

Statistik

Anbei ein Datensatz welcher die Monatsgehälter (in EUR) von 20 Mitarbeitern eines Unternehmens enthält.

Es sollen damit die Berechnung grundlegender statistischer Maße wie Mittelwert, Median, Modus, Spannweite, Varianz und Standardabweichung gezeigt werden.

Gehaltsdaten

Mitarbeiter:in	Gehalt [€]	Mitarbeiter:in	Gehalt [€]
MA 1	2000	MA 11	3500
MA 2	2100	MA 12	3600
MA 3	2200	MA 13	3800
MA 4	2300	MA 14	4000
MA 5	2500	MA 15	4200
MA 6	2500	MA 16	4500
MA 7	2600	MA 17	5000
MA 8	2700	MA 18	5500
MA 9	2800	MA 19	6000
MA 10	3000	MA 20	12000

Interpretation der statistischen Maße

Mittelwert (Durchschnitt)

Er zeigt den durchschnittlichen Verdienst in diesem Unternehmen. Der Mittelwert wird stark von Ausreißern beeinflusst.

$$\text{Mittelwert} = \frac{\sum x_i}{n}$$

$\sum x_i$ = Summe aller Gehälter

n = Anzahl der Personen im Datensatz

Mittelwert (Durchschnitt)

1. Summe aller Gehälter:

$$2000 + 2100 + 2200 + \dots + 12000 = 77000 \text{ EUR}$$

2. Anzahl der Personen:

$$n = 20$$

3. Mittelwert berechnen:

$$\text{Mittelwert} = \frac{77000}{20} = 3850 \text{ EUR}$$

Median

Er teilt den Datensatz in zwei Hälften: Die eine Hälfte der Mitarbeiter verdient weniger, die andere mehr. Der Median ist robuster gegenüber Ausreißern als der Mittelwert.

- Datensatz sortieren
- Anzahl der Werte feststellen
- mittleren Werte sind an Stelle 10 und 11

$$\text{Median} = \frac{3000 + 3500}{2} = 3250 \text{ EUR}$$

Modus

Der Modus zeigt den häufigsten Wert.

Eine Clusterung in Bereiche kann Sinn machen.

Der Modus ist nützlich für kategoriale Daten - z.B. häufige Antworten bei Umfragen.

- Zählen der Häufigkeit jedes Wertes
- Der am häufigsten vorkommende Wert ist der Modus
- Der Wert **2500** EUR kommt **2** mal vor.

Spannweite

Sie gibt an, wie groß der Unterschied zwischen dem höchsten und niedrigsten Wert ist. Die Spannweite ist ein sehr einfaches Maß für die Streuung.

$$\text{Spannweite} = \max(x_i) - \min(x_i)$$

$\max(x_i)$ = der größte Wert im Datensatz

$\min(x_i)$ = der kleinste Wert im Datensatz

Maximalwert (höchstes Gehalt) = 12.000 EUR

Minimalwert (niedrigstes Gehalt) = 2.000 EUR

$$\text{Spannweite} = 12.000 - 2.000 = 10.000 \text{ EUR}$$

Varianz

Sie misst die durchschnittliche quadratische Abweichung vom Mittelwert.

$$\text{Varianz} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

- Mittelwert berechnen
- Abweichung jedes Wertes zum Mittelwert berechnen
- Abweichungen quadrieren (damit werden sie positiv)
- Summe bilden
- durch die Anzahl der Werte dividieren

Standardabweichung

Die Standardabweichung misst die durchschnittliche Abweichung der Werte vom Mittelwert in einem Datensatz. Sie zeigt an, wie stark die einzelnen Werte um den Mittelwert streuen.

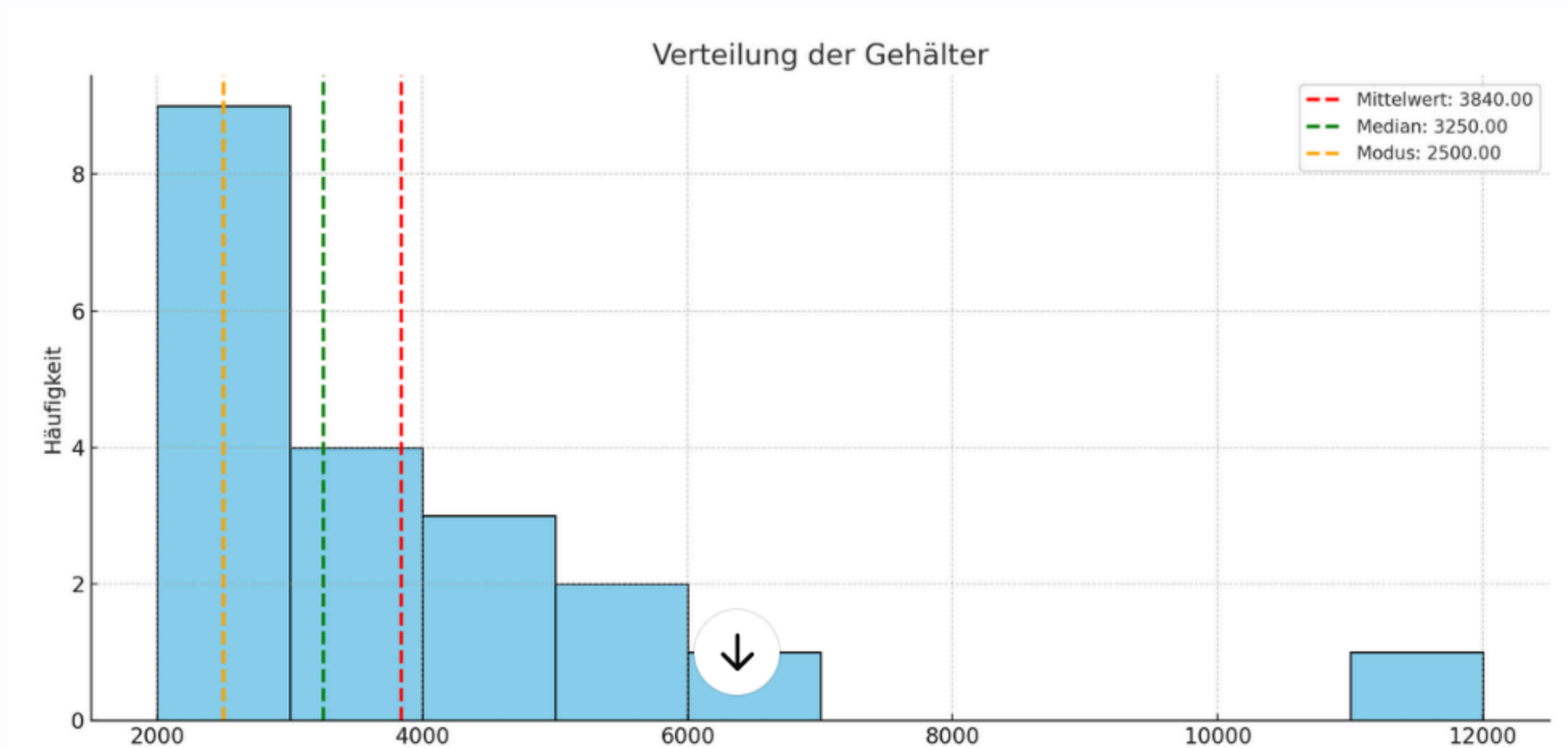
Je größer die Standardabweichung, desto weiter liegen die Werte auseinander. Eine kleine Standardabweichung bedeutet, dass die Werte eng um den Mittelwert gruppiert sind.

$$\text{Standardabweichung} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

Statistische Maße

Maß	Wert
Mittelwert	3840 EUR
Median	3250 EUR
Modus	2500 EUR
Spannweite	10000 EUR
Varianz	5042526 EUR
Standardabweichung	2245 EUR

Diagramm



Noch Fragen?

d.steiner@htl-leonding.ac.at

