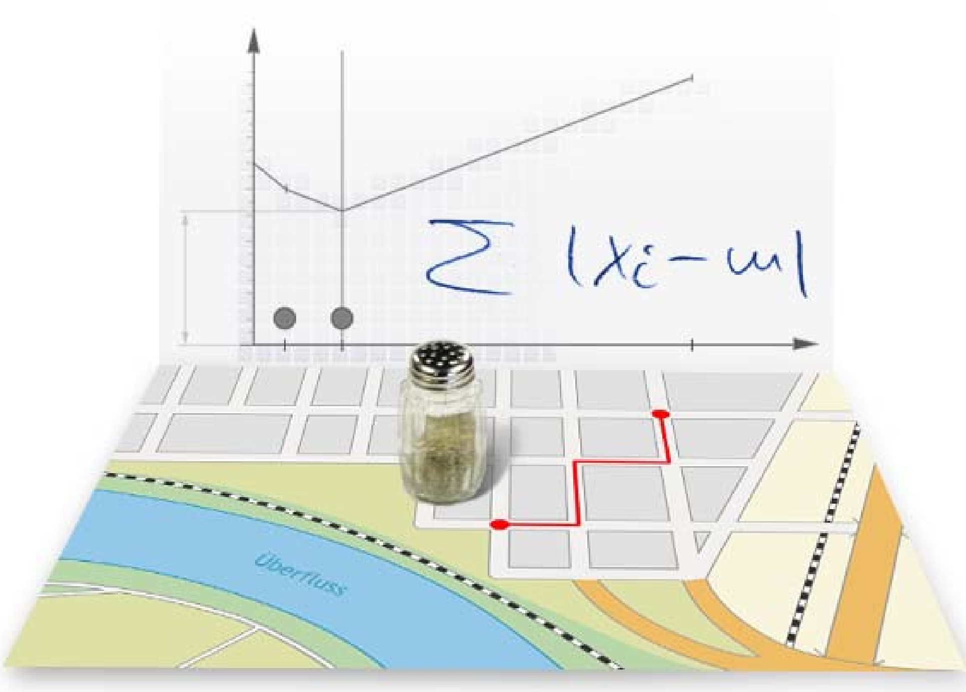
ASM - Alternative Streuungsmaße



Lernziele und Überblick

In dieser Lerneinheit werden Sie weitere Streuungsmaße kennen lernen.

Lernziele

Nach dem Durcharbeiten dieser Lerneinheit sollen Sie das Folgende erklären und berechnen können:

- die Spannweite R,
- den Quartilsabstand IQR,
- die mittlere absolute Abweichung vom arithmetischen Mittel,
- die mittlere absolute Abweichung vom Median.

Gliederung der Lerneinheit

Zeitbedarf und Umfang

Für die Durcharbeitung dieser Lerneinheit benötigen Sie ca. 90 Minuten und für die Übungen mit der Statistiksoftware R ca. 120 Minuten.

1 Einleitung

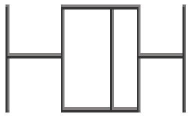
Die Methoden der Statistik sind bunt. Das kennen wir schon vom Studium der Lagemaße. Nicht nur der Mittelwert, auch andere Kenngrößen wie Quantile und Median haben Sinn und Bedeutung. Eigentlich könnten wir neue Kenngrößen für die Streuung unserer Daten selbst erfinden.

Wenn wir wissen, dass zwischen unterem und oberem Quartil ($x_{0,25}$, $x_{0,75}$) die Hälfte der Werte liegt, dann sagt der Abstand zwischen den beiden Quartilen uns auch etwas über Variabilität.

Genauso könnten wir mit Minimum und Maximum verfahren. Vielleicht fallen Ihnen noch mehr Möglichkeiten ein. Aber jetzt haben wir schon fast zu viel verraten. Folgen wir lieber unserer Lerneinheit.

Wir kennen bis jetzt die Varianz und Standardabweichung. In dieser Lerneinheit werden wir Ihnen weitere Streuungsmaße vorstellen:

- die Spannweite R,
- den Quartilsabstand IQR,
- die mittlere absolute Abweichung (durchschnittliche Abweichung) MA.



Hinweis

Wenn Sie die Bedeutung und Berechnung von Quantilen nicht mehr parat haben, sollten Sie sich zunächst noch einmal die Grundlagen aus Lerneinheit *„QBX - Quantile und Boxplot“* Abschnitt 1 und in der Lerneinheit *„MMO - Median und Modus“* den Abschnitt 2 anschauen.

2 Die Spannweite

Das Einfachste aller Streuungsmaße ist die Spannweite
(engl: range - die Weite, die Variationsbreite, die Spanne, die Spannweite).

Beispiel

Spannweite: Einfache Zahlen

Bestimmen Sie die Spannweite der Menge der Zahlen
2, 3, 4, 5, 5, 7, 7, 7, 8, 8, 4, 8, 10, 15, 20.

Die Spannweite beträgt 20 - 2 = 18.

Anstelle der Spannweite werden manchmal nur das Minimum und/oder das Maximum angegeben.

Beispiel

Spannweite: Preise in Geschäften

In 15 Geschäften wird der Preis für einen bestimmten Geschirrspüler ermittelt. Das günstigste Angebot lag bei $x_{(1)}$ = 590 €, das teuerste Angebot bei $x_{(15)}$ = 749 €.

Bestimmen Sie die Spannweite für die Preise eines Geschirrspülers.

Die Zahlen $x_{(1)}$ = 590 € und $x_{(15)}$ = 749 € geben den Streubereich der Häufigkeitsverteilung der Preise für den Geschirrspüler an. Die Spannweite oder die Breite eines Streubereichs für die Preise eines Geschirrspülers ist:

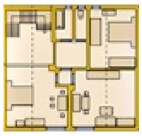
$R = x_{(15)} - x_{(1)} = 749 - 590 = 159$ €.

Beispiel

Spannweite der Mieten der Einzimmerwohnungen

Bestimmen Sie die Spannweite der Mieten der 99 freien Einzimmerwohnungen in Berlin.

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 günstigste Angebot lag bei x_{min} = 231 €, das teuerste Angebot bei x_{max} = 1429 €.

Die Spannweite für die Mieten der 99 freien Einzimmerwohnungen in Berlin ist:

$R = x_{max} - x_{min} = 1429 - 231 = 1198$ €.

2.1 Beispiel: Klausurergebnisse zweier Studierenden

Im folgenden Beispiel werden Sie sehen, inwiefern sich der Range wirklich als Streuungsmaß eignet.

Beispiel

Klausurergebnisse von Anna und Wong


Die Studierenden Anna und Wong haben in diesem Semester 7 Klausuren geschrieben.

Wong erreichte dabei 84, 73, 97, 67, 88, 93 und 79 Punkte, Anna 82, 67, 85, 97, 87, 89 und 88 Punkte.

Bestimmen Sie die Spannweite der Punktzahlen von Anna und Wong.

Man ordnet diese zwei Reihen von Punktzahlen zunächst der Größe nach an:

Wong: 67, 73, 79, 84, 88, 93, 97,
Anna: 67, 82, 85, 87, 88, 89, 97.

Die Spannweite der Punktzahlen bei Wong errechnet sich mit
R_1 = größte Zahl - kleinste Zahl = $x_{max} - x_{min} = x_{(n)} - x_{(1)} = x_{(7)} - x_{(1)} = 97 - 67 = 30$.

Die Spannweite der Punktzahlen bei Anna ist:
R_2 = größte Zahl - kleinste Zahl = $x_{max} - x_{min} = x_{(n)} - x_{(1)} = x_{(7)} - x_{(1)} = 97 - 67 = 30$.

Da die Spannweite keinen Unterschied zwischen den Punktzahlen der beiden angibt, ist sie in diesem Fall kein gutes Maß für die Streuung.

Wir haben den Eindruck, dass die Punktzahlen von Wong stärker streuen als die von Anna.

Was geschieht, wenn die Extremwerte weggelassen werden?

Die Spannweite für die Punktzahlen von Wong ist dann gleich $R'_1 = 93 - 73 = 20$, während die Spannweite für die Punktzahlen von Anna den Wert $R'_2 = 89 - 82 = 7$ ergibt.

Interpretation

Dieses neue Ergebnis verdeutlicht die Tatsache, dass bei den Punktzahlen von Wong eine größere Streuung als bei den Punktzahlen von Anna vorliegt.

Spannweitenvergleich

Man kann die Spannweiten verschiedener Beobachtungsreihen nur dann sinnvoll vergleichen, wenn die Beobachtungsreihen die gleiche Länge n haben.

Die Verlängerung einer Reihe kann nur zur Folge haben, dass sich die Spannweite der Häufigkeitsverteilung entweder nicht verändert oder dass sie sich vergrößert.

2.2 Definition der Spannweite

Die folgende Definition und die Anmerkung fassen noch einmal die wesentlichen Eigenschaften der Spannweite (des Range) zusammen.

Definition

Spannweite

Ist X ein kardinales Merkmal, dann heißt die Differenz zwischen dem größten und dem kleinsten Merkmalswert Spannweite.

$$R = x_{max} - x_{min}$$

Synonym

Range (engl: range - die Weite, die Variationsbreite, die Spanne, die Spannweite), Wertebereich.

Anmerkungen

Anmerkungen zur Spannweite

- Die Spannweite ist das einfachste Streuungsmaß. Die Aussagekraft dieser Kennzahl hängt sehr vom Datentyp ab.
- Die Spannweite nutzt die in den Daten enthaltene Information nur teilweise aus. Bei Ausreißern reagiert sie äußerst sensibel. Deshalb wird die Spannweite nicht oft als Streuungsmaß verwendet.
- Die Spannweite kennzeichnet eher den Spielraum der Beobachtungswerte als ihre Streuung.
- Die Berechnung der Spannweite ist nur bei Einzelwerten üblich (nicht bei Häufigkeitsverteilungen).
- Die Spannweite findet in der Ausdehnung eines Boxplots ihre anschauliche Darstellung.
- Die Spannweite verwendet man in der statistischen Qualitätsprüfung und bei Ausreißertests.

3 Der Quartilsabstand

Da die Spannweite lediglich vom kleinsten und vom größten Merkmalswert einer Beobachtungsreihe abhängt, führen extreme Werte immer zu extrem großen Spannweiten. Die Spannweite ist also nicht robust, sondern empfindlich gegenüber extremen Werten, sogenannten Ausreißern.

Deshalb liegt es nahe, zur Ermittlung der Streuung einer Datenreihe auf weniger extreme Werte zurückzugreifen. Versuchen wir es mit unseren Quartilen, die kennen wir schon vom Boxplot. Die Länge der Box beim Boxplot entspricht der Spannweite zwischen oberem und unterem Quartil.

Hinweis

Damit wir im Folgenden nicht immer unsere berühmte Schlange einsetzen müssen, vereinfachen wir die Bezeichnungsweise und verwenden neben $\tilde{x}_{0.25}$ und $\tilde{x}_{0.75}$ die Symbole Q_1 für das erste und Q_3 für das dritte Quartil.

Alles klar?!

Beispiel

Quartilsabstände der Punktzahlen der Klausuren von zwei Studierenden

Wie Sie schon im Beispiel mit Anna und Wong gesehen haben, gibt die Spannweite R = 30 keine geeignete Zusammenfassung der erreichten Punktzahlen (67, 82, 85, 87, 88, 89, 97) aus den 7 Klausuren von Anna an. Wenn man die beiden Ausreißer nicht berücksichtigt, würde sich die Spannweite auf R = 7 verringern.

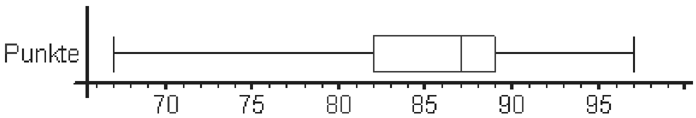


Abb: Boxplot Punktzahlen

Nun berechnen wir die folgende Differenz IQR = Q_3 - Q_1.

Mit n = 7 gilt $Q_1 = x_{(2)} = 82$. Für das obere Quartil erhalten wir den Wert $Q_3 = x_{(6)} = 89$.

Jetzt berechnen wir noch den Interquartilsrange: IQR = $Q_3 - Q_1 = 89 - 82 = 7$.

Sie können die zugehörigen Werte auch am Boxplot der Punktzahlen ablesen. Natürlich ist bei n = 7 Werten diese Darstellung im Vergleich zur Angabe der Einzelwerte keine empfehlenswerte Zusammenfassung.

Die ermittelte Differenz heißt Quartilsabstand. Das Kürzel IQR kommt wieder einmal aus dem Englischen. Es steht für „inter-quartile-range". Er hat rein gar nichts mit dem berühmten IQ zu tun!

Der Quartilsabstand der Punktzahlen 67, 82, 85, 87, 88, 89, 97 von Anna ist IQR = 7.
Zufälligerweise haben wir gleichen Ergebnisse erreicht: Der Wert 7 entspricht genau dem Wert der Spannweite ohne Berücksichtigung der beiden Ausreißer 67 und 97.
Der Quartilsabstand der Klausurpunktzahlen (67, 73, 79, 84, 88, 93, 97) errechnet sich daher mit:
IQR = 93 - 73 = 20 .

Interpretation

Zwischen dem ersten Quartil Q_1 und dem dritten Quartil Q_3 liegen 50 % aller beobachteten Merkmalsausprägungen. Der Interquartilsrange gibt die Länge dieses Bereichs an. Ein Vergleich der Quartilsabstände IQR = 20 und IQR = 7 zeigt, dass Anna konstantere Leistungen als Wong in ihren Klausuren erzielt hat.

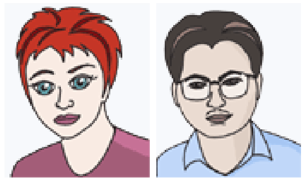
Bei Anna unterscheiden sich die Punktzahlen ihren „mittleren" Klausuren um 7 Punkte. Anscheinend hat Wong zweimal nicht genug Zeit für eine gründliche Klausurvorbereitung gefunden, mit dem Resultat eines IQR von 20.

3.1 Beispiele

Wenn Sie die folgenden Beispiele gründlich durcharbeiten, sollten Sie Range und IQR im Schlaf beherrschen.

Beispiel

Klausuren: Spannweite und Quartilsabstand



Gegeben seien die Punktzahlen der sieben Klausuren eines anderen Studierenden 69, 85, 85, 85, 85, 85, 99 .

Wir bestimmen die Spannweite und den Quartilsabstand. Die Spannweite der Punktzahlen von sieben Klausuren ist:

$R = x_{max} - x_{min} = 99 - 69 = 30.$

Der Quartilsabstand der Punktzahlen von sieben Klausuren ist:
$IQR = Q_3 - Q_1 = x_{(6)} - x_{(2)} = 85 - 85 = 0.$

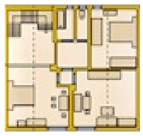
Das sind doch erstaunlich konstante Leistungen. Die Ausreißer 69 und 99 machen sich nur beim Range bemerkbar.

Richtig zur Geltung kommen die Streuungsmaße bei großen Stichproben wie das folgende Beispiel zeigen soll.

Beispiel

Mieten: IQR

Bestimmen Sie den Quartilsabstand der Mieten der 99 freien Einzimmerwohnungen in Berlin. (Mieten in €)



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

Zunächst bestimmen wir wieder Q_1 und Q_3.

Unteres Quartil Q_1:
Mit $n \cdot p = 99 \cdot 0{,}25 = 24{,}75$ folgt: $Q_1 = x_{(25)} = 368$ €.

Oberes Quartil Q_3:
Mit $n \cdot p = 99 \cdot 0{,}75 = 74{,}25$ folgt: $Q_3 = x_{(75)} = 546$ €.

Der Quartilsabstand der Mieten der 99 freien Einzimmerwohnungen in Berlin beträgt:
$IQR = Q_3 - Q_1 = x_{(75)} - x_{(25)} = 546 - 368 = 178$ € .

Interpretation

Die mittleren 50 % der Mieten unserer 99 freien Einzimmerwohnungen unterscheiden sich um maximal 178 €.

Ein Boxplot der Mieten wurde schon in der Lösung der Übung QBX-09 in Lerneinheit QBX, Abschnitt 3.3 dargestellt. Daran können Sie sich die berechneten Werte verdeutlichen.

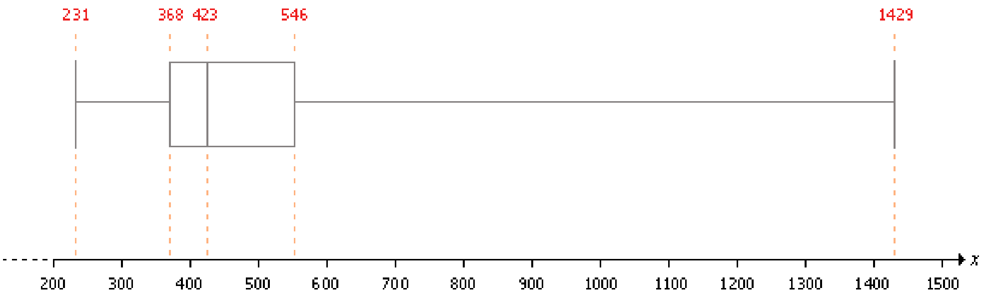


Abb.: Boxplot für die Mieten der Einzimmerwohnungen

3.2 Definition des Quartilsabstands

Das kennen Sie schon: wir schreiben alles noch einmal ganz genau auf.

Definition

Quartilsabstand

Die Differenz zwischen dem dritten Q_3 und ersten Q_1 Quartil heißt Quartilsabstand:

$$IQR = Q_3 - Q_1$$

Anmerkungen

Anmerkungen zum Quartilsabstand:

Der Quartilsabstand ist sehr robust gegenüber statistischen Ausreißern.

Quantile, Dezile

Man kann die Daten auch in andere gleich häufig besetzte Abschnitte (Quantile) zerlegen.

Bei einer Zehner-Teilung spricht man z. B. von einer Zerlegung in neun Dezile:

$D_1 = \tilde{x}_{0,1}$, $D_2 = \tilde{x}_{0,2}$,..., $D_9 = \tilde{x}_{0,9}$

Dezil, Dezilabstand

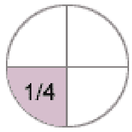
Die Differenz zwischen dem neunten D_9 und dem ersten D_1 Dezil wäre dann der Dezilabstand.

Quintile, Quintilsabstand

Bei einer Fünfer-Teilung spricht man beispielsweise von einer Zerlegung in fünf Quintile. Die Differenz zwischen dem vierten und dem ersten Quintil wäre dann der Quintilsabstand.

Quartile, Quartilsabstand

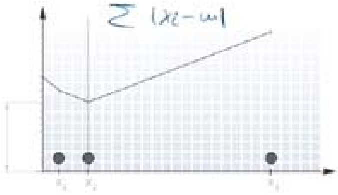
Der Quartilsabstand wird als Streuungsmaß in der explorativen Datenanalyse bei der Konstruktion eines Boxplots verwendet. Der Quartilsabstand gibt den Bereich an, in den die mittleren 50 % aller beobachteten Merkmalswerte fallen.



4 Die mittlere absolute Abweichung

Obwohl wir bei der Einführung der Varianz und der Standardabweichung gelernt haben, dass Quadratsummen nicht weh tun, gibt es natürlich die Möglichkeit, Streuung durch die mittlere absolute Abweichung zu messen.

Im folgenden Abschnitt betrachten wir nun solche Abweichungen vom arithmetischen Mittel bzw. vom Median.



Wie wählt man den Wert c so, dass die Abweichungen minimal werden? Diese Frage können wir schon beantworten! Wir haben gesehen, dass der Ausdruck

$$\frac{1}{n}\sum_{i=1}^{n}|x_i - c|$$

minimal wird, wenn für c der Median eingesetzt wird.

Die Summe der absoluten Abweichungen vom Median

$$\frac{1}{n}\sum_{i=1}^{n}|x_i - \tilde{x}_{0,5}|$$

ist daher ein Charakteristikum der Daten, analog zur mittleren quadratischen Abweichung vom arithmetischen Mittelwert.

In der Praxis haben sich (obwohl nicht ganz konsequent) sowohl die mittlere absolute Abweichung vom Median $\tilde{x}_{0,5}$ als auch vom arithmetischen Mittel $\overline{x}$ durchgesetzt.

Die mittlere absolute Abweichung vom Median heißt: $MA_{\tilde{x}_{0,5}}$.
Die mittlere absolute Abweichung vom arithmetischen Mittel heißt: $MA_{\overline{x}}$.

Die folgenden Beispiele sollen Sie mit der Berechnung vertraut machen.

4.1 Beispiel: Absolute Abweichung vom Median

Zunächst bestimmen wir unser neues Streuungsmaß an den schon hinreichend bekannten Klausurergebnissen.

Beispiel

Klausurergebnisse

Gegeben seien die sieben Messwerte
87, 78, 97, 67, 91, 92, 79 (Punktzahlen, die eine Studentin bei Klausuren erreichte) mit $\tilde{x}_{0.5} = 87$.

Wir bestimmen die mittlere absolute Abweichung $\text{MA}_{\tilde{x}_{0.5}}$ vom Median.

$$\begin{aligned}
\text{MA}_{\tilde{x}_{0,5}} &= \frac{1}{n}\sum_{i=1}^{n} |x_i - \tilde{x}_{0.5}| \\
&= \frac{1}{7}\Big(|87-87| + |78-87| + |97-87| + |67-87| \\
&\qquad + |91-87| + |92-87| + |79-87| \Big) \\
&= \frac{1}{7}(0+9+10+20+4+5+8) = \frac{56}{7} = 8
\end{aligned}$$

Im Mittel unterscheiden sich die Klausurergebnisse um 8 Punkte vom Median.

4.2 Beispiel: Mittlere absolute Abweichung vom arithmetischen Mittel

Zum Vergleich berechnen wir ebenso die mittlere Abweichung vom arithmetischen Mittel am selben Datensatz.

Beispiel

Klausurpunktzahlen

Wir bestimmen die mittlere absolute Abweichung $\text{MA}_{\overline{x}}$ vom arithmetischen Mittel der Klausurpunktzahlen: 87, 78, 97, 67, 91, 92, 79.

Das arithmetische Mittel der Messwerte
87, 78, 97, 67, 91, 92, 79 ist $\overline{x}$ = 84,4

$$\begin{aligned}
\text{MA}_{\overline{x}} &= \frac{1}{n}\sum_{i=1}^{n}|x_i - \overline{x}| \\
&= \frac{1}{7}\Big(|87 - 84{,}4| + |78 - 84{,}4| + |97 - 84{,}4| + |67 - 84{,}4| \\
&\qquad + |91 - 84{,}4| + |92 - 84{,}4| + |79 - 84{,}4|\Big) \\
&= \frac{1}{7}(2{,}6 + 6{,}4 + 12{,}6 + 17{,}4 + 6{,}6 + 7{,}6 + 5{,}4) = \frac{58{,}6}{7} \\
&= 8{,}37
\end{aligned}$$

Die mittlere absolute Abweichung vom arithmetischen Mittel beträgt $\text{MA}_{\overline{x}} = 8{,}37$.

Für diese sieben Klausuren beträgt die mittlere absolute Abweichung vom Median $\text{MA}_{\tilde{x}_{0,5}} = 8$, die mittlere absolute Abweichung vom arithmetischen Mittel $\text{MA}_{\overline{x}} = 8{,}37$ und die Standardabweichung s = 9,18.

Ein Vergleich von $\text{MA}_{\tilde{x}_{0,5}}$, $\text{MA}_{\overline{x}}$ und s zeigt:

$$\text{MA}_{\tilde{x}_{0,5}} \leq MA_{\overline{x}} \leq s$$

Anmerkung

Diese Relation gilt übrigens immer. Zum Beweis würden wir aber eine Überdosis Mathematik benötigen. Wir verzichten lieber darauf.
Genügt ein Ehrenwort?

4.3 Beispiel: Punktzahlen von Studierenden

Beispiel

Die Punktzahlen, die unser Student Wong in sieben Klausuren erreichte, waren: 84, 73, 97, 67, 88, 93 und 79. Wong erreichte: 82, 67, 85, 97, 87, 89, 88 Punkte.
Man bestimme die mittlere absolute Abweichung vom Median $\text{MA}_{\tilde{x}_{0,5}}$ und die mittlere absolute Abweichung vom arithmetischen Mittel $\text{MA}_{\overline{x}}$.

Wong:

Mittlere absolute Abweichung vom Median $\text{MA}_{\tilde{x}_{0,5}}$:

$$\begin{aligned}\text{MA}_{\tilde{x}_{0,5}} &= \frac{1}{n}\sum_{i=1}^{n}|x_i - \tilde{x}_{0,5}| \\ &= \frac{1}{7}\Big(|84-84| + |73-84| + |97-84| + |67-84| \\ &\qquad + |88-84| + |93-84| + |79-84|\Big) \\ &= 8{,}43\end{aligned}$$

Mittlere absolute Abweichung vom arithmetischen Mittel $\text{MA}_{\overline{x}}$:

$$\begin{aligned}\text{MA}_{\overline{x}} &= \frac{1}{n}\sum_{i=1}^{n}|x_i - \overline{x}| \\ &= \frac{1}{7}\Big(|84-83| + |73-83| + |97-83| + |67-83| \\ &\qquad + |88-83| + |93-83| + |79-83|\Big) \\ &= 8{,}57\end{aligned}$$

Anna:

Mittlere absolute Abweichung vom Median $\text{MA}_{\tilde{x}_{0,5}}$:

$$\begin{aligned}\text{MA}_{\tilde{x}_{0,5}} &= \frac{1}{n}\sum_{i=1}^{n}|x_i - \tilde{x}_{0,5}| \\ &= \frac{1}{7}\Big(|82-87| + |67-87| + |85-87| + |97-87| \\ &\qquad + |87-87| + |89-87| + |88-87|\Big) \\ &= 5{,}71\end{aligned}$$

Mittlere absolute Abweichung vom arithmetischen Mittel $\text{MA}_{\overline{x}}$:

$$\begin{aligned}\text{MA}_{\overline{x}} &= \frac{1}{n}\sum_{i=1}^{n}|x_i - \overline{x}| \\ &= \frac{1}{7}\Big(|82-85| + |67-85| + |85-85| + |97-85| \\ &\qquad + |87-85| + |89-85| + |88-85|\Big) \\ &= 6{,}00\end{aligned}$$

Anmerkung

Überlegen Sie doch bitte, warum $\text{MA}_{\tilde{x}_{0,5}}$ immer kleiner gleich $\text{MA}_{\overline{x}}$ gelten muss.

4.4 Definition der mittleren absoluten Abweichung

Nachdem wir die Möglichkeiten zur Ermittlung der absoluten Abweichungen vom Median und vom arithmetischen Mittel kennen gelernt haben, schreiben wir (wie immer) die neuen Streuungsmaße noch einmal ganz genau auf.

Definition

Mittlere absolute Abweichung vom Median und vom arithmetischen Mittel

Ist X ein kardinales Merkmal, dann heißt die Maßzahl

$\mathrm{MA}_{\tilde{x}_{0,5}} = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \tilde{x}_{0,5}\right|$ mittlere absolute Abweichung vom Median und

$\mathrm{MA}_{\overline{x}} = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \overline{x}\right|$ mittlere absolute Abweichung vom arithmetischen Mittel.

Synonyme

Mittlere absolute Abweichung vom Median $\mathrm{MA}_{\tilde{x}_{0,5}}$

(mean absolute deviation from the median)

Mittlere absolute Abweichung vom arithmetischen Mittel $\mathrm{MA}_{\overline{x}}$

(mean absolute deviation from the mean)

Anmerkungen

Anmerkungen zur mittleren absoluten Abweichung

- Mittlere absolute Abweichungen sind besonders anschaulich: Sie geben den durchschnittlichen Abstand einer Beobachtung vom Median bzw. vom arithmetischen Mittel an.
- Es gilt stets: $\mathrm{MA}_{\tilde{x}_{0,5}} \leq \mathrm{MA}_{\overline{x}} \leq s$

4.5 Beispiele: Reklamationen und Garne

Betrachten wir nun zwei Beispiele deren Daten Sie schon aus anderen Lerneinheiten kennen.

Beispiel

Reklamationen

In einer Weinhandlung wurden von Januar 2021 bis Dezember 2022 monatlich folgende Anzahl von Reklamationen registriert. Wir ignorieren den zeitlichen Verlauf der Reklamationen.

	Jan.	Feb.	März	April	Mai	Juni	Juli	Aug.	Sept.	Okt.	Nov.	Dez.
2021	0	4	1	4	2	8	2	4	3	2	1	6
2022	7	1	6	1	9	7	8	8	4	3	7	0

Tab.: Beispiel für registrierte Reklamationen

Bestimmen wir zunächst den Quartilsabstand der Erhebung.Der Quartilsabstand der Reklamationen:

IQR= $\tilde{x}_{0{,}75}$ - $\tilde{x}_{0{,}25}$ = 7 - 1,5 = 5,5

Interpretation
Für die Hälfte der Monate liegt die Anzahl der Reklamationen zwischen 7 und 1,5.

Beispiel

Reklamationen: Absolute Abweichung vom Median

Wir bestimmen die mittlere absolute Abweichung vom Median.

$$\begin{aligned}
\text{MA}_{\tilde{x}_{0{,}5}} &= \frac{1}{n}\sum_{i=1}^{n}|x_i - \tilde{x}_{0,}| \\
&= \frac{1}{24}\Big(|0-4|+|4-4|+|1-4|+|4-4|+|2-4|+|8-4| \\
&\quad + |2-4|+|4-4|+|3-4|+|2-4|+|1-4|+|6-4| \\
&\quad + |7-4|+|1-4|+|6-4|+|1-4|+|9-4|+|7-4| \\
&\quad + |8-4|+|8-4|+|4-4|+|3-4|+|7-4|+|0-4|\Big) \\
&= \frac{1}{24}(4+0+3+0+2+4+2+0+1+2+3+2+3+3 \\
&\quad +2+3+5+3+4+4+0+1+3+4) \\
&= \frac{1}{24}\cdot 58 = 2{,}42
\end{aligned}$$

Interpretation
Im Mittel unterscheidet sich die Anzahl der Reklamationen um 2,42 vom Median (4).

Beispiel

Reißfestigkeit verschiedener Garnsorten

Wir haben uns in der Lerneinheit VAR mit der Reißfestigkeit von zwei Garnsorten beschäftigt. Sie haben gesehen, dass Garn2 deutlich stärker als Garn1 streut.

Wir wollen die verschiedenen Streuungsmaße, die wir in dieser Lerneinheit kennengelernt haben für das Beispiel an dieser Stelle gegenüberstellen.

Range und Interquartilsrange ergeben sich direkt aus der 5-Punkte-Zusammenfassung, zusätzlich wurden die mittleren absoluten Abweichungen von Median und Mittelwert bestimmt.

5-Punkte Zusammenfassung

Tab.: Beispiel für 5-Punkte Zusammenfassung

	Min	Q_1	Median	Q_3	Max
Garn1	3,8	4,7	4,99	5,42	6,0
Garn2	1,6	4,45	5,77	7,46	10,39

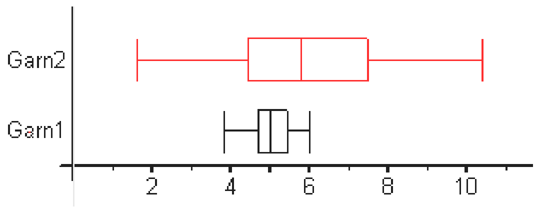


Abb: Boxplot Reißfähigkeit

In der folgenden Tabelle sind alle Streuungsmaße, die wir betrachtet haben, für beide Garnsorten ausgerechnet.

Verschiedene Streuungsmaße

Tab.: Beispiel für verschiedene Streuungsmaße

	$\overline{x}$	S	Range	IQR	MA - $\tilde{x}_{0,5}$	MA - $\overline{x}$
Garn1	5,01	0,49	2,2	0,72	0,39	0,39
Garn2	5,91	2,17	8,79	3,01	1,78	1,79

Alle berechneten Streuungsmaße zeigen den gleichen Sachverhalt:
Die Variabilität bei Garn2 beträgt etwas mehr als das Vierfache der Variabilität von Garn1.

5 Übungsaufgaben

Hinweis

Im folgenden Beispiel haben wir absichtlich eine sehr kleine Anzahl von Beobachtungen gewählt. Sie wissen inzwischen, wie Sie für größere Datenmengen die Statistiksoftware **R** als Handwerkszeug einsetzen können.

Hier sollen Sie vor allem Ihr Formelverständnis vertiefen. Machen Sie sich keine Gedanken, ob z. B. die Angabe des Quartils bei 4 Beobachtungen sinnvoll ist.

Sie werden in den folgenden Übungen verschiedene Kenngrößen für die in der ersten Übung angegebenen Einkommen bestimmen.

Berechnen

Übung ASM-01

Einkommen - Mittelwert und Median

Für die Mitarbeiter und Mitarbeiterinnen in den vier Abteilungen unserer Weinhandlung ergibt sich je nach Einsatzgebiet und Dauer der Betriebszugehörigkeit folgendes Einkommen (in €) pro Monat:

Einkommen (in €)	Person 1	Person 2	Person 3	Person 4	Person 5	Person 6	Person 7
Abteilung 1	1350	1720	1550	1200			
Abteilung 2	1470	1400	1680	1950	2000		
Abteilung 3	1420	1800	1580	2100	1900	2300	
Abteilung 4	1500	1760	2660	1890	2910	2280	2400

Bestimmen und interpretieren Sie für jede Abteilung das Durchschnittseinkommen und den Median.

Lösung (Siehe Anhang)

Bearbeitungszeit: 15 Minuten

Berechnen

Übung ASM-02

Einkommen - Spannweite

Bestimmen und interpretieren Sie für jede Abteilung die Spannweite.

Lösung (Siehe Anhang)

Bearbeitungszeit: 10 Minuten

Berechnen

Übung ASM-03

Einkommen - IQR

Bestimmen und interpretieren Sie für jede Abteilung den Quartilsabstand.

Lösung (Siehe Anhang)

Bearbeitungszeit: 15 Minuten

Berechnen

Übung ASM-04

Einkommen - MA Median

Bestimmen und interpretieren Sie für jede Abteilung die mittlere absolute Abweichung vom Median.

Lösung (Siehe Anhang)

Bearbeitungszeit: 15 Minuten

Berechnen

Übung ASM-05

Einkommen - MA Mittelwert

Bestimmen und interpretieren Sie für jede Abteilung die mittlere absolute Abweichung vom arithmetischen Mittel.

Lösung (Siehe Anhang)

Bearbeitungszeit: 15 Minuten

Zusammenfassung

- Neben Varianz und Standardabweichung kommen auch alternative Streuungsmaße zur Anwendung.
- Die Spannweite charakterisiert eher alle möglichen Ausprägungen und ist deshalb als Streuungsmaß nur begrenzt sinnvoll.
- Der Quartilsabstand IQR entspricht der Spannweite zwischen erstem und drittem Quartil.
- Der Quartilsabstand gibt die Streuung der zentralen Hälfte der Daten an und ist somit robust, d. h. unempfindlich gegenüber Ausreißern.
- Mittlere absolute Abweichungen können an Stelle der Standardabweichung verwendet werden.
- Der Median minimiert die Summe der absoluten Abweichungen der Einzelwerte.
- Die genannten Streuungsmaße verhalten sich bei linearen Transformationen genau wie die Standardabweichung.

Sie sind am Ende dieser Lerneinheit angelangt. Auf der folgenden Seite finden Sie noch die Übungen zur Wissensüberprüfung und weitere Übungen.

Wissensüberprüfung

Multiple Choice

Übung ASM-06

Welche Aussagen sind falsch und welche richtig?

	Richtig	Falsch	Auswertung
Die Berechnung der Spannweite ist nur bei Einzelwerten üblich.			
Die Spannweite wird kleiner, wenn die Zahl der Werte größer wird.			
Der Quartilsabstand gibt die Länge des Intervalls an, in das die mittleren 50% aller beobachteten Merkmalswerte fallen.			
Die mittlere absolute Abweichung vom arithmetischen Mittel kann negativ sein.			
Die Spannweite ergibt Null, wenn die Differenz zwischen dem größten und dem kleinsten Merkmalswerte Null ist.			

Multiple Choice

Übung ASM-07

Welche Aussagen sind falsch und welche richtig?

	Richtig	Falsch	Auswertung
Die Summe des Betrags der Abweichungen der Einzelwerte vom Median ist minimal. Deshalb ist die Standardabweichung das zum Median passende Streuungsmaß.			
Der Quartilsabstand wird als Streuungsmaß in der explorativen Datenanalyse bei der Konstruktion eines Box-Plots verwendet.			
Die Spannweite findet in der Ausdehnung eines Box-Plots ihre anschauliche Darstellung.			
Der Quartilsabstand, genauso wie der Mittelwert, ist sehr empfindlich gegenüber den statistischen Ausreißern.			
Die Spannweite kennzeichnet mehr den Spielraum der Beobachtungswerte als ihre Streuung.			

Übungen mit der Statistiksoftware R

Statistiksoftware R

Die in der Lerneinheit behandelten Themen können Sie anhand der folgenden Übungsaufgaben mit der Statistiksoftware **R** bearbeiten. Die Aufgabenstellung finden Sie in der jeweiligen Übung. Um die Übungen zu bearbeiten, muss die Software „**R**“ auf Ihrem Rechner installiert sein.

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

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

Berechnen

Übung ASM-08a

Spannweite von Mieten

Die Datei **`kdaten.txt`** enthält eine Liste mit Mietpreisen von 1-3 Zimmerwohnungen. Leider nur ein Beispiel.

`kdaten.txt` (2 KB)

1. Bestimmen Sie die Spannweite der Mieten der Wohnungen und interpretieren Sie das Ergebnis.

Lösung mit R und Excel (Siehe Anhang)

Bearbeitungszeit: 20 Minuten

Berechnen

Übung ASM-08b

Mieten - Quartilsabstand

Betrachten wir noch einmal den Datensatz aus der Datei **`kdaten.txt`**.

`kdaten.txt` (2 KB)

1. Bestimmen Sie für jede Zimmeranzahl den Quartilsabstand der Mieten und interpretieren Sie das Ergebnis.

Lösung mit R und Excel (Siehe Anhang)

Bearbeitungszeit: 20 Minuten

Berechnen

Übung ASM-08c

Klausurnoten - Absolute Abweichung vom Median

Wir betrachten wieder sieben Punktzahlen die eine Studentin bei den Klausuren erzielt hat.

87, 78, 97, 67, 91, 92, 79

1. Bestimmen Sie die mittlere absolute Abweichung vom Median.

Lösung mit R und Excel (Siehe Anhang)

Bearbeitungszeit: 20 Minuten

Berechnen

Übung ASM-08d

Klausurnoten - Interquartilsrange

Ein Student erreicht in seinen Klausuren folgende Punktzahlen:

67, 82, 85, 87, 88, 89, 97.

1. Berechnen Sie die Spannweite und den Quartilsabstand bevor und nachdem Sie die Ausreißer 67 und 97 gestrichen haben.
2. Interpretieren Sie die Ergebnisse.

Lösung mit R und Excel (Siehe Anhang)

Bearbeitungszeit: 20 Minuten

Berechnen

Übung ASM-08e

Spannweite bei den Klausurergebnissen von Anna und Wong

Die Studierenden Anna und Wong haben in diesem Semester 7 Klausuren geschrieben.

Wong erreichte dabei 84, 73, 97, 67, 88, 93 und 79 Punkte,
Anna 82, 67, 85, 97, 87, 89 und 88 Punkte.

1. Bestimmen Sie die Spannweite der Punktzahlen von Anna und Wong.
2. Interpretieren Sie die Ergebnisse.

Das Beispiel hatten Sie im vorderen Teil schon einmal gerechnet. Versuchen Sie nun noch die Lösung mit R zu finden.

Lösung mit R und Excel (Siehe Anhang)

Bearbeitungszeit: 20 Minuten

Zusätzliche Übungsaufgaben

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

Berechnen

Übung ASM-09

Spannweite: Körpergröße

Bestimmen Sie die Spannweite der folgenden Körpergrößen:

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

Lösung (Siehe Anhang)

Bearbeitungszeit: 2 Minuten

Berechnen

Übung ASM-10

Spannweite: Studierendenalter

Bestimmen Sie die Spannweite des Alters von 25 Studierenden:

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

Lösung (Siehe Anhang)

Bearbeitungszeit: 2 Minuten

Berechnen

Übung ASM-11

Spannweite: Dummy-Zahlen

Bestimmen Sie die Spannweite für die folgenden Zahlen:

6,3	6,7	7,6	5,8	5,9	7,4	5,6	7,2	6,7	6,3
6,5	7,2	7,1	5,7	6,4	6,8	7,2	5,8	6,4	6,2

Lösung (Siehe Anhang)

Bearbeitungszeit: 2 Minuten

Berechnen

Übung ASM-12

Spannweite: Tagestemperaturen

Folgende Tagestemperaturen wurden in den Monaten Juni und Juli gemessen:

Juni	14	14,5	17	18	19,5	23	24	24,5	25	25	26	27
Juli	18	19,5	23	23	25	25	25	26	26,5	27	28	28,5

Bestimmen Sie die Spannweite der beiden Monate.

Lösung (Siehe Anhang)

Bearbeitungszeit: 4 Minuten

Berechnen

Übung ASM-13

Master und Bachelor: Gehälter im Vergleich

In dieser Übung wollen wir die Gehälter von Fachhochschulabsolventen mit Master und Bachelor Abschlüssen vergleichen. Es wurden mehrere Personen nach ihrem Abschluss, Fachrichtung und Monatsgehalt (in €) befragt und in der unten folgenden Tabelle notiert.

Bachelor

BWL	3100	2900	2999	2780	2850
Naturwissenschaften	2600	2450	2700	2650	2900
Jura	3050	3400	2800	2950	3100
Ingenieure / Ingenieurinnen	3100	3400	3350	3500	3250
Geisteswissenschaften	2100	2320	2400	2150	2330

Master

BWL	3300	3550	3680	3450	3700
Naturwissenschaften	3200	3520	3480	3720	3490
Jura	3900	3780	3400	3850	3650
Ingenieure / Ingenieurinnen	3800	4100	3780	3890	4300
Geisteswissenschaften	3100	2750	2650	2900	3050

1. **Mittelwert und Median**
   Bestimmen und interpretieren Sie für jeden Abschluss und deren Fachrichtung, das Durchschnittseinkommen und den Median.
2. **Spannweite**
   Bestimmen und interpretieren Sie für jeden Abschluss und deren Fachrichtung die Spannweite.
3. **IQR**
   Bestimmen und interpretieren Sie für jede Abteilung den Quartilsabstand.
4. **MA Median**
   Bestimmen und interpretieren Sie für jede Abteilung die mittlere absolute Abweichung vom Median.
5. **MA Mittelwert**
   Bestimmen und interpretieren Sie für jede Abteilung die mittlere absolute Abweichung vom arithmetischen Mittel.

Lösung (Siehe Anhang)

Bearbeitungszeit: 20 Minuten

Appendix

Lösung für Übung ASM-01

Einkommen - Mittelwert und Median

Für das arithmetische Mittel oder den Mittelwert $\overline{x}$ der Einkommen in den vier Abteilungen ergibt sich:

- für **Abteilung 1** ein Betrag (in €) von

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

$$= \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{4}\sum_{i=1}^{4} x_i$$

$$= \frac{1}{4}(1350 + 1720 + 1550 + 1200)$$

$$= \frac{1}{4} \cdot 5820 = 1455$$

- für **Abteilung 2** ein Betrag (in €) von

$$\overline{x} = \frac{1}{5}\sum_{i=1}^{5} x_i$$

$$= \frac{1}{5}(1470 + 1400 + 1680 + 1950 + 2000)$$

$$= \frac{1}{5} \cdot 8500 = 1700$$

- für **Abteilung 3** ein Betrag (in €) von

$$\overline{x} = \frac{1}{6}\sum_{i=1}^{6} x_i \quad = \frac{1}{6}(1420 + 1800 + 1580 + 2100 + 1900 + 2300)$$

$$= \frac{1}{6} \cdot 111000 = 1850$$

- für **Abteilung 4** ein Betrag von

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

$$= \frac{1}{7}(1500 + 1760 + 2660 + 1890 + 2910 + 2280 + 2400)$$

$$= \frac{1}{7} \cdot 15400 = 2200$$

Interpretation
Das Durchschnittseinkommen der Mitarbeiter in den jeweiligen Abteilungen beträgt: 1455 €, 1700 €, 1850 € bzw. 2200 €.

Median
Um den Median bestimmen zu können, müssen die Einkommen der vier Abteilungen zunächst der Größe nach geordnet und dann nach folgendem Ausdruck errechnet werden:

$$Z = \tilde{x}_{0.5} = \begin{cases} x_{\left(\frac{(n+1)}{2}\right)} & \text{falls } n \text{ ungerade} \\ \frac{1}{2}\left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{(n+1)}{2}\right)}\right) & \text{falls } n \text{ gerade} \end{cases}$$

Für Abteilung 1:
Mit $\left(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}\right)$ = (1200, 1350, 1550, 1720) und n = 4 (gerade Zahl) ist

$$Z = \frac{1}{2}\left(x_{\left(\frac{n}{2}\right)} + x_{\left(\left(\frac{n}{2}\right)+1\right)}\right) = \frac{1}{2}\left(x_{\left(\frac{4}{2}\right)} + x_{\left(\left(\frac{4}{2}\right)+1\right)}\right)$$

$$= \frac{1}{2}\left(x_{(2)} + x_{(3)}\right) = \frac{1}{2}(1350 + 1550) = 1450$$

Für Abteilung 2:
Mit $\left(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}, x_{(5)}\right)$ = (1400, 1470, 1680, 1950, 2000) und n = 5 (ungerade Zahl) ist

$$Z = x_{\left(\frac{(n+1)}{2}\right)} = x_{\left(\frac{(5+1)}{2}\right)} = x_{\left(\frac{6}{2}\right)} = x_{(3)} = 1680$$

Für Abteilung 3:
Mit $\left(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}, x_{(5)}, x_{(6)}\right)$= (1420, 1580, 1800, 1900, 2100, 2300) und n = 6 (gerade Zahl) ist

$$Z = \frac{1}{2}\left(x_{\left(\frac{n}{2}\right)} + x_{\left(\left(\frac{n}{2}\right)+1\right)}\right) = \frac{1}{2}\left(x_{\left(\frac{6}{2}\right)} + x_{\left(\left(\frac{6}{2}\right)+1\right)}\right)$$

$$= \frac{1}{2}\left(x_{(3)} + x_{(4)}\right) = \frac{1}{2}(1800 + 1900) = 1850$$

Für Abteilung 4:
Mit $\left(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}, x_{(5)}, x_{(6)}, x_{(7)}\right)$ = (1500, 1760, 1890, 2280, 2400, 2660, 2910) und n = 7 (ungerade Zahl) ist

$$Z = x_{\left(\frac{(n+1)}{2}\right)} = x_{\left(\frac{(7+1)}{2}\right)} = x_{\left(\frac{8}{2}\right)} = x_{(4)} = 2280$$

Interpretation
50 % der Einkommen in den jeweiligen Abteilungen betragen höchstens: 1.450 €, 1.680 €, 1.850 € und 2.280 € bzw.

50 % der Einkommen in den jeweiligen Abteilungen betragen mindestens 1.450 €, 1.680 €, 1.850 € und 2.280 €.

Lösung für Übung ASM-02

Einkommen - Spannweite

Die Spannweite R, die die Breite des Streubereichs der kardinalen Merkmalsausprägungen angibt, ist definiert als R = x_{max} - x_{min} , wobei x_{max} den größten und x_{min} den kleinsten Merkmalswert bezeichnet.

Um die Spannweite bestimmen zu können, müssen die Einkommenswerte der vier Abteilungen zunächst der Größe nach geordnet werden. Die Spannweite beträgt jeweils

Für Abteilung 1: Mit $\left(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}\right)$ = (1200,1350,1550,1720)

R = x_{max} - x_{min} =1720 - 1200 = 520 ,

Für Abteilung 2: Mit $\left(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}, x_{(5)}\right)$ = (1400,1470,1680,1950,2000)

R = x_{max} - x_{min} = 2000 - 1400 = 600 ,

Für Abteilung 3: Mit $\left(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}, x_{(5)}, x_{(6)}\right)$ = (1420,1580,1800,1800,2100,2300)

R = x_{max} - x_{min} = 2300 - 1420 = 880,

Für Abteilung 4: Mit $\left(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}, x_{(5)}, x_{(6)}, x_{(7)}\right)$ = (1500,1760,1890,2280,2400,2660,2910)

R = x_{max} - x_{min} = 2910 - 1500 = 1410 .

Interpretation:
Die Spannweite oder die Breite eines Streubereichs für die Einkommen in den jeweiligen Abteilungen beträgt: 520 €, 600 €, 880 € und 1410 €.

Lösung für Übung ASM-03

Einkommen - IQR

Um den Quartilsabstand bestimmen zu können, müssen die Mitarbeitereinkommen in den 4 Abteilungen zunächst der Größe nach geordnet werden und können dann mit der Formel
IQR = Q_3 - Q_1 berechnet werden.

Q_1 bezeichnet dabei das 1. Quartil $\tilde{x}_{0,25}$ und Q_3 das 3. Quartil $\tilde{x}_{0,75}$.

Für Abteilung 1:
$\left(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}\right)$ = (1200,1350,1550,1720)

Q_3 = 1635, Q_1 =1275

IQR= Q_3 - Q_1 = 1635 - 1275 = 360

Denn: mit n = 4 und n × p = 4 x 0.75 = 3 (ganzzahlig) bzw. n × p = 4 x 0.25 = 1 (ganzzahlig) ist

$$Q_3 = \frac{1}{2}\left[x_{(np)} + x_{(np+1)}\right] = \frac{1}{2}\left[x_{(3)} + x_{(4)}\right] = \frac{1}{2}(1550 + 1720) = 1635$$

und

$$Q_1 = \frac{1}{2}\left[x_{(np)} + x_{(np+1)}\right] = \frac{1}{2}\left[x_{(1)} + x_{(2)}\right] = \frac{1}{2}(1200 + 1350) = 1275$$

Für Abteilung 2:
$\left(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}, x_{(5)}\right)$ = (1400,1470,1680,1950,2000)

Q_3 = 1950, Q_1 =1470

IQR= Q_3 - Q_1 = 1950 - 1470 = 480

Denn: mit n = 5 und n · p = 5 · 0,75 = 3,75 (nicht ganzzahlig) bzw. n · p = 5 · 0,25 = 1,75 (nicht ganzzahlig) ist $Q_3 = x_{(4)}$ = 1950 und $Q_1 = x_{(2)}$ = 1470 .

Für Abteilung 3:
$\left(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}, x_{(5)}, x_{(6)}\right)$ = (1420,1580,1800,1800,2100,2300)

Q_3 = 2100, Q_1 =1580

IQR= Q_3 - Q_1 = 2100 - 1580 = 520

Denn: mit n = 6 und n · p = 6 · 0,75 = 4,5 (nicht ganzzahlig) bzw. n · p = 6 · 0,25 = 1,5 (nicht ganzzahlig) ist $Q_3 = x_{(5)}$ = 2100 und $Q_1 = x_{(2)}$ 1580 .

Für Abteilung 4:
$\left(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}, x_{(5)}, x_{(6)}, x_{(7)}\right)$= (1500,1760,1890,2280,2400,2660,2910)

Q_3 = 2660, Q_1 =1760

IQR= Q_3 - Q_1 = 2660 - 1760 = 900

Denn: mit n = 7 und n · p = 7 · 0,75 = 5,25 (nicht ganzzahlig) bzw. n · p = 7 · 0,25 = 1,75 (nicht ganzzahlig) ist $Q_3 = x_{(6)}$ = 2660 und $Q_1 = x_{(2)}$ = 1760.

Interpretation:
Zwischen dem ersten Quartil Q_1 und dem dritten Quartil Q_3 liegen 50 % aller Einkommenswerte in den jeweiligen Abteilungen.

Der Quartilsabstand liefert somit die Größe des Bereichs, in dem etwa die Hälfte aller Einkommenswerte in den jeweiligen Abteilungen liegen. Beispielsweise liegen 50 % aller Einkommen in Abteilung 1 zwischen 1.635 € und 1.275 € usw.

Lösung für Übung ASM-04

Einkommen - MA Median

Die mittlere absolute Abweichung vom Median $MA_{\tilde{x}_{0,5}}$ in der Abteilung 1 beträgt:

$$\begin{aligned}
\text{MA}_{\tilde{x}_{0,5}} &= \frac{1}{n}\sum_{i=1}^{n}|x_i - \tilde{x}_{0,5}| \\
&= \frac{1}{4}\big(|1350-1450| + |1720-1450| + |1550-1450| + |1200-1450|\big) \\
&= \frac{1}{4}(100+270+100+250) = \frac{720}{4} = 180
\end{aligned}$$

Die mittlere absolute Abweichung vom Median $\text{MA}_{\tilde{x}_{0,5}}$ beträgt:

in Abteilung 2:

$$\begin{aligned}
\text{MA}_{\tilde{x}_{0,5}} &= \frac{1}{n}\sum_{i=1}^{n}|x_i - \tilde{x}_{0,5}| \\
&= \frac{1}{5}\big(|1470-1680| + |1400-1680| + |1680-1680| + |1950-1680| \\
&\quad + |2000-1680|\big) \\
&= \frac{1}{5}(210+280+0+270+320) = \frac{1080}{5} = 216
\end{aligned}$$

in Abteilung 3:

$$\begin{aligned}
\text{MA}_{\tilde{x}_{0,5}} &= \frac{1}{n}\sum_{i=1}^{n}|x_i - \tilde{x}_{0,5}| \\
&= \frac{1}{6}\big(|1420-1850| + |1800-1850| + |1580-1850| + |2100-1850| \\
&\quad + |1900-1850| + |2300-1850|\big) \\
&= \frac{1}{6}(430+50+270+250+50+450) = \frac{1500}{6} = 250
\end{aligned}$$

und in Abteilung 4:

$$\begin{aligned}
\text{MA}_{\tilde{x}_{0,5}} &= \frac{1}{n}\sum_{i=1}^{n}|x_i - \tilde{x}_{0,5}| \\
&= \frac{1}{7}\big(|1500-2280| + |1760-2280| + |2660-2280| + |1890-2280| \\
&\quad + |2910-2280| + |2280-2280| + |2400-2280|\big) \\
&= \frac{1}{7}(780+520+380+390+630+0+120) = \frac{2820}{7} = 402{,}86
\end{aligned}$$

Interpretation:

Die mittlere absolute Abweichung vom Median ($MA_{\tilde{x}_{0,5}}$) ist ein Durchschnittswert, der aus den Abständen der einzelnen Einkommen vom Median berechnet wird.

Lösung für Übung ASM-05

Einkommen - MA Mittelwert

Für die mittlere absolute Abweichung vom Mittelwert können Sie sich an der Lösung zu Übung ASM-04 orientieren. Ersetzen Sie dort einfach den Median durch das arithmetische Mittel je Abteilung.

Die mittlere absolute Abweichung vom arithmetischen Mittel beträgt:
Abteilung 1: 180
Abteilung 2: 220
Abteilung 3: 250
Abteilung 4: 414,3

Lösung Übung ASM-08a

Spannweiten von Mieten

Anzahl Zimmer	Spannweite	Quartilsabstand	Häufigkeit
1	241	151	28
2	1198	172,25	102
3	2853	451	61
4	1926	596	39
5	3081	580	11
6	1227	495	9
7	1070	535	3
8	0	0	1

Interpretation: Vergleichbarkeit ist nicht gegeben, da die Häufigkeiten der Zimmeranzahlen sehr unterschiedlich sind und bei höheren Mietpreisen sowieso mit höherer Spannweite zu rechnen ist. Um das zu verdeutlichen können die Häufigkeiten betrachtet werden.

Lösung mit R

kdaten_loesung.R

```
# Einlesen der Daten aus der Datei kdaten.txt
mietwohnungen<-read.table("kdaten.txt",sep="\t",header=TRUE)

#  Aufspaltung des Data Frames nach Zimmeranzahl
mietwohnungen1<-subset(mietwohnungen,Zimmer==1)
mietwohnungen2<-subset(mietwohnungen,Zimmer==2)
mietwohnungen3<-subset(mietwohnungen,Zimmer==3)
mietwohnungen4<-subset(mietwohnungen,Zimmer==4)
mietwohnungen5<-subset(mietwohnungen,Zimmer==5)
mietwohnungen6<-subset(mietwohnungen,Zimmer==6)
mietwohnungen7<-subset(mietwohnungen,Zimmer==7)
mietwohnungen8<-subset(mietwohnungen,Zimmer==8)

#  Bestimmung der Spannweite der Mieten nach Zimmeranzahl
max(mietwohnungen1$Preis)-min(mietwohnungen1$Preis)
max(mietwohnungen2$Preis)-min(mietwohnungen2$Preis)
max(mietwohnungen3$Preis)-min(mietwohnungen3$Preis)
max(mietwohnungen4$Preis)-min(mietwohnungen4$Preis)
max(mietwohnungen5$Preis)-min(mietwohnungen5$Preis)
max(mietwohnungen6$Preis)-min(mietwohnungen6$Preis)
max(mietwohnungen7$Preis)-min(mietwohnungen7$Preis)
max(mietwohnungen8$Preis)-min(mietwohnungen8$Preis)
```

Lösung mit Excel

WMS_ASM_08_Mieten.xlsx (34 KB)

Aufgabe 1: *Bestimmen Sie für jede Zimmeranzahl die Spannweite der Mieten in Berlin und interpretieren Sie das Ergebnis.*

Um die Mieten sortiert nach ihrer Zimmeranzahl betrachten zu können, ist in Excel ein zusätzlicher Schritt nötig. Aufgrund des Umfangs der Tabelle kann das hier nicht gezeigt werden. Es ist aber in der Excel Datei gut zu erkennen, was gemacht wurde.

Für alle vorkommenden Zimmeranzahlen, also 1 bis 8, wurde eine zusätzliche Spalte an die Tabelle angehängt. In diesen Spalten fragen wir mit der Funktion `WENN` jeweils ab, was die Zimmeranzahl in der entsprechenden Zeile ist. Der Preis wird dann in die jeweils richtige Spalte eingetragen, alle anderen Spalten bleiben in dieser Zeile leer. So können nachher Spaltenweise die Zimmer getrennt nach ihrer Zimmeranzahl betrachtet werden. Ausschnittweise ist in der Abbildung unten zu sehen, wie das geht.

D2 | f_x | =WENN(A2=1;B2;"")

	B	C	D	E	F	
1	Preis		Preis 1 Zimmer	Preis 2 Zimmer	Preis 3 Zimmer	Preis 4
2	203		=WENN(A2=1;B2;"	=WENN(A2=2;B2;"")	=WENN(A2=3;B2;"")	=WENI
3	206		=WENN(A3=1;B3;"	=WENN(A3=2;B3;"")	=WENN(A3=3;B3;"")	=WENI
4	213		=WENN(A4=1;B4;"	=WENN(A4=2;B4;"")	=WENN(A4=3;B4;"")	=WENI
5	213		=WENN(A5=1;B5;"	=WENN(A5=2;B5;"")	=WENN(A5=3;B5;"")	=WENI
6	216		=WENN(A6=1;B6;"	=WENN(A6=2;B6;"")	=WENN(A6=3;B6;"")	=WENI
7	221		=WENN(A7=1;B7;"	=WENN(A7=2;B7;"")	=WENN(A7=3;B7;"")	=WENI

In der Klammer des Funktionsaufrufs steht zunächst die Bedingung. Dann steht als zweites, was in die markierte Zelle eingetragen wird, wenn die Bedingung erfüllt ist. Als drittes steht in der Klammer, was in die markierte Zelle eingetragen wird, wenn die Bedingung nicht erfüllt ist. In unserem Fall hier bleibt die Zelle dann leer.

Nun können mit den Funktionen `MIN`, `MAX` und `QUARTILE.INKL` – wie in folgender Abbildung zu sehen – für jede Zimmeranzahl Minimum, Maximum, 1. Quartil und 3. Quartil berechnet werden. Die Spannweite ergibt sich aus der Differenz zwischen Maximum und Minimum, der Quartilsabstand aus der Differenz zwischen 3. Quartil und 1. Quartil.

F288 | f_x |

	A	B	C	D	E	F	G
274	Ar	Minimum	Maximum	Spannweite	1. Quartil	3. Quartil	Quartilsabsta
275	1	=MIN(D2:D255)	=MAX(D2:D255)	=C275-B275	=QUARTILE.INKL(D2:D255;1)	=QUARTILE.INKL(D2:D255;3)	=F275-E275
276	2	=MIN(E2:E255)	=MAX(E2:E255)	=C276-B276	=QUARTILE.INKL(E2:E255;1)	=QUARTILE.INKL(E2:E255;3)	=F276-E276
277	3	=MIN(F2:F255)	=MAX(F2:F255)	=C277-B277	=QUARTILE.INKL(F2:F255;1)	=QUARTILE.INKL(F2:F255;3)	=F277-E277
278	4	=MIN(G2:G255)	=MAX(G2:G255)	=C278-B278	=QUARTILE.INKL(G2:G255;1)	=QUARTILE.INKL(G2:G255;3)	=F278-E278
279	5	=MIN(H2:H255)	=MAX(H2:H255)	=C279-B279	=QUARTILE.INKL(H2:H255;1)	=QUARTILE.INKL(H2:H255;3)	=F279-E279
280	6	=MIN(I2:I255)	=MAX(I2:I255)	=C280-B280	=QUARTILE.INKL(I2:I255;1)	=QUARTILE.INKL(I2:I255;3)	=F280-E280
281	7	=MIN(J2:J255)	=MAX(J2:J255)	=C281-B281	=QUARTILE.INKL(J2:J255;1)	=QUARTILE.INKL(J2:J255;3)	=F281-E281
282	8	=MIN(K2:K255)	=MAX(K2:K255)	=C282-B282	=QUARTILE.INKL(K2:K255;1)	=QUARTILE.INKL(K2:K255;3)	=F282-E282

Wir erhalten als Ergebnis folgende Tabelle:

Anzahl Zimmer	Spannweite	Quartilsabstand
1	241	151
2	1198	172,25
3	2853	451
4	1926	596
5	3081	580
6	1227	495
7	1070	535
8	0	0

Interpretation: Vergleichbarkeit ist nicht gegeben, da die Häufigkeiten der Zimmeranzahlen sehr unterschiedlich sind und bei höheren Mietpreisen sowieso mit höherer Spannweite zu rechnen ist. Um das deutlich zu machen, führen wir noch diesen Schritt durch:

B262 | f_x | =ZÄHLENWENN(A2:A255;"=1")

	A	B
260		
261	Anzahl Zimmer	Häufigkeit
262	1	=ZÄHLENWENN(A2:A255;"=1")
263	2	=ZÄHLENWENN(A2:A255;"=2")
264	3	=ZÄHLENWENN(A2:A255;"=3")
265	4	=ZÄHLENWENN(A2:A255;"=4")
266	5	=ZÄHLENWENN(A2:A255;"=5")
267	6	=ZÄHLENWENN(A2:A255;"=6")
268	7	=ZÄHLENWENN(A2:A255;"=7")
269	8	=ZÄHLENWENN(A2:A255;"=8")
270		

Anzahl Zimmer	Häufigkeit
1	28
2	102
3	61
4	39
5	11
6	9
7	3
8	1

Die Funktion `ZÄHLEWENN` zählt alle Zellen, die im vorgegebenen Bereich die Bedingung erfüllen. Dabei steht in der Klammer zunächst der Zellbereich und dann die Bedingung.

Lösung für Übung ASM-08b

Quartilsabstand von Mieten

	Min	Max	1 Quartil	3 Quartil	Spannweite	Quartilsabstand
Mit Ausreißern	67	97	83,5	88,5	30	5
Ohne Ausreißer	82	89	85	88	7	3

Interpretation: Die Spannweite wird durch die Ausreißer stark beeinflusst. Da die Quartile im Vergleich dazu stabiler sind, verändert sich der Quartilsabstand nicht so stark.

Lösung mit R

mabweichung_loesung.R

```
# Einlesen der Daten aus der Datei kdaten.txt
mietwohnungen<-read.table("kdaten.txt",sep="\t",header=TRUE)

#  Aufspaltung des Data Frames nach Zimmeranzahl
mietwohnungen1<-subset(mietwohnungen,Zimmer==1)
mietwohnungen2<-subset(mietwohnungen,Zimmer==2)
mietwohnungen3<-subset(mietwohnungen,Zimmer==3)
mietwohnungen4<-subset(mietwohnungen,Zimmer==4)
mietwohnungen5<-subset(mietwohnungen,Zimmer==5)
mietwohnungen6<-subset(mietwohnungen,Zimmer==6)
mietwohnungen7<-subset(mietwohnungen,Zimmer==7)
mietwohnungen8<-subset(mietwohnungen,Zimmer==8)

#  Bestimmung des Quartilsabstands der Mieten nach Zimmeranzahl
quantile(mietwohnungen1$Preis,0.75)-quantile(mietwohnungen1$Preis,0.25)
quantile(mietwohnungen2$Preis,0.75)-quantile(mietwohnungen2$Preis,0.25)
quantile(mietwohnungen3$Preis,0.75)-quantile(mietwohnungen3$Preis,0.25)
quantile(mietwohnungen4$Preis,0.75)-quantile(mietwohnungen4$Preis,0.25)
quantile(mietwohnungen5$Preis,0.75)-quantile(mietwohnungen5$Preis,0.25)
quantile(mietwohnungen6$Preis,0.75)-quantile(mietwohnungen6$Preis,0.25)
quantile(mietwohnungen7$Preis,0.75)-quantile(mietwohnungen7$Preis,0.25)
quantile(mietwohnungen8$Preis,0.75)-quantile(mietwohnungen8$Preis,0.25)
```

Lösung mit Excel

WMS_ASM_08_SpannweiteQuartilsabstand.xlsx (9 KB)

***Aufgabe 1**: Bestimmen Sie für jede Zimmeranzahl den Quartilsabstand der Mieten und interpretieren Sie das Ergebnis.*

Wie in der Übung ASM-08a "Spannweite von Mieten" berechnen wir zunächst die Werte Minimum, Maximum, 1. Quartil und 3. Quartil. Dafür helfen uns wieder die Funktionen `MIN`, `MAX` und `QUARTILE.INKL`.

Die Spannweite ergibt sich wieder aus der Differenz zwischen Maximum und Minimum, der Quartilsabstand aus der Differenz zwischen 3. Quartil und 1. Quartil.

	A	B	C	D	E	F
10						
11	Aufgabe					
12						
13	Mit Ausreißern:					
14						
15	Minimum	Maximum	1. Quartil	3. Quartil	Spannweite	Quartilsabstand
16	=MIN(B2:B8)	=MAX(B2:B8)	=QUARTILE.INKL(B2:B8;1)	=QUARTILE.INKL(B2:B8;3)	=B16-A16	=D16-C16
17						
18						
19	Ohne Ausreißer:					
20						
21	Minimum	Maximum	1. Quartil	3. Quartil	Spannweite	Quartilsabstand
22	=MIN(B3:B7)	=MAX(B3:B7)	=QUARTILE.INKL(B3:B7;1)	=QUARTILE.INKL(B3:B7;3)	=B22-A22	=D22-C22

Oben mit Ausreißern, unten ohne die Ausreißer 67 und 97.

Als Ergebnis erhalten wir:

Mit Ausreißern:					
Minimum	Maximum	1. Quartil	3. Quartil	Spannweite	Quartilsabstand
67	97	83,5	88,5	30	5
Ohne Ausreißer:					
Minimum	Maximum	1. Quartil	3. Quartil	Spannweite	Quartilsabstand
82	89	85	88	7	3

Interpretation: Die Spannweite wird durch die Ausreißer stark beeinflusst. Da die Quartile im Vergleich dazu stabiler sind, verändert sich der Quartilsabstand nicht so stark.

Lösung für Übung ASM-08c

Absolute Abweichung vom Median

Der Median liegt bei 87. Die mittlere Abweichung vom Median ist 8.

Lösung mit R

mmedian_loesung.R

```
#  Einlesen der Punktzahlen in den Vektor "punkte"
punkte<-c(87,78,97,67,91,92,79)

#  Berechnung und Ausgabe des Medians
med<-median(punkte)
med

#  Berechnung der mittleren absoluten Abweichungen vom Median
abs.mit.abw<-mean(abs(punkte-med))
abs.mit.abw
```

Lösung mit Excel

WMS_ASM_08_AbsAbwMedian.xlsx (9 KB)

***Aufgabe 1**: Gegeben seien die sieben Messwerte 87, 78, 97, 67, 91, 92, 79 (Punktzahlen, die eine Studentin bei Klausuren erreichte). Bestimmen Sie die mittlere absolute Abweichung vom Median..*

Zunächst wird mit der Funktion `MEDIAN` der Median der Messwerte berechnet. Dann können wir für jeden Messwert die Differenz zum Median bestimmen. Dazu verwenden wir die Funktion `ABS`, die den Betrag der Differenzen zurück gibt.

A14 | f_x =MEDIAN(B2:B8)

	A	B	C	D	E
10					
11	Aufgabe				
12					
13	Median berechnen:				
14	87				
15					

B17 | f_x =ABS(B2-A14)

	A	B
16	Messung	Abweichung vom Mediar
17	1	=ABS(B2-A14)
18	2	=ABS(B3-A14)
19	3	=ABS(B4-A14)
20	4	=ABS(B5-A14)
21	5	=ABS(B6-A14)
22	6	=ABS(B7-A14)
23	7	=ABS(B8-A14)
24		

Die mittlere Abweichung vom Median berechnen wir dann schließlich mit der Funktion `MITTELWERT`, indem wir den Mittelwert der einzelnen Abweichungsbeträge bestimmen.

SUMME | =MITTELWERT(B17:B23)

	A	B	C	D	E	F
16	Messung	Abweichung vom Median				
17	1	0				
18	2	9				
19	3	10				
20	4	20				
21	5	4				
22	6	5				
23	7	8				
24						
25	Mittlere Abweichung vom Median:					
26		=MITTELWERT(B17:B23)				

Das Ergebnis ist dann:

Mittlere Abweichung vom Median:		
	8	

Lösung für Übung ASM-08d

Klausurnoten - Interquartilsrange

	Spannweite	Quartilsabstand
Mit Ausreißern	30	5
Ohne Ausreißer	7	3

Das Streichen der Ausreißer wirkt sich besonders auf die Spannweite aus, weil die Ausreißer ja genau das Minimum und das Maximum der erreichten Punkte sind, also genau die Werte, die für die Berechnung der Spannweite genutzt werden.

Lösung mit R

quantile_loesung.R

```
#  Spannweite und Quartilsabstand bei Klausurenpunkten - Aufgabe

#  Ein Student erreicht in seinen Klausuren folgende Punktzahlen:
#  67, 82, 85, 87, 88, 89, 97. Berechnen Sie die Spannweite und den
#  Quartilsabstand bevor und nachdem Sie die Ausreißer 67 und 97
#  gestrichen haben. Interpretieren Sie die Ergebnisse.

##########

#  Spannweite und Quartilsabstand bei Klausurenpunkten - Lösung

#  Erstellung der Vektoren mit und ohne die Ausreißer
punkte.mit<-c(67,82,85,87,88,89,97)
punkte.ohne<-c(82,85,87,88,89)

#  Bestimmung der Spannweite
max(punkte.mit)-min(punkte.mit)
max(punkte.ohne)-min(punkte.ohne)

#  Bestimmung des Quartilsabstandes
quantile(punkte.mit,0.75)-quantile(punkte.mit,0.25)
quantile(punkte.ohne,0.75)-quantile(punkte.ohne,0.25)

#  Interpretation:
#  Das Streichen der Ausreißer wirkt sich besonders auf die Spannweite
#  aus, weil die Ausreißer ja genau das Minimum und das Maximum der
#  erreichten Punkte sind, also genau die Werte, die für die Berechnung
#  der Spannweite genutzt werden.
```

Lösung für Übung ASM-08f

Spannweite bei Klausurergebnissen

1. Beide haben dieselbe Spannweite von 30
2. Interpretation:
   Da die höchste und die niedrigste Punktzahl beider Studenten die gleiche ist, haben beide Studenten die gleiche Spannweite, obwohl Sie zum größten Teil unterschiedliche Punkte erreicht haben.

Lösung mit R

zstudenten_loesung.R

```r
#  Einlesen der Punktzahlen in Vektoren
punkte1<-c(84,73,97,67,88,93,79)
punkte2<-c(82,67,85,97,87,89,88)

#  Berechnung der Spannweiten
max(punkte1)-min(punkte1)
max(punkte2)-min(punkte2)
```

Lösung für Übung ASM-09

Spannweite

Die Spannweite beträgt 181 - 168 = 13.

Lösung für Übung ASM-10

Spannweite: Studierendenalter

Die Spannweite beträgt 28 - 20 = 8

Lösung für Übung ASM-11

Spannweite: Dummy-Zahlen

Die Spannweite beträgt 7:6 - 5:6 = 2

Lösung für Übung ASM-12

Spannweite - Tagestemperaturen

Zunächst werden die Tagestemperaturen der Größe nach angeordnet:

Juni: 14, 14.5, 17, 18, 19.5, 23, 24, 24.5, 25, 25, 26, 27
Juli: 18, 19.5, 23, 23, 25, 25, 25, 26, 26.5, 27, 28, 28.5

Die Spannweite für Monat Juni: 27 - 14 = 13

Die Spannweite für Monat Juli: 28,5 - 18 = 10,5

Lösung für Übung ASM-13

Master und Bachelor: Gehälter im Vergleich

Lösung: Mittelwert und Median

Bachelor

Für das arithmetische Mittel im Einkommen in den fünf verschiedenen Fachrichtungen ergibt sich:

- für **BWL** ein Betrag (in €) von:
  $\overline{x} = \frac{1}{5}(3100 + 2900 + 2999 + 2780 + 2850) = 2925.80$
- für **Naturwissenschaften** ein Betrag (in €) von:
  $\overline{x} = \frac{1}{5}(2600 + 2450 + 2700 + 2650 + 2900) = 2660$
- für **Jura** ein Betrag (in €) von:

$\overline{x} = \frac{1}{5}(3050 + 3400 + 2800 + 2950 + 3100) = 3060$

- für **Ingenieure/Ingenieurinnen** ein Betrag (in €) von:
  $\overline{x} = \frac{1}{5}(3100 + 3400 + 3350 + 3500 + 3250) = 3320$
- für **Geisteswissenschaften** ein Betrag (in €) von:
  $\overline{x} = \frac{1}{5}(2100 + 2320 + 2400 + 2150 + 2330) = 2260$

Interpretation:

Das Durchschnittseinkommen der befragten Personen in den jeweiligen Fachrichtungen beträgt:

2925,80 €	2660 €	3060 €	3320 €	2260 €

Lösung: Median

- für **BWL** ein Betrag (in €) von:
  $\tilde{x}_{0,5} = x_{\left(\frac{(n+1)}{2}\right)} = x_3 = 2900$
- für **Naturwissenschaften** ein Betrag (in €) von:
  $\tilde{x}_{0,5} = x_{\left(\frac{(n+1)}{2}\right)} = x_3 = 2650$
- für **Jura** ein Betrag (in €) von:
  $\tilde{x}_{0,5} = x_{\left(\frac{(n+1)}{2}\right)} = x_3 = 3050$
- für **Ingenieure/Ingenieurinnen** ein Betrag (in €) von:
  $\tilde{x}_{0,5} = x_{\left(\frac{(n+1)}{2}\right)} = x_3 = 3350$
- für **Geisteswissenschaften** ein Betrag (in €) von:
  $\tilde{x}_{0,5} = x_{\left(\frac{(n+1)}{2}\right)} = x_3 = 2320$

Interpretation:

50 % der Einkommen in den jeweiligen Fachrichtungen betragen höchstens:

2900 €	2650€	3050 €	3350 €	2320 €

Lösung Mittelwert und Median

Master

Für das arithmetische Mittel im Einkommen in den fünf verschiedenen Fachrichtungen ergibt sich:

- für **BWL** ein Betrag (in €) von:
  $\overline{x} = \frac{1}{5}(3300 + 3550 + 3680 + 3450 + 3700) = 3536$
- für **Naturwissenschaften** ein Betrag (in €) von:
  $\overline{x} = \frac{1}{5}(3200 + 3520 + 3480 + 3720 + 3490) = 3482$
- für **Jura** ein Betrag (in €) von:
  $\overline{x} = \frac{1}{5}(3900 + 3780 + 3400 + 3850 + 3650) = 3716$
- für **Ingenieure/Ingenieurinnen** ein Betrag (in €) von:
  $\overline{x} = \frac{1}{5}(3800 + 4100 + 3780 + 3890 + 4300) = 3974$

- für **Geisteswissenschaften** ein Betrag (in €) von:
  $\overline{x} = \frac{1}{5}(3100 + 2750 + 2650 + 2900 + 3050) = 2890$

Das Durchschnittseinkommen der befragten Personen in den jeweiligen Fachrichtungen beträgt:

3536 €	3482 €	3716 €	3974 €	2890 €

Median

- für **BWL** ein Betrag (in €) von:
  $\tilde{x}_{0,5} = x_{\left(\frac{(n+1)}{2}\right)} = x_3 = 3550$
- für **Naturwissenschaften** ein Betrag (in €) von:
  $\tilde{x}_{0,5} = x_{\left(\frac{(n+1)}{2}\right)} = x_3 = 3490$
- für **Jura** ein Betrag (in €) von:
  $\tilde{x}_{0,5} = x_{\left(\frac{(n+1)}{2}\right)} = x_3 = 3780$
- für **Ingenieure/Ingenieurinnen** ein Betrag (in €) von:
  $\tilde{x}_{0,5} = x_{\left(\frac{(n+1)}{2}\right)} = x_3 = 3890$
- für **Geisteswissenschaften** ein Betrag (in €) von:
  $\tilde{x}_{0,5} = x_{\left(\frac{(n+1)}{2}\right)} = x_3 = 2900$

Interpretation:

50 % der Einkommen in den jeweiligen Fachrichtungen betragen höchstens:

3550 €	3490 €	3780 €	3890 €	2900 €

Lösung Spannweite

Bachelor

Um die Spannweite bestimmen zu können, müssen die Einkommenswerte der fünf Fachrichtungen zunächst der Größe nach geordnet werden. Die Spannweite beträgt jeweils:

- für **BWL** ein Betrag (in €) von:
  $x_{max} - x_{min} = 3100 - 2780 = 320$
- für **Naturwissenschaften** ein Betrag (in €) von:
  $x_{max} - x_{min} = 2900 - 2450 = 450$
- für **Jura** ein Betrag (in €) von:
  $x_{max} - x_{min} = 3400 - 2800 = 600$
- für **Ingenieure/Ingenieurinnen** ein Betrag (in €) von:
  $x_{max} - x_{min} = 3500 - 3100 = 400$
- für **Geisteswissenschaften** ein Betrag (in €) von:
  $x_{max} - x_{min} = 2400 - 2100 = 300$

Interpretation

Die Spannweite oder die Breite eines Streubereichs für das Einkommen in den jeweiligen Fachrichtungen beträgt:

320 €	450 €	600 €	400 €	300 €

Lösung Spannweite

Master

Um die Spannweite bestimmen zu können, müssen die Einkommenswerte der fünf Fachrichtungen zunächst der Größe nach geordnet werden. Die Spannweite beträgt jeweils:

- für **BWL** ein Betrag (in €) von:
  $x_{max} - x_{min} = 3700 - 3300 = 400$
- für **Naturwissenschaften** ein Betrag (in €) von:
  $x_{max} - x_{min} = 3720 - 3200 = 520$
- für **Jura** ein Betrag (in €) von:
  $x_{max} - x_{min} = 3900 - 3400 = 500$
- für **Ingenieure/Ingenieurinnen** ein Betrag (in €) von:
  $x_{max} - x_{min} = 4300 - 3780 = 520$
- für **Geisteswissenschaften** ein Betrag (in €) von:
  $x_{max} - x_{min} = 3100 - 2650 = 450$

Interpretation

Die Spannweite oder die Breite eines Streubereichs für das Einkommen in den jeweiligen Fachrichtungen beträgt:

400 €	520 €	500 €	520 €	450 €

Lösung IQR

Bachelor

Um den Quartilsabstand bestimmen zu können, müssen die Einkommen in den 5 Fachrichtungen zunächst der Größe nach geordnet werden und können dann mit der Formel IQR = Q_3 - Q_1 berechnet werden.

- für **BWL** ein Betrag (in €) von:
  IQR = Q_3 - Q_1 = 2999 - 2850 = 149
- für **Naturwissenschaften** ein Betrag (in €) von:
  IQR =Q_3 - Q_1 = 2700 - 2600 = 100
- für **Jura** ein Betrag (in €) von:
  IQR = Q_3 - Q_1 = 3100 - 2950 = 150
- für **Ingenieure/Ingenieurinnen** ein Betrag (in €) von:
  IQR = Q_3 - Q_1 = 3400 - 3250 = 150
- für **Geisteswissenschaften** ein Betrag (in €) von:
  IQR = Q_3 - Q_1 = 2330 - 2150 = 180

Interpretation

Zwischen dem ersten Quartil Q1 und dem dritten Quartil Q3 liegen 50 % aller Einkommenswerte in den jeweiligen Fachrichtungen. Der Quartilsabstand liefert somit die Größe des Bereichs, in dem etwa die Hälfte aller Einkommenswerte in den jeweiligen Fachrichtungen liegen.
Beispielsweise liegen 50 % aller Einkommen in der
Fachrichtung BWL (Bachelor) zwischen 2.999 € und 2.850 € usw.

Lösung IQR

Master

Um den Quartilsabstand bestimmen zu können, müssen die Einkommen in den 5 Fachrichtungen zunächst der Größe nach geordnet werden und können dann mit der Formel IQR = Q_3 - Q_1 berechnet werden.

- für **BWL** ein Betrag (in €) von:
  IQR = Q_3 - Q_1 = 3680 - 3450 = 230

- für **Naturwissenschaften** ein Betrag (in €) von:
  IQR =Q_3 - Q_1 = 3520 - 3480 = 40

- für **Jura** ein Betrag (in €) von:
  IQR = Q_3 - Q_1 = 3850 - 3650 = 200

- für **Ingenieure/Ingenieurinnen** ein Betrag (in €) von:
  IQR = Q_3 - Q_1 = 4100 - 3800 = 300

- für **Geisteswissenschaften** ein Betrag (in €) von:
  IQR = Q_3 - Q_1 = 3100 - 2750 = 350

Interpretation

Zwischen dem ersten Quartil Q1 und dem dritten Quartil Q3 liegen 50 % aller Einkommenswerte in den jeweiligen Fachrichtungen. Der Quartilsabstand liefert somit die Größe des Bereichs, in dem etwa die Hälfte aller Einkommenswerte in den jeweiligen Fachrichtungen liegen.
Beispielsweise liegen 50 % aller Einkommen in der
Fachrichtung BWL (Master) zwischen 3.680 € und 3.450 € usw.

Lösung MA Median

Bachelor

Die mittlere absolute Abweichung vom Median $\text{MA}_{\tilde{x}_{0,5}}$ in der Fachrichtung

- für **BWL** ein Betrag (in €) von:
  $\text{MA}_{\tilde{x}_{0,5}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0,5}| \right) = \frac{469}{5} = 93.8$

- für **Naturwissenschaften** ein Betrag (in €) von:
  $\text{MA}_{\tilde{x}_{0,5}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0,5}| \right) = \frac{550}{5} = 110$

- für **Jura** ein Betrag (in €) von:
  $\text{MA}_{\tilde{x}_{0,5}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0,5}| \right) = \frac{750}{5} = 150$

- für **Ingenieure/Ingenieurinnen** ein Betrag (in €) von:
  $\text{MA}_{\tilde{x}_{0,5}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0,5}| \right) = \frac{550}{5} = 110$

- für **Geisteswissenschaften** ein Betrag (in €) von:

$\text{MA}_{\tilde{x}_{0,5}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0,5}| \right) = \frac{480}{5} = 96$

Master

Die mittlere absolute Abweichung vom Median $\text{MA}_{\tilde{x}_{0,5}}$ in der Fachrichtung

- für **BWL** ein Betrag (in €) von:
  $\text{MA}_{\tilde{x}_{0,5}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0,5}| \right) = \frac{630}{5} = 126$
- für **Naturwissenschaften** ein Betrag (in €) von:
  $\text{MA}_{\tilde{x}_{0,5}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0,5}| \right) = \frac{560}{5} = 112$
- für **Jura** ein Betrag (in €) von:
  $\text{MA}_{\tilde{x}_{0,5}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0,5}| \right) = \frac{700}{5} = 140$
- für **Ingenieure/Ingenieurinnen** ein Betrag (in €) von:
  $\text{MA}_{\tilde{x}_{0,5}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0,5}| \right) = \frac{820}{5} = 164$
- für **Geisteswissenschaften** ein Betrag (in €) von:
  $\text{MA}_{\tilde{x}_{0,5}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0,5}| \right) = \frac{750}{5} = 150$

Interpretation

Die mittlere absolute Abweichung vom Median $\text{MA}_{\tilde{x}_{0,5}}$ ist ein Durchschnittswert, der aus den Abständen der einzelnen Einkommen vom Median berechnet wird.

Lösung MA Mittelwert

Bachelor

Die mittlere absolute Abweichung vom Mittelwert $\text{MA}_{\overline{x}}$ in der Fachrichtung

- für **BWL** ein Betrag (in €) von:
  $\text{MA}_{\overline{x}} = \frac{1}{n} \sum \left(|x_i - \overline{x}| \right) = \frac{494{,}8}{5} = 98{,}96$
- für **Naturwissenschaften** ein Betrag (in €) von:
  $\text{MA}_{\overline{x}} = \frac{1}{n} \sum \left(|x_i - \overline{x}| \right) = \frac{494{,}8}{5} = 98{,}96$
- für **Jura** ein Betrag (in €) von:
  $\text{MA}_{\overline{x}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0,5}| \right) = \frac{560}{5} = 112$
- für **Ingenieure/Ingenieurinnen** ein Betrag (in €) von:
  $\text{MA}_{\overline{x}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0,5}| \right) = \frac{580}{5} = 116$
- für **Geisteswissenschaften** ein Betrag (in €) von:
  $\text{MA}_{\overline{x}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0,5}| \right) = \frac{540}{5} = 108$

Master

Die mittlere absolute Abweichung vom Mittelwert MA$_{\overline{x}}$ in der Fachrichtung

- für **BWL** ein Betrag (in €) von:
  MA$_{\overline{x}} = \frac{1}{n} \sum \left(|x_i - \overline{x}| \right) = \frac{644}{5} = 128{,}8$
- für **Naturwissenschaften** ein Betrag (in €) von:
  MA$_{\overline{x}} = \frac{1}{n} \sum \left(|x_i - \overline{x}| \right) = \frac{568}{5} = 113{,}6$
- für **Jura** ein Betrag (in €) von:
  MA$_{\overline{x}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0{,}5}| \right) = \frac{765}{5} = 153$
- für **Ingenieure/Ingenieurinnen** ein Betrag (in €) von:
  MA$_{\overline{x}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0{,}5}| \right) = \frac{904}{5} = 180{,}8$
- für **Geisteswissenschaften** ein Betrag (in €) von:
  MA$_{\overline{x}} = \frac{1}{n} \sum \left(|x_i - \tilde{x}_{0{,}5}| \right) = \frac{860}{5} = 172$