

## Lagemaße in der beschreibenden Statistik.

Zur Interpretation einer Beobachtungsreihe kann man neben der grafischen Darstellung weitere charakteristische Größen heranziehen.

### Mittelwert und Median

#### Arithmetisches Mittel einer Datenreihe

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

$n$ : Anzahl der Beobachtungswerte  $x_i$        $x_i$ :  $i$ -ter Beobachtungswert  $i \in \mathbb{N}$

Beispiel:

Bestimmen Sie aus der Liste einer Schülerbefragung die durchschnittliche Körpergröße aller befragten Schüler.

Körpergröße

$x_i$ : 164 ; 162 ; 168 ; 190 ; 166 ; 166 ; 162 ; 160 ; 150 ; 175 ; 180 ; 160 ; 168 ; 175 ; 168

Anzahl der Beobachtungswerte (Schüler)  $n = 15$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{15} \sum_{i=1}^{15} x_i$$

$$= \frac{1}{15} (164 + 162 + 168 + 190 + 166 + 166 + 162 + 160 + 150 + 175 + 180 + 160 + 168 + 175 + 168)$$

$$= \frac{2514}{15} = \underline{\underline{167,6}} \text{ (Durchschnittsgröße)}$$

Weitere Beispiele für Mittelwerte:

Durchschnittsabiturnote: 1,8

Durchschnittsgewicht aller Schüler einer Klasse: 62,3 kg

#### Der Median (Zentralwert einer Datenreihe)

Der Median  $x_{\text{Med}}$  ist derjenige Wert (Merkmalsausprägung), der in der Mitte steht, wenn alle Beobachtungswerte  $x_i$  der Größe nach geordnet sind.

Wir ordnen alle Werte aus unserem Beispiel der Größe nach und bestimmen die Mitte.

Körpergröße geordnet

$x_i$ : 150 ; 160 ; 160 ; 162 ; 162 ; 164 ; 166 ; **166** ; 168 ; 168 ; 168 ; 175 ; 175 ; 180 ; 190

Der 8. Wert, also  $x_8 = 166$  bildet den Median.

Wie verändern sich Mittelwert und Median, wenn der größte Schüler die Klasse verlässt und für ihn eine kleine Schülerin mit der Körpergröße 150 dazu kommt?  
Körpergröße geordnet

$x_i$ : 150 : 150 ; 160 ; 160 ; 162 ; 162 ; 164 ; **166** ; 166 ; 168 ; 168 ; 175 ; 175 ; 180

Der Median, also  $x_8 = 166$  bleibt unverändert.

$$\text{Mittelwert: } \bar{x} = \frac{1}{15} \sum_{i=1}^{15} x_i$$

$$= \frac{1}{15} (164 + 162 + 168 + 150 + 166 + 166 + 162 + 160 + 150 + 175 + 180 + 160 + 168 + 175 + 168)$$

$$= \frac{2474}{15} = \underline{\underline{164,9}}$$

Der Mittelwert wird nach unten gezogen.

Wie verändert sich der Median, wenn ein Schüler mit der Körpergröße 180 dazu kommt?

Körpergröße geordnet

$x_i$ : 150 : 150 ; 160 ; 160 ; 162 ; 162 ; 164 ; **166** ; **166** ; 168 ; 168 ; 168 ; 175 ; 175 ; 180 ; 180

Der Median liegt zwischen den Werten von  $x_8$  und  $x_9$ , bleibt also bei 166

Allgemeine Rechenvorschrift zur Berechnung des Medians

Ist  $n$  ist die Anzahl der Beobachtungswerte  $x_i$ , dann gilt:

$$n \text{ ungerade} \Rightarrow x_{\text{Med}} = x_{\frac{n+1}{2}} \quad n \text{ gerade} \Rightarrow x_{\text{Med}} = \frac{1}{2} \left( x_{\frac{n}{2}} + x_{\frac{n}{2}+1} \right)$$

Vorbetrachtungen zur Varianz

Wir betrachten wieder unser Anfangsbeispiel mit dem Mittelwert 167,6 und bilden die Summe der Abweichungen von diesem.

$x_i$	$\bar{x}$	$x_i - \bar{x}$
164	167,6	-3,6
162	167,6	-5,6
168	167,6	0,4
190	167,6	22,4
166	167,6	-1,6
166	167,6	-1,6
162	167,6	-5,6
160	167,6	-7,6
150	167,6	-17,6
175	167,6	7,4
180	167,6	12,4
160	167,6	-7,6
168	167,6	0,4
175	167,6	7,4
168	167,6	0,4
	$\Sigma$	0,0

Die Summe bestätigt nur den Mittelwert, sie hat keine Aussagekraft für die Streuung.

Die positiven und negativen Differenzen heben sich auf.

Um die negativen Differenzen zu vermeiden, berechnen wir die Quadrate der Differenzen und bilden davon den Mittelwert.

$x_i$	$\bar{x}$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
164	167,6	-3,6	12,96
162	167,6	-5,6	31,36
168	167,6	0,4	0,16
190	167,6	22,4	501,76
166	167,6	-1,6	2,56
166	167,6	-1,6	2,56
162	167,6	-5,6	31,36
160	167,6	-7,6	57,76
150	167,6	-17,6	309,76
175	167,6	7,4	54,76
180	167,6	12,4	153,76
160	167,6	-7,6	57,76
168	167,6	0,4	0,16
175	167,6	7,4	54,76
168	167,6	0,4	0,16
	$\Sigma$	0,0	1271,60

Mittelwertbildung:

$$\bar{x} = \frac{1}{15} \sum_{i=1}^{15} (x_i - \bar{x})^2 = \frac{1}{15} \cdot 1271,60 = 84,77$$

Dieser Wert wird Varianz  $s^2$  genannt.

Er ist der Mittelwert der quadratischen Abweichungen vom Mittelwert.

Die Wurzel daraus nennen wir Standardabweichung  $s$

$$s = \sqrt{s^2} = \sqrt{84,77} \approx 9,21$$

Die Standardabweichung ist ein Maß für die Streuung um den Mittelwert.

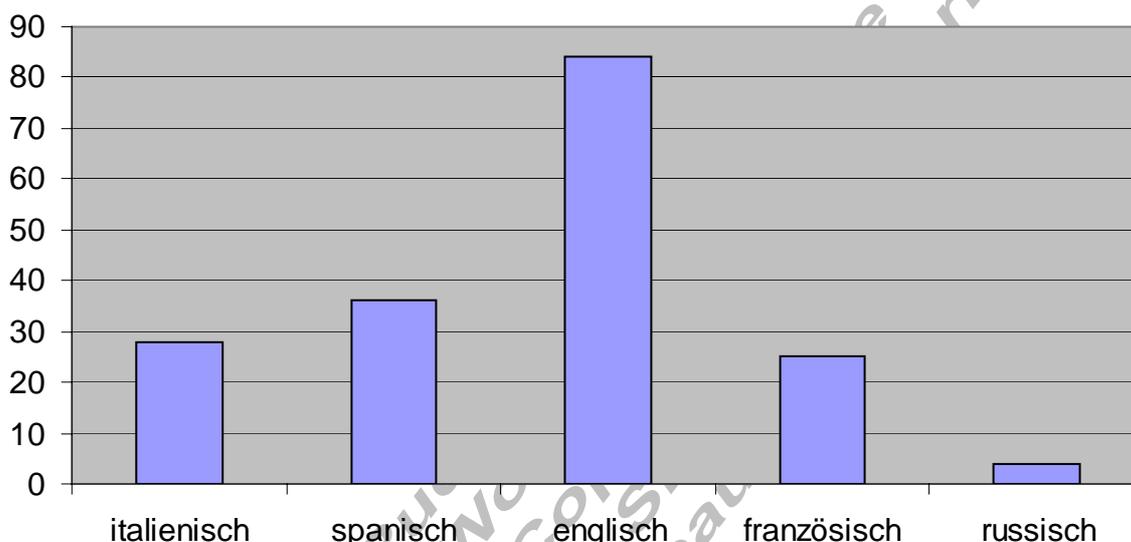
## Modalwert (Modus)

Bei Merkmalsausprägungen wie z.B. „rot, blau, grün“, also bei nominal skalierten Größen kann kein arithmetisches Mittel berechnet werden.

Hier lässt sich lediglich die Frage nach der Merkmalsausprägung mit der größten Häufigkeit stellen.

Beispiel:

Anzahl der Schüler am Fremdsprachenunterricht



Die Fremdsprache englisch kommt mit der größten Häufigkeit vor (84 mal)  
Somit ist der Modalwert  $x_{\text{Mod}} = \text{englisch}$ .

Der Modalwert  $x_{\text{Mod}}$  ist der Merkmalswert, der am häufigsten vorkommt.

Bemerkung zum Modalwert:

Gibt es mehrere Merkmalsausprägungen mit der gleichen maximalen Häufigkeit, so existiert kein Modalwert.

Bei einer Klasseneinteilung ist der Modalwert die Mitte der am dichtesten besetzten Klasse.

Die Verwendung des Modus ist bei jedem Skalenniveau möglich.

## Ergänzungen zum Median

Beispiel:

Ein Baurupp mit 9 Personen hat folgende monatliche Einkünfte in Euro.

1160 | 1050 | 980 | 1200 | 970 | 1800 | 6600 | 1180 | 1090

arithmetisches Mittel:  $\bar{x} \approx 1781,11 \text{ €}$

Dieser Durchschnitt liefert ein falsches Bild, weil die Mehrzahl (7 von 9 Personen) höchstens 1200 € verdient.

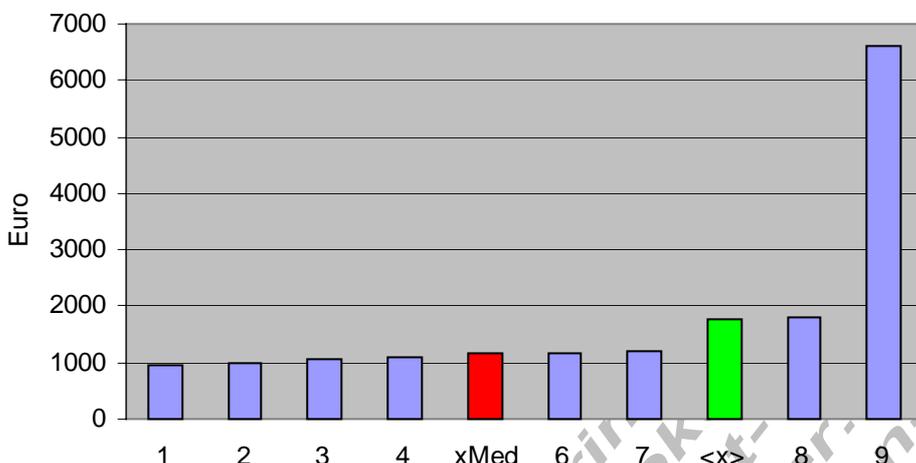
Der Wert 6600 zieht den Mittelwert nach oben.

Wir berechnen den Median

970 | 980 | 1050 | 1090 | 1160 | 1180 | 1200 | 1800 | 6600 n = 9 ungerade

$$\Rightarrow x_{\text{Med}} = x_{\frac{n+1}{2}} = x_{\frac{9+1}{2}} = x_5 = 1160$$

Der Median beschreibt die Verteilung besser als der Mittelwert. Man nennt ihn auch Zentralwert.



Ausreißer haben auf den Median keinen Einfluss.

Beispiel :

Die Anzahl n der Merkmalsausprägungen ist ungerade, z.B. das Alter von 7 Mathematiklehrern ( n = 7 )

28	31	40	45	52	53	62
$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$

$$x_{\text{Med}} = x_{\frac{n+1}{2}} = x_4 = 45$$

3 Werte	$x_{\text{Med}}$	3 Werte
---------	------------------	---------

allgemein:  $x_{\text{Med}} = x_{\frac{n+1}{2}}$

In der Tabelle stehen links und rechts neben dem Median gleich viele Werte.

Beispiel 2:

Die Anzahl der Merkmalsausprägungen ist gerade, z.B. das Alter von 8 Mathematiklehrern ( n = 8 )

28	31	40	45	52	53	58	62
$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$

$$x_{\text{Med}} = \frac{x_4 + x_5}{2} = \frac{45 + 52}{2} = 48,5$$

3 Werte	$x_{\text{Med}}$	3 Werte
---------	------------------	---------

allgemein:  $x_{\text{Med}} = \frac{1}{2} \left( x_{\frac{n}{2}} + x_{\frac{n}{2}+1} \right)$

Bei einer geraden Anzahl von Werten ( n = 8 ) berechnet man den Median aus den beiden mittleren Werten.

**Bemerkungen zum Median.**

Falls das betrachtete Merkmal nur ordinal skaliert ist (z.B. Zeugnisnoten), so ist bei geradem  $n$  zu beachten dass der Median nur dann existiert, wenn beide infrage kommenden Merkmalsausprägungen gleich sind.

Z.B. bei den Zeugnisnoten 1 2 **3 4** 5 6 existiert kein Median, denn 3,5 als Zeugnisnote ist nicht üblich.

Aber: 1 2 **3 3** 4 5 hat den Median 3.

Für den Fall, dass metrische Daten in Klassen gruppiert vorliegen, kann die exakte Merkmalsausprägung des Medians nicht bestimmt werden.

**Berechnung des arithmetischen Mittels aus einer Häufigkeitstabelle**

Fall I: Absolute Häufigkeit  $n_i$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^j x_i \cdot n_i = \frac{1}{n} (x_1 \cdot n_1 + x_2 \cdot n_2 + \dots + x_j \cdot n_j) \quad n = \sum_{i=1}^j n_i = n_1 + n_2 + \dots + n_j$$

Fall II: Relative Häufigkeit  $h_i = \frac{n_i}{n}$

$$\bar{x} = \sum_{i=1}^j x_i \cdot h_i = (x_1 \cdot h_1 + x_2 \cdot h_2 + \dots + x_j \cdot h_j)$$

$n_i$  : absolute Häufigkeit der Merkmalsausprägung  $x_i$

$n$  : Summe der absoluten Häufigkeiten

$h_i$  : relative Häufigkeit der Merkmalsausprägung  $x_i$

$j$  : Anzahl der Merkmalsausprägungen  $x_i$

Beispiel:

Das Ergebnis einer Vergleichsarbeit ist untenstehender Tabelle zu entnehmen.

Berechnen Sie den Notendurchschnitt.

Häufigkeitstabelle

Note ( $x_i$ )	1	2	3	4	5	6
Anz. d. Schüler ( $n_i$ )	5	8	14	16	5	2

Schüler insgesamt:  $n = \sum_{i=1}^6 n_i = 50$

Durchschnittsnote:  $\bar{x} = \frac{1}{50} \sum_{i=1}^6 x_i \cdot n_i = \frac{1 \cdot 5 + 2 \cdot 8 + 3 \cdot 14 + 4 \cdot 16 + 5 \cdot 5 + 6 \cdot 2}{50} = \frac{164}{50} = \underline{\underline{3,28}}$

## Berechnung des arithmetischen Mittels bei klassierten Daten

Fall I: Absolute Häufigkeit  $n_i$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k m_i \cdot n_i = \frac{1}{n} (m_1 \cdot n_1 + m_2 \cdot n_2 + \dots + m_k \cdot n_k) \quad n = \sum_{i=1}^k n_i = n_1 + n_2 + \dots + n_k$$

Fall II: Relative Häufigkeit  $h_i = \frac{n_i}{n}$

$$\bar{x} = \sum_{i=1}^k m_i \cdot h_i = m_1 \cdot h_1 + m_2 \cdot h_2 + \dots + m_k \cdot h_k$$

$n_i$  : absolute Häufigkeit der  $i$ -ten Klasse

$n$  : Summe der absoluten Häufigkeiten

$h_i$  : relative Häufigkeit der  $i$ -ten Klasse

$k$  : Anzahl der Klassen

$m_i$  : Klassenmitte der  $i$ -ten Klasse

Beispiel:

Bestimmen Sie aus der klassierten Häufigkeitstabelle für das Körpergewicht den arithmetischen Mittelwert.

Klasse $x_i$	51 bis 60	61 bis 70	71 bis 80	81 bis 90
Häufigkeit $n_i$	11	10	5	1
Klassenmitte $m_i$	55,5	65,5	75,5	85,5

$$\text{Klassenmitte} = \frac{\text{Klassenanfang} + \text{Klassenende}}{2} \quad \text{z.B.} \quad \frac{61+70}{2} = 65,5$$

Der Häufigkeit wird die Klassenmitte zugeordnet.

Man geht davon aus, dass alle 10 Schüler der Klasse  $x_2$  das Körpergewicht 65,5 kg haben.

$$\text{Anzahl der Schüler: } n = \sum_{i=1}^4 x_i = 11 + 10 + 5 + 1 = 27$$

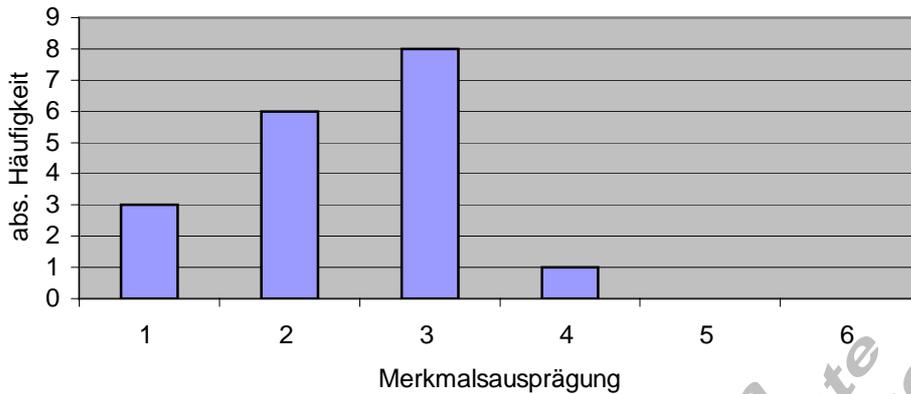
$$\begin{aligned} \bar{x} &= \frac{1}{27} \sum_{i=1}^4 m_i \cdot n_i \\ &= \frac{1}{27} (55,5 \cdot 11 + 65,5 \cdot 10 + 75,5 \cdot 5 + 85,5 \cdot 1) = \frac{1728,5}{27} \approx \underline{\underline{64,0}} \quad (\text{Durchschnittsgewicht}) \end{aligned}$$

Zusammenfassung der Eigenschaft von Lagemaßen.

Lagemaß	Skala	Definition	Beispiel: 23;19;17;44;19;44;44
Mittelwert $\bar{x}$	metrische	$\frac{1}{n} \sum_{i=1}^n x_i$	$\bar{x} = \frac{23 + 19 + 17 + 44 + 19 + 44 + 44}{7} = 30$
Modalwert $x_{\text{Mod}}$	alle	häufigste Merkmalsausprägung	$x_{\text{Mod}} = 44$ kommt 3 mal vor
Median $x_{\text{Med}}$	ordinal/metrisch	Wert in der Mitte	17;19;19; <b>23</b> ;44;44;44 $\Rightarrow x_{\text{Med}} = 23$

## Vergleich von Lagemaßen anhand eines Säulendiagramms

### Mathematiknoten einer Klassenarbeit



Die Noten werden in diesem Beispiel metrisch skaliert, dh. es soll auch Zwischennoten geben.

#### Häufigkeitstabelle

	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$\Sigma$
Note ( $x_i$ )	1	2	3	4	5	6	
abs. Häufigkeit ( $n_i$ )	3	6	8	1	0	0	$n = \sum_{i=1}^6 n_i = 18$

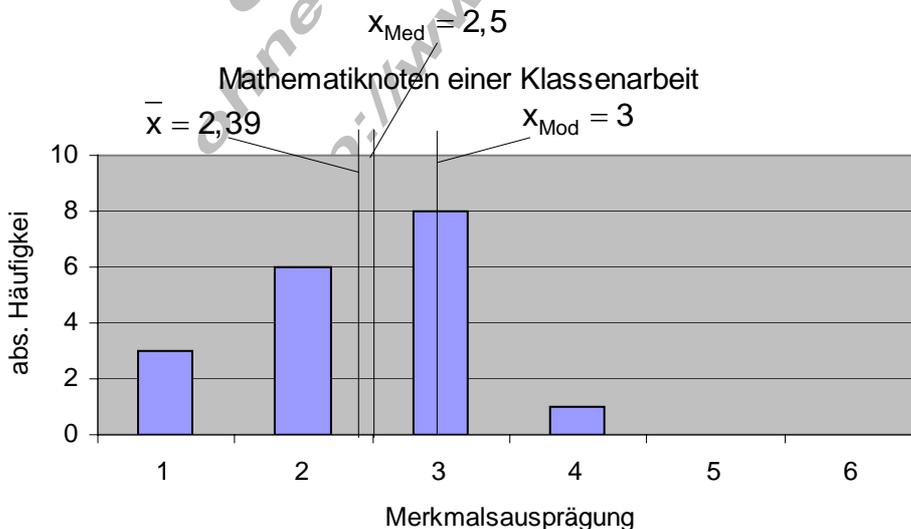
Mittelwert:  $\bar{x} = \frac{1}{n} \sum_{i=1}^6 x_i \cdot n_i = \frac{1}{18} (1 \cdot 3 + 2 \cdot 6 + 3 \cdot 8 + 4 \cdot 1) = \underline{\underline{2,39}}$

Modalwert:  $x_{\text{Mod}} = x_3 = \underline{\underline{3}}$  (kommt 8 mal vor)

Median:

1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4  $\Rightarrow x_{\text{Med}} = \underline{\underline{2,5}}$

Lagemaße im Säulendiagramm eingezeichnet



## Das Stängel – Blatt – Diagramm

Zur Bestimmung des Medians müssen die Daten (Merkmalsausprägungen) geordnet werden. Das kann mühsam sein.

Eine Erleichterung bietet hier das Stängel – Blatt – Diagramm

Beispiel:

Die Daten einer Urliste lauten

Nr.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
kg	52	67	60	55	63	63	70	78	84	68	63	57	67	58	70	73	55	72	51	60	64	51	54	68	75	76	59

Die Daten werden im Stängel – Blatt – Diagramm geordnet.

Einheit der Stängel 10, Einheit der Blätter 1

Stängel	Blätter
5	1 1 2 4 5 5 7 8 9
6	0 0 0 3 3 4 7 7 8 8    Stengel 6 Blatt 4 steht für 64
7	0 0 2 3 5 6 8
8	4

Die Daten werden nach den Stängeln (Zehnerzahlen) geordnet.

Zu jedem Stängel werden dann die Blätter (Einerzahlen) der Größe nach hinzugeschrieben.

Die meisten Daten liegen im 2. Stängel.

Der Wert der größten Häufigkeit (Modalwert) ist  $x_{\text{Mod}} = 60$

An der 14. Stelle steht der Median  $x_{\text{Med}} = 63$