)+(6\cdot 4)+(2\cdot 6)+(9\cdot 7)+(3\cdot 3)+(4\cdot 5)+(3\cdot 8)}{3+6+2+9+3+4+3}$$
$$= \frac{6+24+12+63+9+20+24}{30} = \frac{158}{30} = 5{,}27$$

Wir sprechen hier von einem gewichteten arithmetischen Mittel. Manchmal werden Daten auch aus anderen Gründen als der Häufigkeit ihres Auftretens gewichtet, wie im folgenden Beispiel.

Beispiel

Punktzahl

Wenn für den Datensatz in der Übung ARM-02 die Abschlussklausur für Wirtschaftsmathematik (79 Punkte) dreimal so hoch gewertet wird wie die Klausuren in den anderen Fächern (87, 78, 97, 67, 91, 92 Punkte), so berechnet sich die durchschnittliche Punktzahl wie folgt:

$$\overline{x} = \frac{(1)\cdot(87)+(1)\cdot(78)+(1)\cdot(97)+(1)\cdot(67)+(1)\cdot(91)+(1)\cdot(92)+(3)\cdot(79)}{1+1+1+1+1+1+3}$$
$$= \frac{749}{9} = 83{,}2$$

Diese durchschnittliche Punktzahl ist im Vergleich zur Lösung der Übung ARM-02 etwas gesunken: (83,2 < 84,4) Pech gehabt!

Den Einsatz von Gewichten bei der Mittelung werden Sie im Kapitel 4 dieser Lerneinheit noch genauer kennenlernen.

2.4 Das gewichtete arithmetische Mittel (Definition und Übung)

Anhand der beiden bisherigen Beispiele haben Sie gesehen wie das gewichtete arithmetische Mittel berechnet wird. Hier folgt nun die Definition des gewichteten arithmetischen Mittels und damit eine allgemeine Berechnungsvorschrift:

Definition

Gewichtetes arithmetisches Mittel

Sind die **absoluten Häufigkeiten** $n_1, \ldots, n_k$ der Daten $x_1, \ldots, x_k$ gegeben, mit $n_{ges} = n_1 + \ldots + n_k$ berechnet sich das gewichtete arithmetische Mittel wie folgt:

$$\overline{x} = \frac{1}{n_{ges}} \sum_{i=1}^{k} n_i \cdot x_i$$

$h_i = \frac{n_i}{n_{ges}}.$

Sind dagegen die relativen Häufigkeiten $h_1, \ldots, h_k$ gegeben, berechnet sich das gewichtete arithmetische Mittel so:

$$\overline{x} = \sum_{i=1}^{k} h_i \cdot x_i$$

Berechnen

Übung ARM-05

1000 Zahlen

Von 1000 Zahlen waren 150 die 3, 50 die 2, 400 die 7, 300 die 5 und der Rest die 6. Bestimmen Sie das arithmetische Mittel der Zahlen.

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

3 Eigenschaften und Interpretationen des arithmetischen Mittels

Wenn Sie das arithmetische Mittel einsetzen wollen, sollten Sie wissen, warum Sie das tun. Deshalb machen wir Sie mit einigen Eigenschaften vertraut, auf die Sie Ihre Entscheidung stützen können. Im Grunde verfahren Sie genau wie im Alltag.

Worauf kommt es an?

- **Tempo**:
  Sie wählen das Verkehrsmittel, mit dem Sie am schnellsten zum Ziel kommen,
- **Durchmesser**:
  Sie wählen den zum Dübel passenden Bohrer.
- **Aroma**:
  Sie wählen den Deckel, der auf den Topf passt, dann verfliegt der Duft nicht.

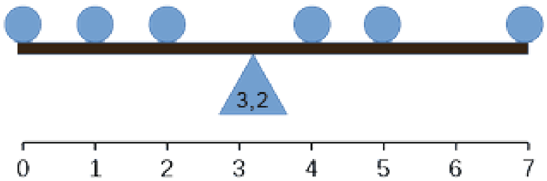
Also, wenn Sie sich mit den Eigenschaften des Mittelwerts vertraut gemacht haben, können Sie besser beurteilen, ob das arithmetische Mittel als Maßzahl für Ihre Fragestellung geeignet ist.

Auf den folgenden Seiten werden verschiedene Eigenschaften des arithmetischen Mittels beschrieben:

- Schwerpunkteigenschaft
- Minimumeigenschaft
- Verträglichkeit mit linearen Transformationen
- Ersatzwerteigenschaft
- Arithmetisches Mittel zusammengefasster, gepoolter Daten

3.1 Schwerpunkteigenschaft

Das arithmetische Mittel tariert die Daten aus, d. h. es bringt sie wie auf einer Waage ins Gleichgewicht.



$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \iff n\cdot\bar{x} = \sum_{i=1}^{n} x_i$

$\iff \left(\sum_{i=1}^{n} x_i\right) - n\cdot\bar{x} = 0$

Dieses Charakteristikum des arithmetischen Mittels nennt man Schwerpunkteigenschaft. Sie haben es bereits im Kap. 2.1 im Beispiel *„Semesterparty“* kennengelernt. Hier werden wir uns etwas intensiver damit beschäftigen.

Schwerpunkt

Die Summe der Abweichungen der Einzelwerte vom arithmetischen Mittel ist stets gleich null.

$$\sum_{i=1}^{n}(x_i - \overline{x}) = 0$$

Rechnen wir das zunächst an einem ganz kleinen Zahlenbeispiel nach:

$x_1, x_2, x_3 = 1, 2, 3$ daraus folgt $\overline{x} = 2$

Prüfen wir nun die Schwerpunkteigenschaft: (1-2)+(2-2)+(3-2)=0 Stimmt!

Beispiel

Gehäufte Werte: Schwerpunkt

Das arithmetische Mittel der Werte 3, 0, 5, 0, 25, 0, 13, 3, 28, 13 kann man als ungewichtetes und als gewichtetes Mittel berechnen.

a) Ungewichtetes arithmetisches Mittel:

$$\overline{x} = \frac{3+0+5+0+25+0+13+3+28+13}{10} = \frac{90}{10} = 9$$

b) Gewichtetes arithmetisches Mittel:

Methode 1:

$$\overline{x} = \frac{(2\cdot 3)+(3\cdot 0)+(1\cdot 5)+(1\cdot 25)+(2\cdot 13)+(1\cdot 28)}{2+3+1+1+2+1}$$

$$= \frac{6+5+25+26+28}{10} = \frac{90}{10} = 9$$

Methode 2:

Man kann auch das gewichtete arithmetische Mittel aus einer Tabelle berechnen. In der folgenden Tabelle bestimmt man in der Spalte 4 die Größen $x_i \times n_i$, deren Summe durch die Summe n_i (Spalte 3) geteilt wird:

$$\overline{x} = \frac{90}{10} = 9$$

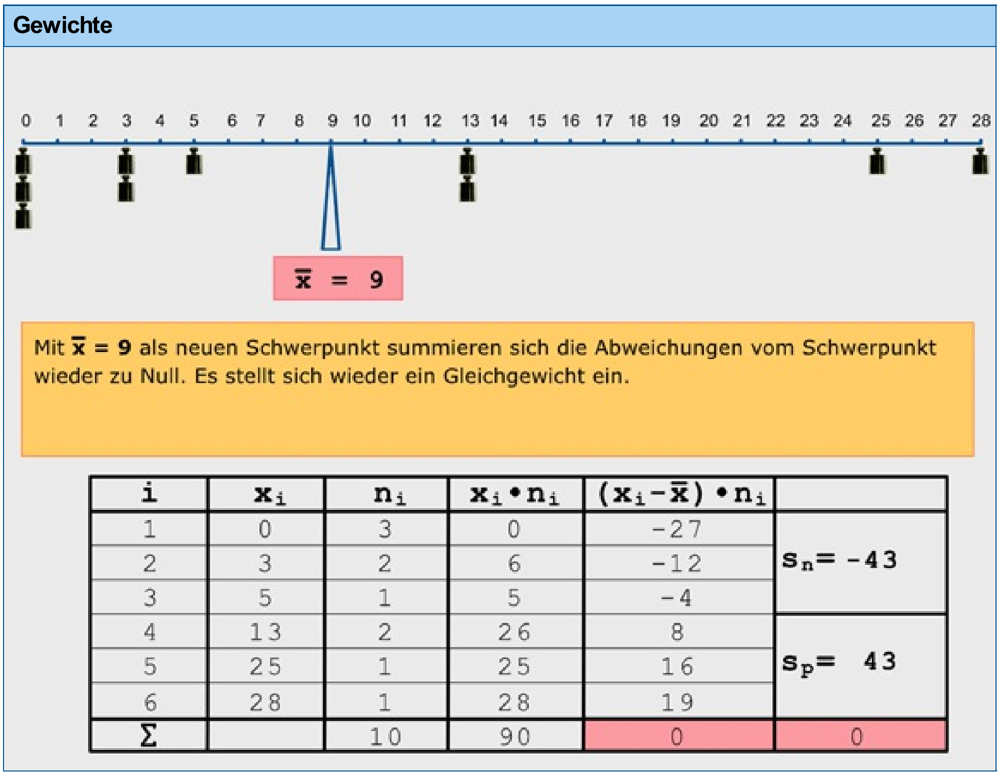
i	x_i	n_i	$x_i \cdot n_i$	$x_i - \overline{x}$	$(x_i - \overline{x}) \cdot n_i$	S
1	0	3	0	-9	-27	
2	3	2	6	-6	-12	S_n = - 43
3	5	1	5	-4	-4	
4	13	2	26	4	8	S_p = + 43
5	25	1	25	16	16	
6	28	1	28	19	19	
∑	-	10	90	-	0	

Tab.: Tabelle zum Berechnen von gewichteten arithmetischen Mitteln

Die Tabelle macht deutlich, warum hier vom „Schwerpunkt" gesprochen wird. Die Summen der negativen (S_n) und positiven (S_p) Abweichungen der Merkmalsausprägungen des arithmetischen Mittel sind gleich: S_n = - 43 und S_p = + 43. Die Gleichheit von S_n und S_p bedeutet, dass sich negative und positive Abweichungen ausgleichen.

In der folgenden Animation erscheinen die Größen n_i als Gewichte () und die Größen x_i als Positionen an den Hebelarmen an einer Balkenwaage. Wird diese im Punkt $\overline{x}$ = 9 unterstützt, dann befindet sich die Waage im Gleichgewicht.

Diashow



i	x_i	n_i	$x_i \cdot n_i$	$(x_i - \overline{x}) \cdot n_i$	
1	0	3	0	-27	
2	3	2	6	-12	S_n= -43
3	5	1	5	-4	
4	13	2	26	8	
5	25	1	25	16	S_p= 43
6	28	1	28	19	
∑		10	90	0	0

Verfolgen Sie, wie sich in der Animation durch Verschieben der Gewichte die positiven und negativen Abweichungen ändern.

3.2 Minimumeigenschaft

Minimum

Es gilt:

$$\sum_{i=1}^{n}(x_i-\overline{x})^2 < \sum_{i=1}^{n}(x_i-c)^2$$

$\overline{x} = \arg\min_{c} \sum_{i=1}^{n} (x_i - c)^2$

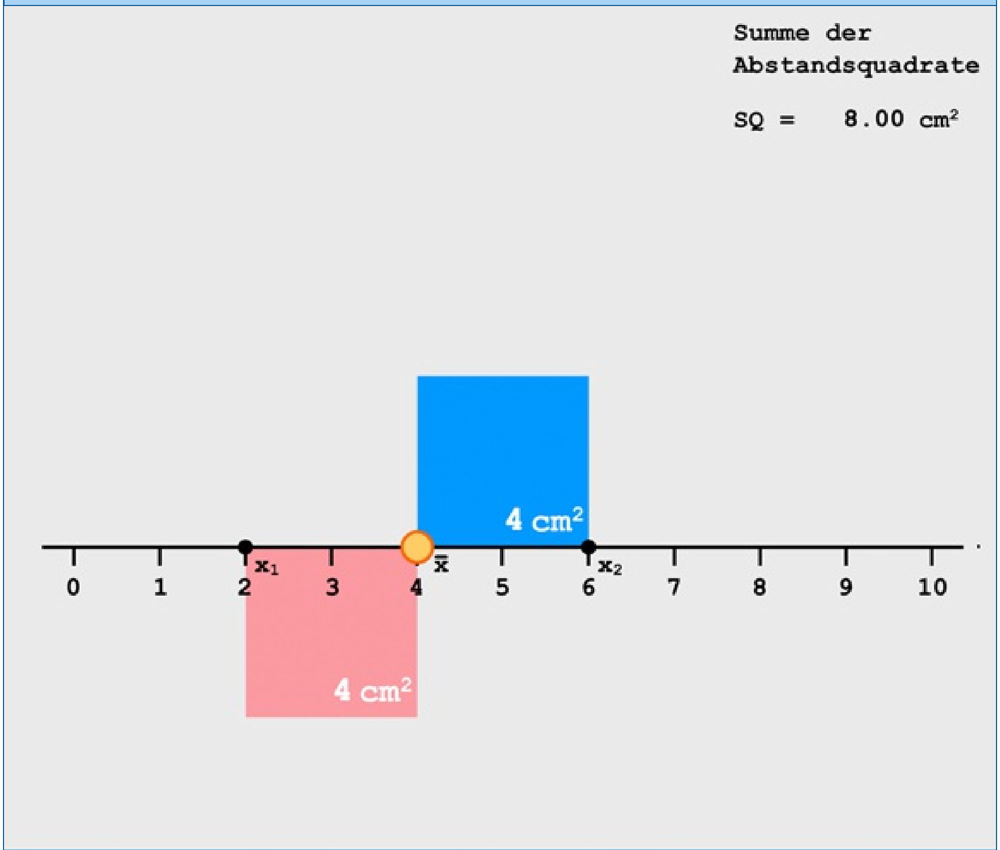
$\forall c \neq \overline{x}$

Nur für das arithmetische Mittel ist die Summe der Abstandsquadrate minimal. Für jede andere Zahl c werden die Abstandsquadrate größer. Die nachfolgende Interaktion demonstriert das sehr schön für zwei Datenwerte. Aber die Minimumeigenschaft gilt natürlich für eine beliebige Zahl an Daten.

In der Interaktion können Sie per Drag und Drop beispielhaft die Summen der Abstandsquadrate anzeigen lassen.

Interaktion

Abstandsquadrate



Beispiele ersetzen nie einen mathematischen Beweis. Sie finden den Beweis zur Minimumseigenschaft in den meisten Lehrbüchern zur Statistik. Wir verzichten hier darauf.

3.3 Verträglichkeit mit linearen Transformationen

Ein wenig Mathematik kann nicht schaden. Manchmal wird das Leben dadurch wesentlich leichter. Muss man Mittelwerte immer neu bestimmen, wenn die Daten transformiert werden?

Lineare Transformation

Werden die Daten $x_1, \dots, x_n$ linear transformiert mit $y_i = a + bx_i, i = 1, \dots, n$dann ist das arithmetische Mittel der transformierten Daten gleich dem transformierten arithmetischen Mittel: $\overline{y} = a + b\overline{x}$

Mittelwertberechnung und lineare Transformationen sind vertauschbar.

Beispiel

Probieren wir das einmal an einem kleinen Beispiel.

1. x-Werte: $1, \dots, 5$

   $\overline{x} = \frac{1}{5}(1 + 2 + 3 + 4 + 5) = 3$
2. $a = 4, \quad b = 3, \quad y_i = 4 + 3 \cdot x_i, \quad i = 1, \dots, 5,$
3. y-Werte:
   $y_1 = 4 + 3 \cdot 1 = 7$
   $y_2 = 4 + 3 \cdot 2 = 10$
   $y_3 = 4 + 3 \cdot 3 = 13$
   $y_4 = 4 + 3 \cdot 4 = 16$
   $y_5 = 4 + 3 \cdot 5 = 19$

   $\overline{y} = \frac{1}{5}(7 + 10 + 13 + 16 + 19) = 13$
4. $\overline{y} = a + b\overline{x} = 4 + 3 \cdot 3 = 13$ ✔

Achtung

Für nicht lineare Transformationen gilt die Vertauschbarkeit **nicht**!

3.4 Ersatzwerteigenschaft

Die Summe von n Werten $x_i, \ldots, x_n$ entspricht n mal dem arithmetischen Mittel dieser Werte:

$$\boxed{\sum_{i=1}^{n} x_i = n \cdot \overline{x}}$$

Diese Eigenschaft ergibt sich direkt aus der Definition des arithmetischen Mittels. Zeigen wir die Ersatzwerteigenschaft an einem kleinen Beispiel.

Beispiel

Ersatzwerteigenschaft

Es liegen uns die folgenden 10 Daten vor:
$x_1, \ldots, x_{10} = 1, 3, 7, 11, 9, 3, 5, 12, 14, 5$
Deren Summe ist:

$$\sum_{i=1}^{10} x_i = 1 + 3 + 7 + 11 + 9 + 3 + 5 + 12 + 14 + 5 = 70$$

Das arithmetische Mittel ist gleich

$$\overline{x} = \frac{1}{10} \sum_{i=1}^{10} x_i = \frac{1}{10} \cdot 70 = 7$$

Die Summe der Daten, nämlich 70, ist gleich der Anzahl der Daten n=10 multipliziert mit dem arithmetischen Mittel $\overline{x} = 7$.

$$\sum_{i=1}^{10} x_i = 10 \cdot 7 = 70$$

Das genau sagt die Ersatzwerteigenschaft.

Dank der Ersatzwerteigenschaft ist es möglich die Summe der Daten mit $n \cdot \overline{x}$ zu berechnen ohne die ursprünglichen Daten zu kennen – vorausgesetzt das Arithmetische Mittel ist bekannt. Dies nutzen wir in der Praxis, um das arithmetische Mittel zusammengefasster Datensätze zu berechnen.

3.5 Das arithmetische Mittel zusammengefasster Datensätze

Sind die arithmetischen Mittelwerte $\overline{x}_1, \ldots, \overline{x}_k$ und die Umfänge $n_1, \ldots, n_k$ von k Datensätzen gegeben, so ist das arithmetische Mittel des zusammengefassten Datensatzes aller Daten

$$\overline{x} = \frac{n_1\overline{x}_1 + n_2\overline{x}_2 + \ldots + n_k\overline{x}_k}{n_1 + n_2 + \ldots + n_k} = \frac{1}{n}\sum_{i=1}^{k} n_i\overline{x}_i$$

Dies folgt direkt aus der Ersatzwerteigenschaft.

zusammengefasst = gepoolt

Man kann also das arithmetische Mittel zusammengefasster – man sagt auch **gepoolter** – Datensätze berechnen ohne die ursprünglichen Daten zu kennen**.** Es reicht die einzelnen arithmetischen Mittelwerte und die jeweiligen, zugehörigen Stichprobenumfänge zu kennen. Diese Tatsache wird häufig bei Umfragen und Hochrechnungen verwendet und wird auch als Aggregationseigenschaft bezeichnet.

Beispiel

Nimers Restaurant

pixabay.com

Nimer kauft auf dem Großmarkt frische Lebensmittel für die leckeren Salate, die sein Restaurant bekannt gemacht haben. Bei 5 Händlern kauft er Gemüse mit einem durchschnittlichen Preis von 2,50 Euro pro Kilo, bei weiteren 4 Händlern verschiedene Blattsalate mit einem durchschnittlichen Kilopreis von 0,55 Euro und schließlich bei 3 Händlern Obst zu einem Durchschnittspreis von 1,90 Euro.

Wie groß ist nun der durchschnittliche Kilopreis seiner frischen Salatzutaten?

$$\overline{x} = \frac{5 \cdot 2{,}5 + 4 \cdot 0{,}5 + 3 \cdot 1{,}90}{5 + 4 + 3} = 1{,}70$$

Er muss also bei seinen Salaten mit einem durchschnittlichen Kilopreis der frischen Zutaten von 1,70 Euro kalkulieren.

Berechnen

Übung ARM-06

Pizza bestellen

Eine Lerngruppe von 10 Studierenden wirft Geld zusammen um Pizza und Getränke zu bestellen. Im Durchschnitt gibt jeder 16,5 Euro.

1. Eine Studentin kommt etwas später und gibt noch einmal 22 Euro dazu. Wieviel tragen die 11 Studierenden jetzt im Durchschnitt bei?
2. Wieviel hätte die verspätete Studentin beitragen müssen um den Durchschnitt um genau einen Euro zu heben? .

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-07

Mittleres Jahreseinkommen

Wenn das mittlere Jahreseinkommen von Arbeitern und Arbeiterinnen aus Treptow bzw. Schöneberg in Berlin 26.000 € bzw. 30.000 € beträgt, ist dann das mittlere Jahreseinkommen für beide Bezirke zusammen gleich 28.000 €?

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-08

Mittleres Monatsgehalt

Berechnen Sie die Mittelwerte der Monatsgehälter in den einzelnen Abteilungen und den Gesamtmittelwert der Monatsgehälter in der Firma:

Abteilung L: 1600 €, 1850 €, 1550 €, 2200 €, 2100 €.
Abteilung M: 2500 €, 1950 €, 2250 €, 1700 €, 2300 €, 1500 €, 2100 €.
Abteilung N: 2400 €, 2350 €, 1875 €.

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

3.6 Mangelnde Robustheit - Verzerrung durch Ausreißer

Das arithmetische Mittel reagiert empfindlich auf extreme Werte. Das können beispielsweise Mess- oder Übertragungsfehler sein aber auch außergewöhnlich hohe oder niedrige Datenwerte. Anders ausgedrückt, das arithmetische Mittel ist nicht **robust** gegen Ausreißer. Wir wollen dies einmal anhand eines Beispiels zeigen.

Beispiel

Ausreißer im Reiseunternehmen

Das Reiseunternehmen Hagen-Alpintours aus dem Allgäu bietet weltweit Wander- und Skireisen an. Am Beginn des neuen Jahres schaut sich das Betreiberehepaar die Zahlen des Vorjahres an. Für die strategische Ausrichtung der nächsten Jahre wollen sie wissen wer der typische Kunde oder Kundin ist und in welcher Höhe die durchschnittliche Buchung liegt.

Im Vorjahr wurden 762 Reisen an Direktkunden, Familien, Singles, Paare etc. in Rechnung gestellt in einem Volumen von insgesamt 2.539.469 Euro.

Die Daten können Sie sich hier als CSV-Datei runterladen und zusammen mit der R-Datei die Rechnungen mit R nachvollziehen.

Alpintours.R

Alpintours.csv (2 KB)

Die Auswertung ergab, dass die durchschnittliche Buchung bei 3.333 Euro liegt. Das scheint dem erfahrenen Unternehmerpaar zu hoch. Aber wie kommt es dann zu diesem Wert?

Bei Durchsicht der Daten fällt auf, dass einige Buchungen von Skiclubs und Vereinen stammten und deren Rechnungen sich über mehrere 10.000 Euro beliefen. Anhand des Boxplots erkennt man deutlich die Ausreißer, die zum Teil weit über 10.000 Euro liegen während der Großteil der Buchungen deutlich unter 5.000 Euro liegt.

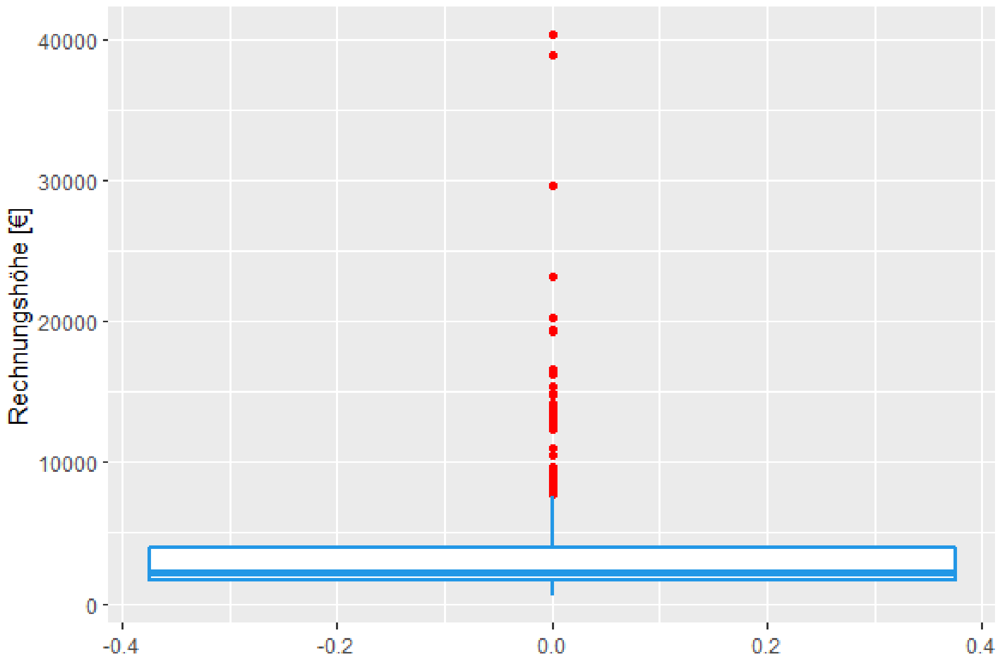


Abb.: Boxplot Reisebuchungen mit Ausreißern (rot)

© Quelle: Daten Hagen Alpintours ,https://www.welt-weit-wandern.de

Werden die Ausreißer aus dem Datensatz entfernt - bzw. bei der Berechnung nicht berücksichtigt, ergibt sich ein arithmetisches Mittel von 2.674 €. Ein Wert, der um mehr als 600 € kleiner und für die strategische Planung deutlich besser geeignet ist.

4 Arithmetisches Mittel bei klassierten Daten

Wir haben schon in der Lerneinheit DHV - Datenerhebung, Häufigkeit, Verteilung gesehen, dass sich Kenngrößen zumindest näherungsweise berechnen lassen, wenn statt der Originaldaten die Häufigkeitsverteilung klassierter Werte vorliegt. In den folgenden Abschnitten werden Sie das Vorgehen, ein Rechenschema und die Anwendung am Beispiel kennen lernen.

Die Rechnungen werden in Tabellenform dargestellt. Es werden die Aggregationseigenschaft und lineare Transformationen verwendet.

Bei klassierten Daten tritt eine Datenreduktion auf, da wir nur noch das Intervall kennen, in dem ein Originalwert liegt, der eigentliche Wert ist nicht bekannt. Die Datenreduktion ist mit Informationsverlust verbunden. Dennoch gelingt es, den Mittelwert mit geringer Verzerrung aus klassierten Werten zu bestimmen. Zur Darstellung erweitern wir die Häufigkeitstabelle um eine weitere Rechenspalte und Summenzeile.

Index k	**Klasse** K_k	**Klassenmitte** m_k	**absolute oder relative Häufigkeit** f_k	**Produkt** $f_k \cdot m_k$
1 . k . a	K_1 . K_k . K_a	m_1 . m_k . m_a	f_1 . f_k . f_a	$f_1 \cdot m_1$. $f_k \cdot m_k$. $f_a \cdot m_a$
		Summe	$\sum f_k$	$\sum f_k \cdot m_k$

Tab.: Erweiterte Häufigkeitstabelle

In der obigen Tabelle können für Häufigkeiten absolute Werte $(f_k = n_k)$ oder relative Werte $f_k = h_k$ eingesetzt werden. Die Klassen werden wie üblich durch Intervallgrenzen $(K_k = (x_{k-1}, x_k])$ festgelegt.

Daraus berechnet man die Klassenmitten $m_k = \frac{x_{k-1}+x_k}{2}$

Wenn wir annehmen, dass die Klassenmitten tatsächlich mit den Mittelwerten der Daten in den einzelnen Klassen übereinstimmen, dann ließe sich der Originalmittelwert mit Hilfe der Ersatzwerteigenschaft berechnen. So verwenden wir die Formel als Näherung für den tatsächlichen Mittelwert.

$$\overline{x} = \frac{f_1 m_1 + f_2 m_2 + \ldots + f_a m_a}{f_1 + f_2 + \ldots + f_a} = \frac{\sum\limits_{k=1}^{a} f_k m_k}{\sum\limits_{k=1}^{a} f_k}$$

Falls absolute Häufigkeiten verwendet werden $f_k = n_k$ gilt:

$$\overline{x} = \frac{1}{n} \sum_{k=1}^{a} f_k m_k$$

Die Formel sieht im Fall von relativen Häufigkeiten wie folgt aus:

$$\overline{x} = \sum_{k=1}^{a} f_k m_k$$

Damit Sie sehen, wie der Rechengang läuft, zunächst ein kleines Zahlenbeispiel:

Häufigkeitstabelle: Beispieldaten

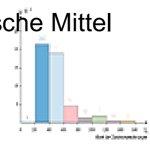
Index k	Klasse K_k	Klassenmitte m_k	Häufigkeit f_k	Produkt $f_k \cdot m_k$
1	(0, 2]	1	30	30
2	(2, 4]	3	50	150
3	(4, 6]	5	20	100
		Summe	100	280

Tab.: Häufigkeitstabelle: Beispieldaten

Beispiel

Der Mittelwert auf der Basis der klassierten Daten ergibt sich als $\frac{280}{100} = 2{,}8$.
Das Dividieren durch 100 hätten wir auch schon bei den Häufigkeiten vornehmen können, dann hätten wir für f_k die relativen Häufigkeiten eingesetzt - mit demselben Ergebnis.

4.1 Das arithmetische Mittel klassierter Daten am Beispiel Mietpreise



Sie kennen schon die Mietpreise von Einzimmerwohnungen in Berlin. Das arithmetische Mittel der Mietpreise haben wir im Kapitel 2.1 in dieser Lerneinheit berechnet. Es beträgt 495 €.

Ein Histogramm der klassierten Daten wurde Ihnen bereits in der Lerneinheit DHV in Kapitel 7.2 vorgestellt: Jetzt berechnen wir das Mittel auf der Basis der klassierten Werte mit Hilfe der folgenden Tabelle.

Index k	Klasse K_k	Klassenmitte m_k	absolute Klassenhäufigkeit n_k	relative Klassenhäufigkeit h_k	Produkt $f_k \cdot m_k$
1	K_1 = (0; 200]	100	0	0,0000	0
2	K_2 = (200; 400]	300	43	0,4343	12900
3	K_3 = (400; 600]	500	38	0,3838	19000
4	K_4 = (600; 800]	700	9	0,0909	6300
5	K_5 = (800; 1000]	900	3	0,0303	2700
6	K_6 = (1000; 1200]	1100	4	0,0404	4400
7	K_7 = (1200; 1400]	1300	1	0,0101	1300
8	K_8 = (1400; 1600]	1500	1	0,0101	1500
	Summe		99	1	48100

Tab.: Häufigkeitstabelle: Mietpreise

Die näherungsweise Berechnung auf der Basis der klassierten Daten ergibt 48100 / 99 = 485,86 € als mittlere Miete. Immerhin ein Unterschied von rund 10 € zum tatsächlichen Wert. Der relative Fehler liegt bei etwa 2 %.

4.2 Arithmetisches Mittel bei klassierten Daten (Übungsaufgaben)

Berechnen

Übung ARM-09

Grundstücke

Die Inhaberin des Weinfachgeschäftes Maestro hat 10 neue Angebote von Grundstücken in Veneto bekommen. Die Grundstücke haben folgende Preise umgerechnet in Tausend €:

Angebot	1	2	3	4	5	6	7	8	9	10
Preise (in Tausend €)	10,5	5	4,3	12	14	10,5	6,2	5	7,3	10,5

Bestimmen und interpretieren Sie den Mittelwert der Preise.

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-10

Verkaufte Flaschen

Die Inhaberin des Weinfachgeschäftes Maestro hat einen neuen Prosecco in ihr Sortiment aufgenommen und interessiert sich für die Anzahl der pro Tag verkauften Flaschen dieser Sorte. Vier Wochen lang hat sie täglich die Anzahl der verkauften Flaschen notiert und in der folgenden Tabelle zusammengefasst:

Flaschenanzahl	Tage
0	3
2	5
3	7
5	4
7	3
8	2

a) Bestimmen und interpretieren Sie das arithmetische Mittel.
b) Wie viele Flaschen des neuen Prosecco konnten im Laufe dieser vier Wochen insgesamt abgesetzt werden?

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-11

Wohnungsanzahl

Bezüglich der Anzahl der Wohnräume für ein bestimmtes Jahr in der Stadt N ergab die Statistik des Wohnungsbestands folgendes Bild:

Raumanzahl	Wohnungsanzahl
1	130
2	615
3	1855
4	2720
5	1147
6	383
7	120

Bestimmen und interpretieren Sie Folgendes: Wie viele Zimmer besitzt eine Wohnung im Durchschnitt?

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-12

Arbeitsunfälle

Im April des Jahres wurde die Anzahl der Arbeitsunfälle in Firma Maestro, durch die jeweils eine beschäftigte Person verletzt wurde, statistisch erfasst und aufbereitet. Es ergab sich, dass 90 % der Beschäftigten keinen Arbeitsunfall hatten, 7 % hatten einen Arbeitsunfall und 3 % hatten genau zwei Arbeitsunfälle.

Bestimmen und interpretieren Sie für diese Daten das arithmetische Mittel.

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Zusammenfassung

- Mittelwerte sind die bekanntesten statistischen Größen, sie fassen die Daten mit einer einzigen Zahl zusammen.
- Das arithmetische Mittel oder der Mittelwert $\overline{x}$ (sprich: „x quer“) charakterisiert das Niveau einer Datenreihe.
- Vorsicht! Das arithmetische Mittel reagiert empfindlich auf Ausreißer.
- Gewogene Mittelwerte werden mit Hilfe von Gewichten bestimmt.
- Für klassierte Daten kann das arithmetische Mittel näherungsweise als gewichtetes Mittel der Klassenmitten bestimmt werden.
- Mittelwerte von Teilreihen lassen sich gewichtet zum Gesamtmittel zusammenfassen.
- Das arithmetische Mittel liegt im Schwerpunkt der Daten.
- Das arithmetische Mittel minimiert die Summe der Abweichungsquadrate.
- Mittelwertberechnung und lineare Transformationen sind vertauschbar..

Sie sind am Ende dieser Lerneinheit angelangt. Auf den folgenden Seiten finden Sie noch die Übungen zur Wissensüberprüfung und die Übungen für die Statistiksoftware R.

Wissensüberprüfung

Mit den folgenden Übungen können Sie ihr Wissen überprüfen. Versuchen Sie zuerst, die Aufgaben selbst zu lösen, bevor Sie sich die Antworten anzeigen lassen.

Berechnen

Übung ARM-13

Wiederholungsprüfungen

Die folgende Liste beinhaltet die Anzahl der Wiederholungsprüfungen im Fach Statistik von 117 Studierenden einer Berliner Fachhochschule, die im Verlauf des WS 2021 ihre Prüfung absolvierten.

0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 2 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0
0 1 0 0 1 0 0 3 0 0 1 0 0 0 0 2 0 1 0 1 0 1 0 0 0 1 0 0 1 1 0 1 0 0 1 0 0 0 0 0 1 0 0 0 1 0
2 1 1 0 0 0 0 0 1 0 1 1 1 1 0 2 0 1 1 2 1 1 0 1 0

Charakterisieren Sie die Verteilung des Erhebungsmerkmals mit Hilfe des arithmetischen Mittels. Interpretieren Sie das Ergebnis.

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-14

Störfälle

In einer Gärtnerei in Treptow wurden in den Monaten Januar 2020 bis Dezember 2021 jeweils folgende Anzahl von Störfällen registriert:

Monat	Jan	Feb	März	April	Mai	Juni	Juli	Aug	Sep	Okt	Nov	Dez
2020	0	4	1	4	2	8	2	4	3	2	1	6
2021	7	1	6	1	9	7	8	8	4	3	7	0

Überprüfen Sie die folgende Aussage:

„Die mittlere Störzahl pro Monat lag 2021 höher als 2020“.

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-15

Körpergewichte

Die folgende Tabelle zeigt die Häufigkeitsverteilung des Körpergewichts von 100 Studierenden einer Sportschule.

Index k	Klasse K_k	Absolute Klassenhäufigkeit $n_k = f_k$
1	(57; 60]	2
2	(60; 63]	5
3	(63; 66]	16
4	(66; 69]	42
5	(69; 72]	25
6	(72; 75]	7
7	(75; 78]	3

Bestimmen Sie das arithmetische Mittel $\overline{x}$ der klassierten Daten.

Lösung (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

Multiple Choice

Übung ARM-16

Welche Aussagen sind falsch und welche richtig?

	Richtig	Falsch	Auswertung
Das arithmetische Mittel $\bar{x}$ einer Gesamtreihe ist gleich dem gewichteten Mittel aus den Mittelwerten $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$ der Teilreihen.			
Die Summe der Abweichungen der Einzelwerte des arithmetischen Mittels ist stets gleich Null. Diese Aussage beschreibt die Minimumeigenschaft des arithmetischen Mittels $\bar{x}$.			
"Ungewichtetes" und "gewichtetes" arithmetisches Mittel nur Bezeichnungen für zwei Arten der Berechnung des gleichen arithmetischen Mittels.			
Die Bildung des arithmetischen Mittels und lineare Transformationen sind vertauschbar.			
Die Summe der quadrierten Abweichungen der Einzelwerte des arithmetischen Mittels ist minimal. Diese Aussage beschreibt die Schwerpunkteigenschaft des arithmetischen Mittels $\bar{x}$.			

Übungen mit der Statistiksoftware R

Die in der Lerneinheit behandelten Themen können Sie anhand der folgenden Übungsaufgaben auch mit der Statistiksoftware **R** oder **R-Studio** bearbeiten. Um die Übungen zu bearbeiten, sollte die Software „**R**" auf Ihrem Rechner installiert sein.

Installationshinweise [Manuals | **R** Installation and Administration]

Versuchen Sie, die auf dieser Seite gestellten Aufgaben mit Hilfe von **R** oder **R-Studio** zu lösen. Ihre Ergebnisse können Sie mit der Lösungsdatei vergleichen. Speichern Sie dazu die R-Datei mit der Lösung auf Ihrem Rechner und öffnen Sie die Datei mit **R**. Überprüfen Sie vor dem Öffnen, ob das Arbeitsverzeichnis von **R** mit dem Speicherort der Datei übereinstimmt. Hilfe beim Einlesen von Daten aus Text- oder CSV-Dateien finden Sie u. a. auf den Hilfeseiten der Firma Posit.

Für einige Übungen stehen auch Musterlösungen für das Programm Excel bereit.

Berechnen

Übung ARM-17a

Streiktage

die Tabelle zeigt die jährlich durch Streiks ausgefallenen Arbeitstage nach Ländern von 2011 bis 2020.

Berechnen Sie

1. den Median,
2. das arithmetische Mittel und
3. interpretieren Sie die Ergebnisse.

Belgien	97
Frankreich	93
Kanada	79
Finnland	52
Spanien	48
Dänemark	44
Norwegen	38
Niederlande	21
Vereinigtes Königreich	18
Deutschland	18
Irland	16
Polen	16
Portugal	13
USA	9
Ungarn	4
Österreich	2
Schweden	2
Schweiz	1

Quelle: Hans-Böckler-Stiftung; WSI (Statista) 2022

Lösung mit R und Excel (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

Berechnen

Übung ARM-17b

Torjäger des FC-Bayern

In der Datei **`bayerntore.txt`** finden Sie einen Datensatz mit der ewigen Torschützenliste des FC Bayern (Stand: Saison 2022/23). Speichern Sie die Datei im gleichen Ordner wie die R-Datei. Mit dem **`read.table()`** Befehl laden Sie die Daten in einen Data Frame (hier: tore), damit können Sie die unten stehenden Aufgaben bearbeiten.

`bayerntore.txt` (2 KB)

Aufgaben:

1. Erstellen Sie ein Histogramm mit Klassenbreite 10! Zeichnen Sie darin folgende Linien ein: Median, arithmetisches Mittel, jeweils unklassiert!
2. Worin liegt es begründet, dass beide (Median, arith. Mittel) stark voneinander abweichen?
3. Schließen Sie auf eine symmetrische, linkssteile oder rechtssteile Verteilung? Begründen Sie ihre Entscheidung!

In der Übung wird schon der Median abgefragt, der später in einer eigenen Lerneinheit genauer betrachtet wird.

Lösung mit R und Excel (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

Berechnen

Übung ARM-17c

Europa

In der Datei **`europa.txt`** finden Sie die Fläche (in 1000 Quadratkilometern), die Bevölkerung (in 1000) und die Arbeitslosenquote in Prozent (Stand 10/22) von einigen europäischen Ländern.

`europa.txt` (2 KB)

Speichern Sie die Datei im gleichen Ordner wie die R-Datei. Mit dem **`read.table()`** Befehl laden Sie die Daten in einen Data Frame (hier: europa), damit können Sie die unten stehenden Aufgabe bearbeiten

Aufgabe:

1. Berechnen Sie für die drei Variablen das arithmetische Mittel und den Median.

In der Übung wird schon der Median abgefragt, der später in einer eigenen Lerneinheit genauer betrachtet wird.

Lösung mit R (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

Berechnen

Übung ARM-17d

Gewicht

Aus der Ihnen schon bekannten Studierendengruppe sehen wir uns nun die 10 kleinsten Studierenden etwas genauer an.

	Name	Größe	Gewicht
1	Gleb	158	48
2	Marc	160	59
3	Malte	163	102
4	Alina	165	57
5	Taner	165	80
6	Anne	168	53
7	Louis	168	58
8	Esra	168	58
9	Sofia	169	66
10	Oliver	170	87

Aufgabe:

1. Berechnen Sie den Median und das arithmetische Mittel des Gewichts sowie die Standardabweichung. Entfernen Sie nun den übergewichtigen Roland aus dem Datensatz und führen Sie die Berechnungen des Medians und des arithmetischen Mittels noch einmal durch. Welche Beobachtungen können Sie machen? Interpretieren sie Ihre Ergebnisse.

In der Übung wird die Standardabweichung betrachtet - diese wird in später noch genauer erläutert.

Lösung mit R und Excel (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

Berechnen

Übung ARM-17e

Notenspiegel

Die folgende Liste enthält die Schulnoten einer 10. Klasse in Mathematik.

1 5 2 4 2 3 3 3 2 2 2 4 3 4 3 5 3 1 3 6 6 3 4 5

Aufgaben:

1. Geben Sie die absoluten Häufigkeiten an.
2. Berechnen Sie das arithmetische Mittel, den Median und den Modus.

In der Übung wird schon der Median abgefragt, der später in einer eigenen Lerneinheit genauer betrachtet wird.

Lösung mit R und Excel (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

Berechnen

Übung ARM-17f

Grundstücke

Die Inhaberin des Weinfachgeschäftes Maestro hat 10 neue Angebote von Grundstücken in Veneto bekommen. Die Grundstücke haben folgende Preise umgerechnet in Tausend Euro:

10,5	5	4,3	12	14	10,5	6,2	5	7,3	10,5

Aufgabe:

1. Bestimmen und interpretieren Sie den Mittelwert der Preise von Angeboten.

Lösung mit R (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

Berechnen

Übung ARM-17g

Prosecco

Die Inhaberin des Weinfachgeschäftes Maestro hat eine neue Sorte Prosecco in ihr Sorti- ment aufgenommen und interessiert sich für die Anzahl der pro Tag verkauften Flaschen dieser Sorte. Vier Wochen lang hat sie täglich die Anzahl der verkauften Flaschen notiert und in der folgenden Tabelle zusammengefasst.

Anzahl Flaschen	Tage
0	3
2	5
3	7
5	4
7	3
8	2

Aufgabe:

1. Bestimmen und interpretieren Sie das arithmetische Mittel.

Lösung mit R (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

Berechnen

Übung ARM-17h

Wohnungsbestand

Bezüglich der Anzahl der Wohnräume für ein bestimmtes Jahr in der Stadt N ergab die Statistik des Wohnungsbestands folgendes Bild:

Raumanzahl	Wohnungsanzahl
1	130
2	615
3	1855
4	2720
5	1147
6	383
7	120

Aufgabe:

1. Bestimmen Sie, wie viele Zimmer eine Wohnung im Durchschnitt besitzt.

Lösung mit R (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-17i

Störfälle

In einem Zulieferbetrieb wurden in den Monaten Januar 2021 bis Dezember 2022 folgende Anzahl von Störfällen registriert.

Monat	Jahr 2021	Jahr 2022
Jan	0	7
Feb	4	1
Mae	1	6
Apr	4	1
Mai	2	9
Jun	8	7
Jul	2	8
Aug	4	8
Sep	3	4
Okt	2	3
Nov	1	7
Dez	6	0

Aufgabe:

1. Überprüfen Sie die folgende Aussage: „*Die mittlere Störzahl lag 2022 höher als 2021*“.

Lösung mit R und Excel (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

Berechnen

Übung ARM-17k

Arbeitsunfälle

Im April diesen Jahres wurde die Anzahl der Arbeitsunfälle in der Firma Maestro, durch die jeweils eine Person in Mitleidenschaft gezogen wurde, statistisch erfasst und aufbereitet. Es ergab sich, dass 90 % der Beschäftigten keinen Arbeitsunfall hatten, 7 % hatten einen Arbeitsunfall und 3 % hatten genau zwei Arbeitsunfälle.

Aufgabe:

1. Bestimmen und interpretieren Sie für diese Daten das arithmetische Mittel.

Lösung mit R (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

Berechnen

Übung ARM-17m

Wiederholungsprüfung

Die folgende Liste beinhaltet die Anzahl der Wiederholungsprüfungen im Fach Statistik von 117 Studierenden einer Berliner Hochschule, die im Verlauf des WS 2022 ihre Prüfung absolvierten.

0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 2 0 0 0 0 0 0 0 1
0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 3 0 0 1 0 0 0 0 2 0 1 0 1 0 1 0 0 0 1 0 0 1 1
0 1 0 0 1 0 0 0 0 0 1 0 0 0 1 0 2 1 1 0 0 0 0 0 1 0 1 1 1 1 0 2 0 1 1 2 1 1
0 1 0

Aufgabe:

1. Bestimmen Sie die Verteilung des Erhebungsmerkmals und berechnen Sie das arithmetische Mittel. Interpretieren Sie ihre Ergebnisse

Lösung mit R (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

Zusätzliche Übungsaufgaben

Zum Vertiefen der vorgestellten Inhalte finden Sie hier zusätzliche Übungsaufgaben zur freiwilligen Bearbeitung.

Berechnen

Übung ARM-18

Körpergröße

Berechnen Sie das arithmetische Mittel der folgenden Körpergrößen:

168, 168, 169, 170, 170, 171, 172, 172, 172, 172, 173, 174, 175, 175, 175, 178, 179, 180, 181

(Angaben in Zentimeter)

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-19

Alter von Studierenden

In der folgenden Liste ist das Alter von 25 Studierenden aufgeführt:

22, 20, 21, 23, 23, 24, 25, 28, 25, 23, 22, 21, 23, 24, 25, 21, 24, 23, 26, 24, 23, 25, 23, 21, 25

(Angaben in Jahren)

Bestimmen Sie das arithmetische Mittel des Alters.

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-20

Verträge

Ein Versicherungsmakler hat in 20 Tagen folgende Anzahl von Verträgen pro Tag abgeschlossen:

12, 15, 20, 23, 30, 13, 18, 17, 19, 25, 26, 14, 26, 19, 30, 32, 26, 28, 25, 20

Berechnen Sie das arithmetische Mittel der Anzahl abgeschlossener Verträge.

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-21

Übertragungsrate

Um die Geschwindigkeit der Internetverbindung zu überprüfen, wird eine Datei mit einer Größe von 4835 kb 6-mal zu unterschiedlichen Tageszeiten verschickt und dabei die Zeit gemessen, bis diese beim Empfänger ankommt. Sie erhalten folgende Werte:

35,8	23,4	15,6	19,8	28,3	31,2

(Angaben in Sekunden)

Berechnen Sie die durchschnittliche Übertragungsrate in kb/s.

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-22

Karies bei Schulkindern

Bei einer medizinischen Untersuchung von 100 Schulkindern wurden folgende Häufigkeiten der kariösen Zähne beobachtet:

Anzahl kariöser Zähne	f_i
0	35
1	27
2	18
3	13
4	7

Bestimmen Sie den Mittelwert kariöser Zähne der 100 Schulkinder.

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-23

Kraftstoffverbrauch

Bestimmen Sie anhand der folgenden Tabelle den Durchschnittsverbrauch eines PKW:

Tankvorgang	Strecke	Durchschnittsverbrauch
1	657 km	9,8 l
2	734 km	9,3 l
3	702 km	9,5 l
4	1020 km	7,4 l
5	812 km	8,9 l

Berechnen Sie das arithmetische Mittel.

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-24

Körpergrößen von Männern und Frauen

Bei 80 Personen wurde die Körpergröße gemessen, davon sind 60 % männlich mit einer mittleren Körpergröße von 182,3 cm und 40 % weiblich mit 165,6 cm mittlerer Körpergröße.

Bestimmen Sie die durchschnittliche Körpergröße dieser 80 Personen.

Lösung (Siehe Anhang)

Bearbeitungszeit: 5 Minuten

Berechnen

Übung ARM-25

Umfrage in einer Kleinstadt

Bei einer Umfrage von 160 Personen in einer Kleinstadt ergab sich folgende Tabelle des monatlichen Bruttoeinkommens in Euro.

Lohnklasse	Einkommen (Euro)	Klassenhäufigkeit
A	[1000; 1500)	35
B	[1500; 2000)	68
C	[2000; 2500)	27
D	[2500; 3500)	17
E	[3500; 5000)	13

Bestimmen Sie das arithmetische Mittel.

Lösung (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

Berechnen

Übung ARM-26

Wohnungsanzeigen

Eugen und Mercan suchen für sich eine Mietwohnung und vergleichen dazu einige Wohnungsanzeigen aus dem Internet. Leider kann sich die junge Familie nicht einigen, ob sie eine Altbau- oder eine Neubauwohnung haben will. In der folgenden Tabelle sind die Mieten für 100 Altbauwohnungen aufgelistet:

i	Miete in Euro	Häufigkeit
1	[0; 300)	2
2	[300; 400)	5
3	[400; 500)	36
4	[500; 600)	47
5	[600; 700)	6
6	[700; 800)	4

Eugen ist der Meinung, dass die Altbauwohnungen grundsätzlich teurer sind. Das arithmetische Mittel für 100 Neubauwohnungen ist gleich 543. Bestimmen Sie nach einer beliebigen Methode das arithmetische Mittel für die Altbauwohnungen und beurteilen Sie Ihre Ergebnisse.

Lösung (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

Berechnen

Übung ARM-27

Rentner in Brandenburg

In der folgenden Tabelle ist der Anteil an der Bevölkerung über 65 Jahren in den Landkreisen und kreisfreien Städte in Brandenburg dargestellt.

i	Anteile (in%)	Häufigkeit
1	[2,5; 3,0)	5
2	[3,0; 3,5)	11
3	[3,5; 4,0)	9
4	[4,0; 4,5)	5

Bestimmen Sie das arithmetische Mittel.

Lösung (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

Appendix

Lösung für Übung ARM-01

Zahlen

$$\overline{x} = \frac{2+4+6+7+3+2+4+5+5+7+6+5+7+7+8+5+4+7+3+7+8+7+4+4+7}{30} = \frac{158}{30} = 5{,}27$$

Lösung für Übung ARM-02

Punktzahl

$$\overline{x} = \frac{87+78+97+67+91+92+79}{7} = \frac{591}{7} = 84{,}4$$

Lösung für Übung ARM-03

Gehälter von acht Angestellten

$$\overline{x} = \frac{1500+2400+1800+2000+2100+5500+2200+1900}{8} = \frac{19400}{8} = 2425$$

Das arithmetische Mittel der Gehälter liegt bei 2.425 €

Lösung für Übung ARM-04

Gehälter von drei Abteilungen

Abteilung L:

$$\overline{x}_L = \frac{1600+1850+1550+2200+2100}{5} = \frac{9300}{5} = 1860$$

Das mittlere Monatsgehalt in der Abteilung L liegt bei 1.860 €

Abteilung M:

$$\overline{x}_M = \frac{2500+1950+2250+1700+2300+1500+2100}{7} = \frac{14300}{7} = 2042{,}9$$

Das mittlere Monatsgehalt in der Abteilung M liegt bei 2.042,90 €

Abteilung N:

$$\overline{x}_N = \frac{2400+2350+1875}{3} = \frac{6625}{3} \approx 2208{,}33$$

Das mittlere Monatsgehalt in der Abteilung N liegt bei rund 2.208,33 €

Gesamter Betrieb:

$$\overline{x}_{Abt.} = \frac{1600+1850+1550+2200+2100+2500+1950+2250+1700+2300+1500+2100+2400+2350+1875}{15}$$

$$= \frac{30225}{15} = 2015$$

Vereinfacht kann auch so gerechnet werden:

$$\overline{x}_{Abt.} = \frac{9300 + 14300 + 6625}{5 + 7 + 3} = \frac{1}{15} \cdot 30225 = 2015$$

Das mittlere Monatsgehalt aller Angestellten liegt bei 2.015 €.

Lösung für Übung ARM-05

1000 Zahlen

$$\overline{x} = \frac{(150 \cdot 3) + (50 \cdot 2) + (400 \cdot 7) + (300 \cdot 5) + (100 \cdot 6)}{1000} = \frac{5450}{1000} = 5{,}45$$

Lösung für Übung ARM-06

Pizzabestellung

1) Wir überprüfen unser Ergebnis mit der Formel für zusammengesetzte Datensätze:

$$\bar{x} = \frac{n_1 \cdot \bar{x}_1 + n_2 \cdot \bar{x}_2}{n_1 + n_2} = \frac{10 \cdot 16{,}5 + 1 \cdot 22}{11} = \frac{187}{11} = 17$$

Im Durchschnitt tragen die 11 Studenten und Studentinnen 17 Euro zur Bestellung bei.

2) Wir überprüfen unsere Annahme wieder mit der Formel:

$$\bar{x} = \frac{n_1 \cdot \bar{x}_1 + n_2 \cdot \bar{x}_2}{n_1 + n_2}$$

$$16{,}5 + 1 = \frac{10 \cdot 16{,}5 + 1 \cdot x}{11}$$

$$17{,}5 = \frac{165 + x}{11} = \frac{165}{11} + \frac{x}{11}$$

$$17{,}5 = 15 + \frac{x}{11}$$

$$2{,}5 = \frac{x}{11}$$

$$x = 27{,}5$$

Sie hätte 27,50 € beitragen müssen.

Lösung für Übung ARM-07

Mittleres Jahreseinkommen

Der Mittelwert wäre nur dann 28.000 €, wenn die Anzahl der Arbeiter und Arbeiterinnen von Treptow und Schöneberg gleich wäre. Um das wahre mittlere Jahreseinkommen zu bestimmen, müssten wir die Anzahl der Arbeiter und Arbeiterinnen in jedem Bezirk kennen. Gibt es z. B. auf 7 Arbeiter/innen aus Schöneberg je eine/n aus Treptow, so wäre der gemeinsame, gepoolte Mittelwert:

$$\overline{x} = \frac{(1 \cdot 26000\,€) + (7 \cdot 30000\,€)}{1+7} = \frac{236000\,€}{8} = 29500\,€$$

Das gemeinsame Durchschnittsjahreseinkommen beider Bezirke wäre bei dem hier angenommenen Verhältnis 29.500 €.

Lösung für Übung ARM-08

Mittleres Monatsgehalt

$$\overline{x}_L = \frac{1600+1850+1550+2200+2100}{5} = 1860$$

$$\overline{x}_M = \frac{2500+1950+2250+1700+2300+1500+2100}{7} \approx 2042{,}86$$

$$\overline{x}_N = \frac{2400+2350+1875}{3} \approx 2208{,}33$$

$$\overline{x}_{Abt.} = \frac{(5 \cdot \overline{x}_L) + (7 \cdot \overline{x}_M) + (3 \cdot \overline{x}_N)}{5+7+3} = \frac{5}{15}\overline{x}_L = \frac{7}{15}\overline{x}_M = \frac{3}{15}\overline{x}_N$$

Das durchschnittliche Firmeneinkommen lässt sich wie folgt berechnen:

$$\overline{x}_{Firma} = \frac{5 \cdot 1860 + 7 \cdot 2042{,}86 + 3 \cdot 2208{,}33}{15} \approx 2015$$

Das monatliche Durchschnittseinkommen in der gesamten Firma beträgt rund 2015 €. Wir haben uns hier die Ersatzwerteigenschaft des arithmetischen Mittels zunutze gemacht.

Lösung für Übung ARM-09

Grundstücke

$$\overline{x} = \tfrac{1}{n}(x_1 + x_2 + \ldots + x_n) = \tfrac{1}{n}\sum_{i=1}^{n} x_i$$

$$\overline{x} = \tfrac{1}{10}(10{,}5+5+4{,}3+12+14+10{,}5+6{,}2+5+7{,}3+10{,}5) = \tfrac{85{,}3}{10} = 8{,}53$$

Interpretation: $\overline{x}$ = 8,53, d. h. der Durchschnittspreis der Grundstücke beträgt 8.530 €.

Lösung für Übung ARM-10

Verkaufte Flaschen

$$\overline{x} = \frac{g_1 x_1 + g_2 x_2 + \ldots + g_n x_n}{g_1 + g_2 + \ldots + g_n} = \frac{\sum_{i=1}^{n} g_i x_i}{\sum_{i=1}^{n} g_i}$$

$$\overline{x} = \frac{(3 \cdot 0) + (5 \cdot 2) + (7 \cdot 3) + (4 \cdot 5) + (3 \cdot 7) + (2 \cdot 8)}{3 + 5 + 7 + 4 + 3 + 2} = \frac{88}{24} = 3{,}7$$

Interpretation: Im Durchschnitt wurden 3,7 Flaschen pro Tag verkauft.

88 Flaschen des neuen Prosecco konnten im Laufe dieser vier Wochen insgesamt abgesetzt werden.

Lösung für Übung ARM-11

Wohnungsanzahl

$$\overline{x} = \frac{g_1 x_1 + g_2 x_2 + \ldots + g_n x_n}{g_1 + g_2 + \ldots + g_n} = \frac{\sum_{i=1}^{n} g_i x_i}{\sum_{i=1}^{n} g_i}$$

$$\overline{x} = \frac{(130 \cdot 1) + (615 \cdot 2) + (1855 \cdot 3) + (2720 \cdot 4) + (1147 \cdot 5) + (383 \cdot 6) + (120 \cdot 7)}{130 + 615 + 1855 + 2720 + 1147 + 383 + 120}$$

$$\overline{x} = \frac{130 + 1230 + 5565 + 10880 + 5735 + 2298 + 840}{6970} = \frac{26678}{6970} = 3{,}83$$

Interpretation: $\overline{x}$ = 3,83, d. h. im Durchschnitt besitzt eine Wohnung 3,83 Zimmer.

Lösung für Übung ARM-12

Arbeitsunfälle

$$\overline{x} = g_1 x_1 + g_2 x_2 + \ldots + g_n x_n = \sum_{i=1}^{n} g_i x_i$$

$$\overline{x} = (0{,}90 \cdot 0) + (0{,}07 \cdot 1) + (0{,}03 \cdot 2) = 0{,}07 + 0{,}06 = 0{,}13$$

Interpretation: Es gab im Mittel 0,13 Unfälle je beschäftigter Person.

Lösung für Übung ARM-13

Wiederholungsprüfungen

$$\overline{x} = \frac{g_1 x_1 + g_2 x_2 + \ldots + g_n x_n}{g_1 + g_2 + \ldots + g_n} = \frac{\sum\limits_{i=1}^{n} g_i x_i}{\sum\limits_{i=1}^{n} g_i}$$

$$\overline{x} = \frac{(81 \cdot 0) + (30 \cdot 1) + (5 \cdot 2) + (1 \cdot 3)}{81 + 30 + 5 + 1} = \frac{30 + 10 + 3}{117} = \frac{43}{117} = 0{,}37$$

Interpretation: $\overline{x}$ = 0,37, d. h. im Durchschnitt entfielen auf einen Studierenden 0,37 Wiederholungsprüfungen.

Lösung für Übung ARM-14

Störfälle

Die mittlere Störzahl pro Monat im Jahr 2020und 2021 wurde als ungewichtetes arithmetisches Mittel berechnet:$\overline{x}_{2000}$

$$\overline{x}_{2020} = \frac{1}{n}(x_1 + x_2 + \ldots + x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{12} \cdot 37 = 3{,}1$$

$$\overline{x}_{2021} = \frac{1}{n}(x_1 + x_2 + \ldots + x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{12} \cdot 61 = 5{,}1$$

Die mittlere Störzahl pro Monat im Jahr 2021 betrug 5,1; im Jahr 2020 gab es im Durchschnitt monatlich 3,1 Störfälle. Die Behauptung „Die mittlere Störzahl pro Monat lag 2021 höher als 2020“ ist also richtig: $\overline{x}$ 2021 > $\overline{x}$ 2020.

Lösung für Übung ARM-15

Körpergewichte

a) Rechnungen des arithmetischen Mittels $\overline{x}$ der Körpergewichte von 100 Studierenden für Formel

$$\overline{x} = \frac{f_1 m_1 + f_2 m_2 + \ldots + f_k m_k}{f_1 + f_2 + \ldots + f_k} = \frac{\sum_{k=1}^{a} f_k m_k}{\sum_{k=1}^{a} f_k}$$

Index k	Klasse K_k	Klassenmitte x_k	absolute Klassenhäufigkeit $n_k = f_k$	Produkte $f_k \cdot x_k$
1	(57; 60]	58	2	116
2	(60; 63]	61	5	305
3	(63; 66]	64	16	1024
4	(66; 69]	67	42	2814
5	(69; 72]	70	25	1750
6	(72; 75]	73	7	511
7	(75; 78]	76	3	228
			$N = \sum_{k=1}^{a} f_k = 100$	$N = \sum_{k=1}^{a} f_k x_k = 6748$

Tabelle ARM-15a

$$\overline{x} = \frac{\sum_{k=1}^{a} f_k x_k}{\sum_{k=1}^{a} f_k} = \frac{1}{N} \sum_{k=1}^{a} f_k x_k = \frac{6748}{100} = 67{,}48$$

Der Mittelwert der klassierten Daten liegt bei 67,48 kg

Index k	Klasse K_k	Klassenmitte x_k	Abweichung $d_k = x_k - A$	absolute Klassenhäufigkeit $n_k = f_k$	Produkte $f_k \cdot x_k$
1	(57; 60]	58	-9	2	-18
2	(60; 63]	61	-6	5	-30
3	(63; 66]	64	-3	16	-48
4	(66; 69]	A → 67	0	42	0
5	(69; 72]	70	3	25	75
6	(72; 75]	73	5	7	42
7	(75; 78]	76	9	3	27
Zwischenergebnisse für Formel $\overline{x} = p_1 x_1 + p_2 x_2 + \ldots + p_k x_k = \sum_{k=1}^{a} p_k x_k$				$N = \sum_{k=1}^{a} f_k = 100$	$N = \sum_{k=1}^{a} f_k d_k = 48$

Tabelle ARM-15b

Dann ist nach Formel

$$\overline{x} = A + \frac{\sum\limits_{k=1}^{a} f_k d_k}{\sum\limits_{k=1}^{a} f_k} = A + \frac{1}{N} \sum_{k=1}^{a} f_k d_k$$

$$\overline{x} = A + \frac{\sum\limits_{k=1}^{a} f_k d_k}{\sum\limits_{k=1}^{a} f_k} = A + \frac{1}{N} \sum_{k=1}^{a} f_k d_k = 67 + \frac{48}{100} = 67 + 0{,}48 = 67{,}48$$

b) Rechnungen des arithmetischen Mittels $\overline{x}$ der Körpergewichte von 100 Studierenden für die Methode der Kodierung

$$\overline{x} = A + B\frac{1}{N} \sum_{k=1}^{a} f_k u_k$$

für geschätzten Mittelwert A = 67. Klassenbreite B = 3, im Intervall liegen nur drei Werte! Dadurch besteht die Klassenmitte aus ganzen Werten und kann leicht abgelesen werden.

Index k	**Klasse K_k**	**Klassenmitte x_k**	**Abweichung** $d_k \frac{x_k - A}{B}$	**absolute Klassenhäufigkeit** $n_k = f_k$	**Produkte $f_k \cdot x_k$**
1	(57; 60]	58	-3	2	-6
2	(60; 63]	61	-2	5	-10
3	(63; 66]	64	-1	16	-16
4	(66; 69]	A → 67	0	42	0
5	(69; 72]	70	1	25	25
6	(72; 75]	73	2	7	14
7	(75; 78]	76	3	3	9
Zwischenergebnisse für Formel $\overline{x} = A + \frac{\sum\limits_{k=1}^{a} f_k d_k}{\sum\limits_{k=1}^{a} f_k} = A + \frac{1}{N} \sum\limits_{k=1}^{a} f_k d_k$				$N = \sum\limits_{k=1}^{a} f_k = 100$	$N = \sum\limits_{k=1}^{a} f_k d_k = 16$

Tabelle ARM-14c

Dann ist nach Formel $\overline{x} = A + B\frac{1}{N} \sum\limits_{k=1}^{a} f_k u_k$

$$\overline{x} = A + B\frac{1}{N} \sum_{k=1}^{a} f_k u_k = 67 + (3)\left(\frac{16}{100}\right) = 67 + 0{,}48 = 67{,}48$$

Lösung für Übung ARM-17a

Streiktage

1. Median = 18
2. das arithmetische Mittel beträgt 31,72
3. **Interpretation**: Da es einige Ausreißer nach oben gibt, verschiebt sich das arithmetische Mittel nach oben. Der Median ist stabiler gegenüber Ausreißern.

Lösung mit R

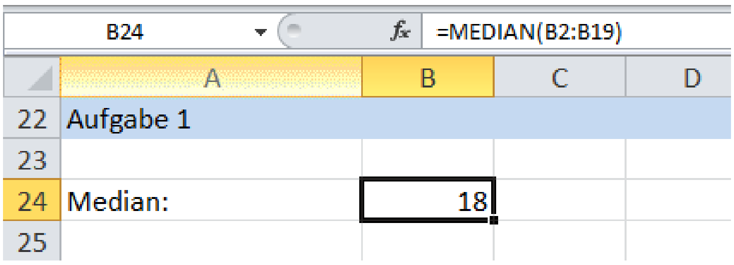
arbeitstage_loesung.R

```
streik <-
  data.frame(
    Land = c(
      "Belgien",
      "Frankreich",
      "Kanada",
      "Finnland",
      "Spanien",
      "Dänemark",
      "Norwegen",
      "Niederlande",
      "UK",
      "Deutschland",
      "Irland",
      "Polen",
      "Portugal",
      "USA",
      "Ungarn",
      "Österreich",
      "Schweden",
      "Schweiz"
    ),
    Tage = c(97, 93, 79, 52, 48, 44, 38, 21, 18, 18, 16, 16, 13, 9, 4, 2, 2, 1)
  )

median(streik$Tage) # Berechnung des Medians Variable Tage

mean(streik$Tage) # Berechnung des arith. Mittels Variable Tage
```

Lösung mit Excel

WMS_ARM_17a_Streiktage.xlsx (9 KB)

Um den Median zu berechnen, benutzen wir die Funktion `MEDIAN`. Im Excel-Tabellenblatt sieht das wie in der nachfolgenden Abbildung aus. Es werden die Zellen mit den Daten markiert.



Genauso gehen wir vor, um das arithmetische Mittel zu berechnen. Die Funktion für das arithmetische Mittel heißt in Excel `MITTELWERT`. Auch hier werden die Zellen mit den Daten markiert.

B29 | f_x =MITTELWERT(B2:B19)

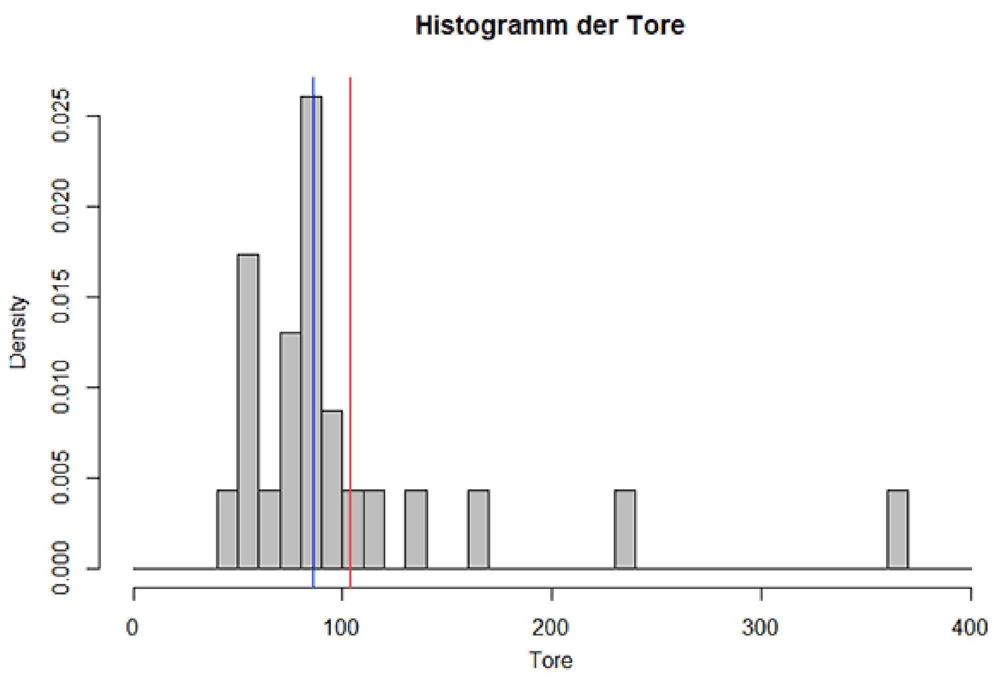
	A	B	C	D
27	Aufgabe 2			
28				
29	Arithmetisches Mittel:	31,7222		
30				

Interpretation: Da es einige Ausreißer nach oben gibt, verschiebt sich das arithmetische Mittel nach oben. Der Median ist stabiler gegenüber Ausreißern.

Lösung für Übung ARM-17b

Torjäger des FC Bayern

1. Histogramm mit Klassenbreite 10, Median (86 blau), arithmetisches Mittel (103,6087 rot).



2. Worin liegt es begründet, dass beide (Median, arith. Mittel) stark voneinander abweichen?

   Gerd Müller, sicherlich ein guter Torschütze (im wahrsten Sinne des Wortes: eine Klasse für sich), ist der Übeltäter. Er stellt in diesem Falle einen Ausreißer dar. Und mit seinen über 120 Toren Vorsprung vor dem zweitbesten Bayerntorschützen bringt er das arithmetische Mittel auf eine immense Höhe. Der Median reagiert dagegen kaum auf den Ausreißer.
3. Schließen Sie auf eine symmetrische, linkssteile oder rechtssteile Verteilung? Begründen Sie ihre Entscheidung!

   Aufgrund der berechneten Lagemaße kann man hier auf eine linkssteile bzw. rechtsschiefe Verteilung schließen. Zum einen ist der Median kleiner als das arithmetische Mittel, zum anderen ist aber auch aus dem Zusammenhang klar, dass viele Fußballer wenige Tore schießen und nur wenige, viele Tore.

Lösung mit R

bayerntore_loesung.R

bayerntore.txt (2 KB)

```
tore<-read.table("bayerntore.txt",sep="\t",header=TRUE)
#Auf die Variablen der Tabelle „tore“ kann jetzt über den Variablennamen direkt zugegriffen werden.
attach(tore)

hist(Tore,breaks=seq(0,400,10),main="Histogramm der Tore",freq=F,col="grey")
abline(v=mean(Tore),col="red") # Hinzufügen einer  vertikalen Gerade im Plot auf Höhe des arithm. Mittels
abline(v=median(Tore),col="blue") # Hinzufügen einer  vertikalen Gerade im Plot auf Höhe des Medians

mean(Tore)
median(Tore)
```

Lösung für Übung ARM-17c

Europa

	Median	Arithmetisches Mittel
Fläche	112	208
Bevölkerung	10.091	22.579,93
Arbeitslosenquote	5,75	6,392857

Lösung mit R

europa_loesung.R

europa.txt (2 KB)

```
europa<-read.table("europa.txt",sep="\t",header=TRUE)

attach(europa)
# Auf die Variablen  der Tabelle „europa“ kann jetzt über den  Variablennamen direkt zugegriffen werden.

#  Berechnung des arithmetischen Mittels
mean(Fl)
mean(Bev)
mean(Arb)

#  Berechnung des Median
median(Fl)
median(Bev)
median(Arb)
```

Lösung für Übung ARM-17d

Gewicht

	Median	Arithmetisches Mittel	Standardabweichung
Alle Personen	58,5	66,8	17,25
ohne Malte	58	62,9	12,75

Interpretation: Wie Sie sehen, hat sich das arithmetische Mittel nach Entfernung des Datzensatzes „Malte“ stark verändert, nämlich um fast vier Kilo. Der Median hingegen ist mit einer Schwankung von gerade mal einem halben Kilo nahezu unverändert geblieben. Er ist unempfindlich gegen Ausreißer und somit ein robuster Lageparameter.

Lösung mit R

gewicht_loesung.R

```
studierendengruppe <-
  data.frame(
    Name = c(
      "Gleb",
      "Marc",
      "Malte",
      "Alina",
      "Taner",
      "Anne",
      "Louis",
      "Esra",
      "Sofia",
      "Oliver"
    ),
    Größe = c(158, 160, 163, 165, 165, 168, 168, 168, 169, 170),
    Gewicht = c(48, 59, 102, 57, 80, 53, 58, 58, 66, 87)
  )

#  Berechnung des arithmetischen Mittels, des Medians und der
#  Standardabweichung
mean(studierendengruppe$Gewicht)
median(studierendengruppe$Gewicht)
sd(studierendengruppe$Gewicht)

#  Entfernen des übergewichtigen Roland
studierendengruppe.neu<-subset(studierendengruppe,Name!="Roland")

#  Nochmals: Berechnung des arithmetischen Mittels, des Medians und
#  der Standardabweichung
mean(studierendengruppe.neu$Gewicht)
median(studierendengruppe.neu$Gewicht)
sd(studierendengruppe.neu$Gewicht)
```

Lösung mit Excel

WMS_ARM_17d_Gewicht.xlsx (9 KB)

Den Median und das arithmetische Mittel bestimmen wir wie in der Aufgabe Streiktage mit den Funktionen `MEDIAN` und `MITTELWERT`. Die Standardabweichung wird mit der Funktion `STABWNA` berechnet. Auch hier werden die Daten, wie nachfolgend zu sehen, markiert.

B20 | f_x | =STABWNA(D2:D11)

	A	B	C	D
17	Mit Malte:			
18	Median:	58,5		
19	Arithmetisches Mittel:	66,8		
20	Standardabweichung:	16,363374		
21				

B26 | f_x | =STABWNA(D2:D3;D5:D11)

	A	B	C	D	E
23	Ohne Malte:				
24	Median:	58			
25	Arithmetisches Mittel:	62,888889			
26	Standardabweichung:	12,022612			
27					

Auf die gleiche Art berechnen wir die gleichen Größen nochmal, nur dieses mal ohne das Gewicht von Malte zu markieren. Ergebnis:

Mit Malte:		
Median:	58,5	
Arithmetisches Mittel:	66,8	
Standardabweichung:	16,36	
Ohne Malte:		
Median:	58	
Arithmetisches Mittel:	62,89	
Standardabweichung:	12,02	

Interpretation: Durch den Ausreißer Malte wird das arithmetische Mittel stark nach oben verschoben, der Median jedoch nicht. Wird Malte heraus genommen, liegen Median und arithmetisches Mittel näher beieinander. Auch die Standardabweichung wird von Ausreißern beeinflusst.

Lösung für Übung ARM-17e

Notenspiegel

1. Absolute Häufigkeiten

1	2	3	4	5	6
2	5	**8**	4	3	2

2. Berechnen Sie das arithmetische Mittel, den Median und den Modus.

Median	Arithmetisches Mittel	Modus
3	3,29	3

Der Modus kann an den absoluten Häufigkeiten abgelesen werden!

Lösung mit R

notenspiegel_loesung.R

```
noten<-c(1,5,2,4,2,3,3,3,2,2,2,4,3,4,3,5,3,1,3,6,6,3,4,5)

#  1)
table(noten)

#  2)
mean(noten)
median(noten)
#  Der Modus kann an den absoluten Häufigkeiten abgelesen werden!
#  Modus = 3
# Es gibt keine Möglichkeit im R Basispaket den Modus zu berechnen.
# Sie können alternativ ein entsprechendes Paket laden:
install.packages("DescTools") #Paket wird installiert
library("DescTools")# Paket wird aufgerufen
Mode (noten)

```

Lösung mit Excel

WMS_ARM_17e_Notenspiegel.xlsx (9 KB)

Wir gehen auch bei dieser Aufgabe wieder sehr ähnlich vor. Zunächst berechnen wir mit der Funktion `HÄUFIGKEIT` die absoluten Häufigkeiten. Dabei stehen die einzelnen Noten in den Feldern B2 bis B25. Nachfolgend ist zu sehen, wie die einzelnen Zellen aussehen müssen.

B32 {=HÄUFIGKEIT(B2:B25;A32:A37)}

	A	B	C	D
28	Aufgabe 1			
29				
30	Absolute Häufigkeiten:			
31	Note	H		
32	1	2		
33	2	5		
34	3	8		
35	4	4		
36	5	3		
37	6	2		
38				

Arithmetisches Mittel, Median und Modus berechnen wir mit den Funktionen `MITTELWERT`, `MEDIAN` und `MODUS.EINF`.

In der folgenden Abbildung ist dies zu sehen, beispielhaft für die Zelle in der der Modus steht.

B44 =MODUS.EINF(B2:B25)

	A	B	C
40	Aufgabe 2		
41			
42	arithmetisches Mittel:	3,29167	
43	Median:	3	
44	Modus:	3	

Wir erhalten folgende Ergebnisse:

Absolute Häufigkeiten:	
Note	H
1	2
2	5
3	8
4	4
5	3
6	2

arithmetisches Mittel:	3,29167
Median:	3
Modus:	3

Lösung für Übung ARM-17f

Grundstücke

1. Das durchschnittliche Angebot beträgt 8530 Euro!

Lösung mit R

grundstuecke_loesung.R

```
angebote<-c(10.5,5,4.3,12,14,10.5,6.2,5,7.3,10.5)

mean(angebote)
```

Lösung für Übung ARM-17g

Prosecco

1. Im Durchschnitt konnten 3,7 Flaschen am Tag verkauft werden.

Lösung mit R

prosecco_loesung.R

```
prosecco <-
  data.frame(Flaschenanzahl = c(0, 2, 3, 5, 7, 8),
             Tage = c(3, 5, 7, 4, 3, 2))

  round(sum(prosecco$Flaschenanzahl * prosecco$Tage) / sum(prosecco$Tage),
        1) # Berechnung des gewichteten arith. Mittels,- gerundet auf eine Nachkommastelle
```

Lösung für Übung ARM-17h

Wohnungsbestand

1. Im Durchschnitt besitzt eine Wohnung 3,8 Zimmer.

Lösung mit R

wbestand_loesung.R

```
wohnungsbestand <-
  data.frame(
    Raumanzahl = 1:7,
    Wohnungsanzahl = c(130, 615, 1855, 2720, 1147, 383, 120)
  )

  round(
    sum(wohnungsbestand$Raumanzahl * wohnungsbestand$Wohnungsanzahl)
 / sum(wohnungsbestand$Wohnungsanzahl),
    1) # Berechnung des  gewichteten arith. Mittels,- gerundet auf eine Nachkommastelle
```

Lösung für Übung ARM-17i

Störfälle

1. Die Aussage, dass die mittlere Störzahl pro Monat im Jahr 2022 höher als im Jahr 2021 lag, ist richtig.

Lösung mit R

stoerfaelle_loesung.R

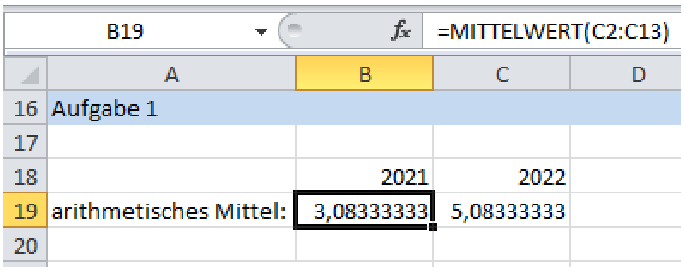
```
stoerfaelle <-
  data.frame(
    Monat = c(
      "Jan", "Feb", "Mae", "Apr", "Mai", "Jun",
      "Jul", "Aug", "Sep", "Okt", "Nov", "Dez"
    ),
    Jahr_2021 = c(0, 4, 1, 4, 2, 8, 2, 4, 3, 2, 1, 6),
    Jahr_2022 = c(7, 1, 6, 1, 9, 7, 8, 8, 4, 3, 7, 0)
  )

aM.2022<-mean(stoerfaelle$Jahr_2022)
aM.2022
aM.2021<-mean(stoerfaelle$Jahr_2021)
aM.2021

aM.2022>aM.2021
```

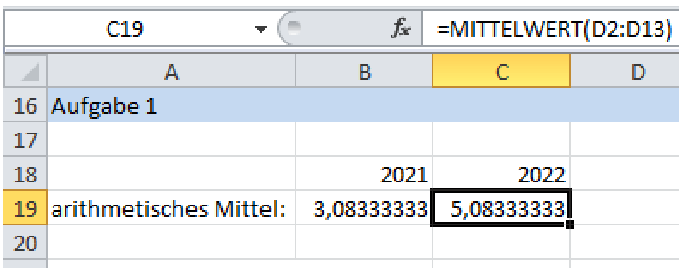
Lösung mit Excel

WMS_ARM_17i_Stoerfaelle.xlsx (9 KB)

Um zu überprüfen, ob die Aussage richtig ist, berechnen wir mit der Funktion `MITTELWERT` jeweils für das Jahr 2021 und das Jahr 2022 die mittlere Störzahl pro Monat.

Wie das geht und welche Werte wir erhalten ist hier zu sehen:





Die Aussage, dass die mittlere Störzahl pro Monat im Jahr 2022 höher als im Jahr 2021 lag, ist also richtig.

Lösung für Übung ARM-17k

Arbeitsunfälle

1. Im Durchschnitt hatte ein Mitarbeiter bzw. eine Mitabeiterin also 0,1 Arbeitsunfälle.

Lösung mit R

arbeitsunfaelle_loesung.R

```
<-round((90*0+7*1+3*2)/100,1)
# Berechnung des gewichteten arith. Mittels,- gerundet auf eine #Nachkommastelle
```

Lösung für Übung ARM-17m

Wiederholungsprüfung

69 % der Studierenden bestehen die Prüfung direkt, 26 % bestehen die erste Wiederholungsprüfung, 4 % bestehen erst die zweite Wiederholungsprüfung und 1 % brauchen einen dritten Versuch. Im Durchschnitt benötigen die Studierenden rund 0,37 Wiederholungen um zu bestehen.

Lösung mit R

wpruefung_loesung.R

```
absolventenliste <-
  c(
  0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,
  0,0,0,0,0,0,0,1,0,0,2,0,0,0,0,0,0,0,1,
  0,0,0,1,0,0,0,0,0,1,0,0,1,0,0,3,0,0,1,
  0,0,0,0,2,0,1,0,1,0,1,0,0,0,1,0,0,1,1,
  0,1,0,0,1,0,0,0,0,0,1,0,0,0,1,0,2,1,1,
  0,0,0,0,0,1,0,1,1,1,1,0,2,0,1,1,2,1,1,
  0,1,0
  )

mean(absolventenliste)
table(absolventenliste) # Absolute Häufigkeiten
round(table(absolventenliste)/117*100) # prozentuale Häufigkeiten, gerundet)
```

Lösung für Übung ARM-18

Körpergröße

$$\overline{x} = \frac{168+168+169+170+170+171+172+172+172+172+173+174+175+175+175+178+179+180+181}{19}$$

$$= \frac{3294}{19} = 173{,}37$$

Lösung für Übung ARM-19

Alter von Studierenden

$$\overline{x} = \frac{22+20+21+23+23+24+25+28+25+23+22+21+23+24+25+21+24+23+26+24+23+25+23+21+25}{25}$$

$$= \frac{584}{25} = 23{,}36$$

Lösung für Übung ARM-20

Verträge

$$\overline{x} = \frac{12+15+20+23+30+13+18+17+19+25+26+14+26+19+30+32+26+28+25+20}{20}$$

$$= \frac{438}{20} = 21{,}9$$

Lösung für Übung ARM-21

Übertragungsrate

$$\frac{(135{,}056 + 206{,}624 + 309{,}936 + 244{,}192 + 170{,}848 + 154{,}968)}{6} = 203{,}604$$

Die durchschnittliche Übertragungsrate beträgt 203,604 kb/s.

Lösung für Übung ARM-22

Karies bei Schulkindern

$$\overline{x} = \frac{(0 \cdot 35) + (1 \cdot 27) + (2 \cdot 18) + (3 \cdot 13) + (4 \cdot 7)}{100} = \frac{130}{100} = 1{,}3$$

Lösung für Übung ARM-23

Kraftstoffverbrauch

$$\overline{x} = \frac{(657 \cdot 9{,}8) + (734 \cdot 9{,}3) + (702 \cdot 9{,}5) + (1020 \cdot 7{,}4) + (812 \cdot 8{,}9)}{3925} = \frac{34708{,}6}{3925} = 8{,}84$$

Lösung für Übung ARM-24

Körpergrößen von Männern und Frauen

Ein Anteil von 60 % bzw. 0,6 der Personen sind Männer mit einer durchschnittlichen Körpergröße von 182,3 cm. 40 % bzw. 0,4 sind anteilig Frauen mit einer Durchschnittsgröße von 165,6. Bei der Berechnung der durchschnittlichen Größe aller Personen müssen wir daher entsprechend gewichten:

$$\overline{x} = \frac{(48 \cdot 182{,}3) + (165{,}6 \cdot 32)}{80} = \frac{14049{,}6}{80} = 175{,}62$$

$$\overline{x} = 0{,}6 \cdot 182{,}3 + 0{,}4 \cdot 165{,}6 = 175{,}62$$

Lösung für Übung ARM-25

Umfrage in einer Kleinstadt

$$\overline{x} = \frac{\sum\limits_{i=1}^{k} f_i m_i}{\sum\limits_{i=1}^{k} f_i}$$

i	Einkommen	f_i	m_i	$f_i \cdot m_i$
1	[1000; 1500)	35	1250	43750
2	[1500; 2000)	68	1750	119000
3	[2000; 2500)	27	2250	60750
4	[2500; 3500)	17	3000	51000
5	[3500; 5000)	13	4250	55250

Damit erhalten wir durch das Einsetzen in die Formel

$$\overline{x} = \frac{43750 + 119000 + 60750 + 51000 + 55250}{160} = 2060{,}94$$

Lösung für Übung ARM-26

Wohnungsanzeigen

i	Miete in Euro	f_i	m_i	$f_i * m_i$
1	[0; 300)	2	150	300
2	[300; 400)	5	350	1750
3	[400; 500)	36	450	16200
4	[500; 600)	47	550	25850
5	[600; 700)	6	650	3900
6	[700; 800)	4	750	3000

Damit erhalten wir als arithmetisches Mittel $\frac{51000}{100} = 510$ und können sagen, dass die Altbauwohnungen im Schnitt sogar günstiger sind als die Neubauwohnungen.

Lösung für Übung ARM-27

Rentner in Brandenburg

i	Anteile in %	f_i	m_i	$f_i \cdot m_i$
1	[2,5; 3,0)	5	2,75	13,75
2	[3,0; 3,5)	11	3,25	35,75
3	[3,5; 4,0)	9	3,75	33,75
4	[4,0; 4,5)	5	4,25	21,25

Damit ist das arithmetische Mittel gleich:

$$\frac{13{,}75 + 35{,}75 + 33{,}75 + 21{,}25}{30} = 3{,}48$$