

# Statistik

**Grundgesamtheit:** Menge aller für eine Fragestellung relevanten Objekte.

**Stichprobe:** Teilmenge der Grundgesamtheit, die befragt wird / für die Merkmale erhoben werden.

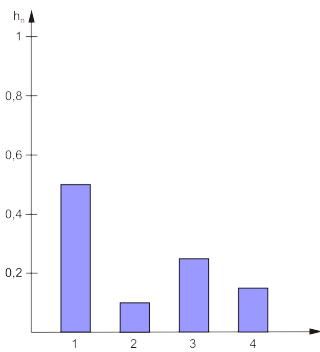
**Nominalskala:** Skala nach Kategorien geordnet (Geschlecht, Sprache, Nationalität).

**Ordinalskala:** Skala nach Grad/Größe geordnet (Bildungsniveau, Dienstgrad).

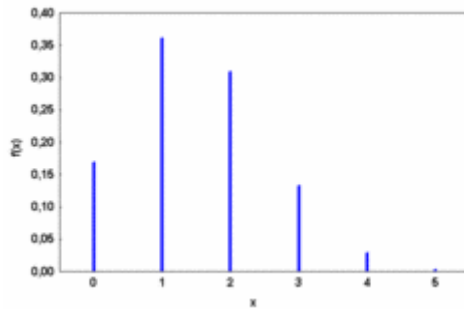
**Metrische Skala:** Skala für diskrete Merkmale (Alter, Einwohnerzahl, Gewicht).

**Absolute Häufigkeit:** Absoluter Wert / absolute Größe von etwas (14 Buben in einer Klasse).

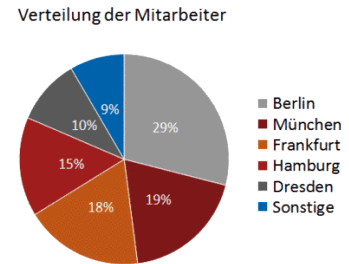
**Relative Häufigkeit:** Wert / Größe als Teilmenge relativ zu einer Grundmenge (14 Buben in einer Klasse mit 20 Schülern =  $14 / 20 = 70\%$ ).



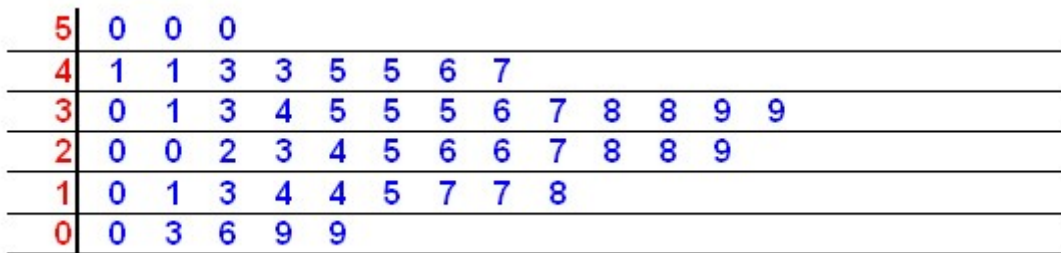
## Säulendiagramm



## Stabdiagramm



### Kreisdiagramm



## Stängelblattdiagramm

**Arithmetisches Mittel** (Mittelwert, Durchschnittswert,  $\bar{x}$ ):  $\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$

**Median** (Zentralwert,  $\tilde{x}$ ): Wert in der Mitte einer Stichprobe. Bei gerader Stichprobenlänge ist der Median das arithmetische Mittel aus den beiden mittleren Werten.

**Modus** (Modalwert,  $x_D$ ): Der / die am Häufigsten vorkommende(n) Wert(e) einer Stichprobe.

**Spannweite:** Die Differenz zwischen dem größten Wert (Maximum) und dem kleinsten Wert (Minimum) einer Stichprobe.

Varianz einer Stichprobe (Mean-Square):

$$V = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$$

Umgeformt:

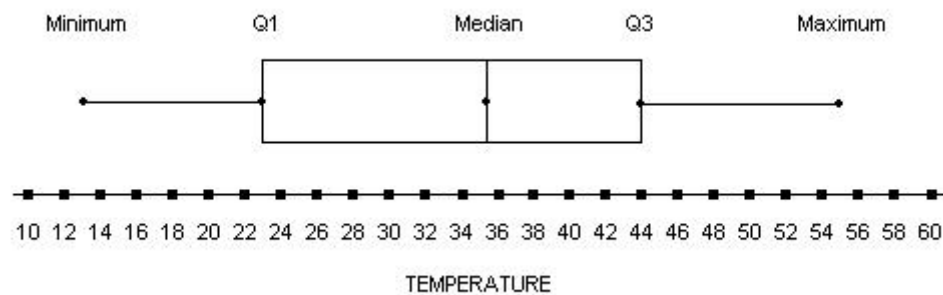
$$V = \frac{x_1^2 + x_2^2 + \dots + x_n^2}{n} - (\bar{x})^2$$

Standardabweichung einer Stichprobe: (Root-Mean-Square,  $\sigma$ ): Wurzel der Varianz

$$\sigma = \sqrt{V}$$

$$\sigma = \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_n^2}{n} - (\bar{x})^2}$$

Bei Varianz und Standardabweichung quadriert man alle Differenzen damit man keine negativen Werte bekommt, wodurch sich manche Werte aufheben würden.



Boxplot Diagramm

**Stichprobe:** Alle Daten eines Boxplot Diagramms  $x_1, x_2, \dots, x_n$

Das **erste Quartil**  $q_1$  ist der Median der Daten, die unter  $\tilde{x}$  liegen (25 %). Hier 23 (  $\frac{22+24}{2}$  )

Das **zweite Quartil**  $q_2$  ist gleich dem Median (Zentralwert)  $\tilde{x}$  (50 %). Hier 35 (  $\frac{34+35}{2}$  )

Das **dritte Quartil**  $q_3$  ist der Median der Daten, die über  $\tilde{x}$  liegen (75 %). Hier 44.

**Minimum:** Unteres Ende der Stichprobe. Hier 13.

**Maximum:** Oberes Ende der Stichprobe. Hier 56.

**Quartilsabstand:** Differenz  $q_3 - q_1$  . Hier 43 (  $56 - 13$  ).