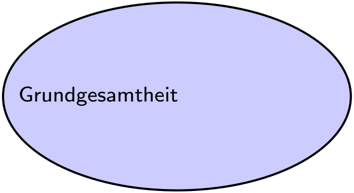
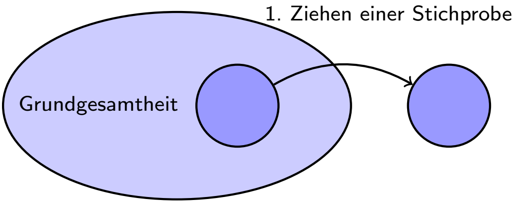
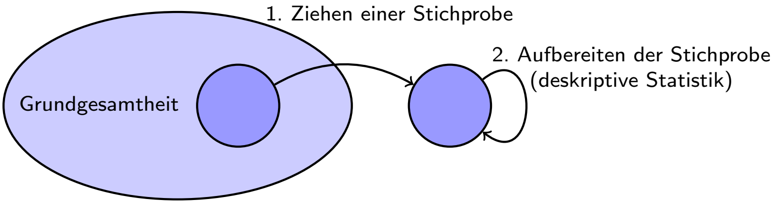
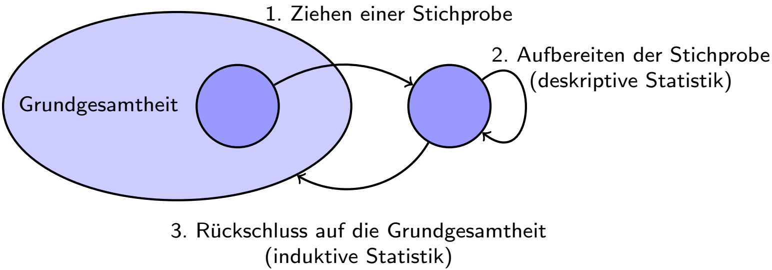
Deskriptive Statistik

Theorie


Grundgesamtheit


1. Ziehen einer Stichprobe
Grundgesamtheit


1. Ziehen einer Stichprobe
2. Aufbereiten der Stichprobe
(deskriptive Statistik)
Grundgesamtheit


1. Ziehen einer Stichprobe
2. Aufbereiten der Stichprobe
(deskriptive Statistik)
Grundgesamtheit
3. Rückschluss auf die Grundgesamtheit
(induktive Statistik)

In der deskriptiven (=beschreibenden) Statistik werden Untersuchungsergebnisse

In der deskriptiven (=beschreibenden) Statistik werden Untersuchungsergebnisse

- übersichtlich dargestellt,

In der deskriptiven (=beschreibenden) Statistik werden Untersuchungsergebnisse

- übersichtlich dargestellt,
- durch Kennzahlen charakterisiert und

In der deskriptiven (=beschreibenden) Statistik werden Untersuchungsergebnisse

- übersichtlich dargestellt,
- durch Kennzahlen charakterisiert und
- grafisch veranschaulicht.

Die *Grundgesamtheit* (oder *Population*) ist eine gedankliche Konstruktion, die alle Elemente eines Untersuchungsgegenstands umfasst.

Die *Grundgesamtheit* (oder *Population*) ist eine gedankliche Konstruktion, die alle Elemente eines Untersuchungsgegenstands umfasst.

Beispiele:

Die *Grundgesamtheit* (oder *Population*) ist eine gedankliche Konstruktion, die alle Elemente eines Untersuchungsgegenstands umfasst.

Beispiele:

- Alle linkshändigen Schülerinnen

Die *Grundgesamtheit* (oder *Population*) ist eine gedankliche Konstruktion, die alle Elemente eines Untersuchungsgegenstands umfasst.

Beispiele:

- Alle linkshändigen Schülerinnen
- Alle am 31.12.2014 in Stans wohnhaften Personen

„Eine Stichprobe stellt eine Teilmenge aller Untersuchungsobjekte dar, die die untersuchungsrelevanten Eigenschaften der Grundgesamtheit möglichst genau abbilden soll.“ (Bortz, 2010)

„Eine Stichprobe stellt eine Teilmenge aller Untersuchungsobjekte dar, die die untersuchungsrelevanten Eigenschaften der Grundgesamtheit möglichst genau abbilden soll.“ (Bortz, 2010)

Warum untersucht man nicht gleich die Grundgesamtheit?

„Eine Stichprobe stellt eine Teilmenge aller Untersuchungsobjekte dar, die die untersuchungsrelevanten Eigenschaften der Grundgesamtheit möglichst genau abbilden soll.“ (Bortz, 2010)

Warum untersucht man nicht gleich die Grundgesamtheit?

- Grundgesamtheiten sind für eine Vollerhebung oft zu gross.

„Eine Stichprobe stellt eine Teilmenge aller Untersuchungsobjekte dar, die die untersuchungsrelevanten Eigenschaften der Grundgesamtheit möglichst genau abbilden soll.“ (Bortz, 2010)

Warum untersucht man nicht gleich die Grundgesamtheit?

- Grundgesamtheiten sind für eine Vollerhebung oft zu gross.
- Manche Untersuchungen zerstören die Merkmalsträger (z. B. Reissfestigkeit von Bergseilen)

- *Einfache Zufallsstichprobe:* Aus der Grundgesamtheit werden zufällig n Objekte (ohne Zurücklegen) gezogen.

- *Einfache Zufallsstichprobe:* Aus der Grundgesamtheit werden zufällig n Objekte (ohne Zurücklegen) gezogen.
- *Klumpenstichprobe:* In zufällig ausgewählten „Klumpen“ (Schulen, Gemeinden, Kliniken) werden *alle* Objekte untersucht.

- *Einfache Zufallsstichprobe:* Aus der Grundgesamtheit werden zufällig n Objekte (ohne Zurücklegen) gezogen.
- *Klumpenstichprobe:* In zufällig ausgewählten „Klumpen“ (Schulen, Gemeinden, Kliniken) werden *alle* Objekte untersucht.
- *Geschichtete Zufallsstichprobe:* Falls bekannt ist, welche Grösse(n) die zu untersuchenden Objekte beeinflussen, ist eine Zerlegung der Grundgesamtheit in entsprechende Kategorien (Schichten) sinnvoll. *Beispiel:* Möchte man die Konsumgewohnheiten in der Schweiz untersuchen, so kann z. B. das Einkommen als Schichtungsmerkmal verwendet werden.

- *Einfache Zufallsstichprobe:* Aus der Grundgesamtheit werden zufällig n Objekte (ohne Zurücklegen) gezogen.
- *Klumpenstichprobe:* In zufällig ausgewählten „Klumpen“ (Schulen, Gemeinden, Kliniken) werden *alle* Objekte untersucht.
- *Geschichtete Zufallsstichprobe:* Falls bekannt ist, welche Grösse(n) die zu untersuchenden Objekte beeinflussen, ist eine Zerlegung der Grundgesamtheit in entsprechende Kategorien (Schichten) sinnvoll. *Beispiel:* Möchte man die Konsumgewohnheiten in der Schweiz untersuchen, so kann z. B. das Einkommen als Schichtungsmerkmal verwendet werden.
- *Ad-hoc-Stichprobe:* Familie, Schulklasse, Freundeskreis. Bequem; aber für eine Verallgemeinerung in der Regel ungeeignet.

Merkmale beschreiben Eigenschaften einer Population und damit auch einer Stichprobe.

Merkmale beschreiben Eigenschaften einer Population und damit auch einer Stichprobe.

Ein Merkmal besteht aus einem *Merkmalsträger* und einer *Merkmalsausprägung* (*Faktor*).

Merkmale beschreiben Eigenschaften einer Population und damit auch einer Stichprobe.

Ein Merkmal besteht aus einem *Merkmalsträger* und einer *Merkmalsausprägung* (*Faktor*).

Beispiel: In der Schweiz erwerbstätige Personen

Merkmale beschreiben Eigenschaften einer Population und damit auch einer Stichprobe.

Ein Merkmal besteht aus einem *Merkmalsträger* und einer *Merkmalsausprägung* (*Faktor*).

Beispiel: In der Schweiz erwerbstätige Personen

Merkmalsträger:

Merkmale beschreiben Eigenschaften einer Population und damit auch einer Stichprobe.

Ein Merkmal besteht aus einem *Merkmalsträger* und einer *Merkmalsausprägung* (*Faktor*).

Beispiel: In der Schweiz erwerbstätige Personen

Merkmalsträger: Person (AHV-Nr. ...)

Merkmale beschreiben Eigenschaften einer Population und damit auch einer Stichprobe.

Ein Merkmal besteht aus einem *Merkmalsträger* und einer *Merkmalsausprägung* (*Faktor*).

Beispiel: In der Schweiz erwerbstätige Personen

Merkmalsträger: Person (AHV-Nr. ...)

Merkmal:

Merkmale beschreiben Eigenschaften einer Population und damit auch einer Stichprobe.

Ein Merkmal besteht aus einem *Merkmalsträger* und einer *Merkmalsausprägung* (*Faktor*).

Beispiel: In der Schweiz erwerbstätige Personen

Merkmalsträger:	Person (AHV-Nr. ...)
Merkmal:	Geschlecht

Merkmale beschreiben Eigenschaften einer Population und damit auch einer Stichprobe.

Ein Merkmal besteht aus einem *Merkmalsträger* und einer *Merkmalsausprägung* (*Faktor*).

Beispiel: In der Schweiz erwerbstätige Personen

Merkmalsträger: Person (AHV-Nr. ...)

Merkmal: Geschlecht

Merkmalsausprägung:

Merkmale beschreiben Eigenschaften einer Population und damit auch einer Stichprobe.

Ein Merkmal besteht aus einem *Merkmalsträger* und einer *Merkmalsausprägung* (*Faktor*).

Beispiel: In der Schweiz erwerbstätige Personen

Merkmalsträger:	Person (AHV-Nr. ...)
Merkmal:	Geschlecht
Merkmalsausprägung:	weiblich

Merkmale beschreiben Eigenschaften einer Population und damit auch einer Stichprobe.

Ein Merkmal besteht aus einem *Merkmalsträger* und einer *Merkmalsausprägung* (*Faktor*).

Beispiel: In der Schweiz erwerbstätige Personen

Merkmalsträger: Person (AHV-Nr. ...)

Merkmal: Geschlecht

Merkmalsausprägung: weiblich

Merkmal:

Merkmale beschreiben Eigenschaften einer Population und damit auch einer Stichprobe.

Ein Merkmal besteht aus einem *Merkmalsträger* und einer *Merkmalsausprägung* (*Faktor*).

Beispiel: In der Schweiz erwerbstätige Personen

Merkmalsträger:	Person (AHV-Nr. ...)
Merkmal:	Geschlecht
Merkmalsausprägung:	weiblich
Merkmal:	Jahreseinkommen

Merkmale beschreiben Eigenschaften einer Population und damit auch einer Stichprobe.

Ein Merkmal besteht aus einem *Merkmalsträger* und einer *Merkmalsausprägung* (*Faktor*).

Beispiel: In der Schweiz erwerbstätige Personen

Merkmalsträger:	Person (AHV-Nr. ...)
Merkmal:	Geschlecht
Merkmalsausprägung:	weiblich
Merkmal:	Jahreseinkommen
Merkmalsausprägung:	

Merkmale beschreiben Eigenschaften einer Population und damit auch einer Stichprobe.

Ein Merkmal besteht aus einem *Merkmalsträger* und einer *Merkmalsausprägung* (*Faktor*).

Beispiel: In der Schweiz erwerbstätige Personen

Merkmalsträger:	Person (AHV-Nr. ...)
Merkmal:	Geschlecht
Merkmalsausprägung:	weiblich
Merkmal:	Jahreseinkommen
Merkmalsausprägung:	CHF 74 000

S. S. Stevens definiert „Messen" wie folgt (1946):

> *Measurement, in the broadest sense, is defined as the assignment of numerals to objects or events according to rule.*

S. S. Stevens definiert „Messen“ wie folgt (1946):

> *Measurement, in the broadest sense, is defined as the assignment of numerals to objects or events according to rule.*

Gesetzlich ist „Messen“ ein *Vergleich* mit einer Skala (DIN 1319, Teil 1).

S. S. Stevens definiert „Messen" wie folgt (1946):

> *Measurement, in the broadest sense, is defined as the assignment of numerals to objects or events according to rule.*

Gesetzlich ist „Messen" ein *Vergleich* mit einer Skala (DIN 1319, Teil 1).

Fasst man also einen Messvorgang so auf, dass man einem Merkmalsträger eine Zahl zuordnet, so stellt sich die Frage, welche mathematischen Operationen auf der jeweiligen Skala möglich sind.

Nominalskala

Die Skalenwerte haben keinen Zusammenhang mit den Objekteigenschaften. Sie dienen nur der Kategorisierung der Objekte.

Nominalskala

Die Skalenwerte haben keinen Zusammenhang mit den Objekteigenschaften. Sie dienen nur der Kategorisierung der Objekte. *Beispiele:*

Nominalskala

Die Skalenwerte haben keinen Zusammenhang mit den Objekteigenschaften. Sie dienen nur der Kategorisierung der Objekte. *Beispiele:*

- Geschlecht

Nominalskala

Die Skalenwerte haben keinen Zusammenhang mit den Objekteigenschaften. Sie dienen nur der Kategorisierung der Objekte. *Beispiele:*

- Geschlecht
- Haarfarbe

Nominalskala

Die Skalenwerte haben keinen Zusammenhang mit den Objekteigenschaften. Sie dienen nur der Kategorisierung der Objekte. *Beispiele:*

- Geschlecht
- Haarfarbe
- Nationalität

Nominalskala

Die Skalenwerte haben keinen Zusammenhang mit den Objekteigenschaften. Sie dienen nur der Kategorisierung der Objekte. *Beispiele:*

- Geschlecht
- Haarfarbe
- Nationalität
- Konfession

Nominalskala

Die Skalenwerte haben keinen Zusammenhang mit den Objekteigenschaften. Sie dienen nur der Kategorisierung der Objekte. *Beispiele:*

- Geschlecht
- Haarfarbe
- Nationalität
- Konfession
- Postleitzahl

Nominalskala

Die Skalenwerte haben keinen Zusammenhang mit den Objekteigenschaften. Sie dienen nur der Kategorisierung der Objekte. *Beispiele:*

- Geschlecht
- Haarfarbe
- Nationalität
- Konfession
- Postleitzahl

Die einzigen Operation auf dieser Skala sind der Test auf Gleichheit ($=$) und der auf Ungleichheit ($\neq$)

Nominalskala

Die Skalenwerte haben keinen Zusammenhang mit den Objekteigenschaften. Sie dienen nur der Kategorisierung der Objekte. *Beispiele:*

- Geschlecht
- Haarfarbe
- Nationalität
- Konfession
- Postleitzahl

Die einzigen Operation auf dieser Skala sind der Test auf Gleichheit ($=$) und der auf Ungleichheit ($\neq$)

Die Ausprägungen eines nominalskalierten Merkmals können also „nur“ kategorisiert werden.

Ordinalskala

Die Skalenwerte erlauben es, eine Ordnungsbeziehung zwischen den Objekten herzustellen.

Ordinalskala

Die Skalenwerte erlauben es, eine Ordnungsbeziehung zwischen den Objekten herzustellen. *Beispiele:*

Ordinalskala

Die Skalenwerte erlauben es, eine Ordnungsbeziehung zwischen den Objekten herzustellen. *Beispiele:*

- Zufriedenheit

Ordinalskala

Die Skalenwerte erlauben es, eine Ordnungsbeziehung zwischen den Objekten herzustellen. *Beispiele:*

- Zufriedenheit
- Rangfolge bei Schönheitswettbewerben

Ordinalskala

Die Skalenwerte erlauben es, eine Ordnungsbeziehung zwischen den Objekten herzustellen. *Beispiele:*

- Zufriedenheit
- Rangfolge bei Schönheitswettbewerben
- Güteklassen von Lebensmitteln

Ordinalskala

Die Skalenwerte erlauben es, eine Ordnungsbeziehung zwischen den Objekten herzustellen. *Beispiele:*

- Zufriedenheit
- Rangfolge bei Schönheitswettbewerben
- Güteklassen von Lebensmitteln
- Windstärke (Beaufort-Skala)

Ordinalskala

Die Skalenwerte erlauben es, eine Ordnungsbeziehung zwischen den Objekten herzustellen. *Beispiele:*

- Zufriedenheit
- Rangfolge bei Schönheitswettbewerben
- Güteklassen von Lebensmitteln
- Windstärke (Beaufort-Skala)
- Zeugnisnoten

Ordinalskala

Die Skalenwerte erlauben es, eine Ordnungsbeziehung zwischen den Objekten herzustellen. *Beispiele:*

- Zufriedenheit
- Rangfolge bei Schönheitswettbewerben
- Güteklassen von Lebensmitteln
- Windstärke (Beaufort-Skala)
- Zeugnisnoten

Bei der Ordinalskala kommen zu den Operationen der Nominalskala ($=$, $\neq$) die Ordnungsbeziehungen *kleiner als* ($<$) bzw. *grösser als* ($>$) hinzu.

Ordinalskala

Die Skalenwerte erlauben es, eine Ordnungsbeziehung zwischen den Objekten herzustellen. *Beispiele:*

- Zufriedenheit
- Rangfolge bei Schönheitswettbewerben
- Güteklassen von Lebensmitteln
- Windstärke (Beaufort-Skala)
- Zeugnisnoten

Bei der Ordinalskala kommen zu den Operationen der Nominalskala ($=$, $\neq$) die Ordnungsbeziehungen *kleiner als* ($<$) bzw. *grösser als* ($>$) hinzu.

Die Ausprägungen eines ordinalskalierten Merkmals können neben der Kategorisierung auch „rangiert“, d. h. in eine Reihenfolge gebracht werden.

Intervallskala

Die Intervallskala erlaubt es, die Differenzen zwischen den Skalenwerten der Objekte zu vergleichen.

Intervallskala

Die Intervallskala erlaubt es, die Differenzen zwischen den Skalenwerten der Objekte zu vergleichen. *Beispiele:*

Intervallskala

Die Intervallskala erlaubt es, die Differenzen zwischen den Skalenwerten der Objekte zu vergleichen. *Beispiele:*

- Temperatur (in Grad Celsius)

Intervallskala

Die Intervallskala erlaubt es, die Differenzen zwischen den Skalenwerten der Objekte zu vergleichen. *Beispiele:*

- Temperatur (in Grad Celsius)
- Jahreszahlen

Intervallskala

Die Intervallskala erlaubt es, die Differenzen zwischen den Skalenwerten der Objekte zu vergleichen. *Beispiele:*

- Temperatur (in Grad Celsius)
- Jahreszahlen
- Zeitpunkte

Intervallskala

Die Intervallskala erlaubt es, die Differenzen zwischen den Skalenwerten der Objekte zu vergleichen. *Beispiele:*

- Temperatur (in Grad Celsius)
- Jahreszahlen
- Zeitpunkte
- IQ

Intervallskala

Die Intervallskala erlaubt es, die Differenzen zwischen den Skalenwerten der Objekte zu vergleichen. *Beispiele:*

- Temperatur (in Grad Celsius)
- Jahreszahlen
- Zeitpunkte
- IQ

Neben der Kategorisierung und der Rangierung ($=$, $\neq$, $<$, $>$) können bei der Intervallskala zusätzlich die Differenzen der Ausprägungen sinnvoll interpretiert werden.

Achtung: Sind $T_1 = 08{:}10$ Uhr und $T_2 = 08{:}55$ Uhr zwei Zeit*punkte*, so ist ihre Differenz $\Delta T = T_2 - T_1 = 45$ Minuten, aber nicht ihre Summe $T_1 + T_2$ sinnvoll definiert. Hingegen können wir zu einem Zeitpunkt T eine Zeitdifferenz ΔT addieren: $T_4 = T_3 + \Delta T = 10{:}00\,\text{Uhr} + 45\ \text{Minuten} = 10{:}45\,\text{Uhr}$.

Verhältnisskala

Die Verhältnisskala erlaubt es, Verhältnisse zwischen den Skalenwerten der Objekte zu vergleichen.

Verhältnisskala

Die Verhältnisskala erlaubt es, Verhältnisse zwischen den Skalenwerten der Objekte zu vergleichen. *Beispiele:*

Verhältnisskala

Die Verhältnisskala erlaubt es, Verhältnisse zwischen den Skalenwerten der Objekte zu vergleichen. *Beispiele:*

- Länge

Verhältnisskala

Die Verhältnisskala erlaubt es, Verhältnisse zwischen den Skalenwerten der Objekte zu vergleichen. *Beispiele:*

- Länge
- Punktzahlen

Verhältnisskala

Die Verhältnisskala erlaubt es, Verhältnisse zwischen den Skalenwerten der Objekte zu vergleichen. *Beispiele:*

- Länge
- Punktzahlen
- Alter

Verhältnisskala

Die Verhältnisskala erlaubt es, Verhältnisse zwischen den Skalenwerten der Objekte zu vergleichen. *Beispiele:*

- Länge
- Punktzahlen
- Alter
- Körpergrösse

Verhältnisskala

Die Verhältnisskala erlaubt es, Verhältnisse zwischen den Skalenwerten der Objekte zu vergleichen. *Beispiele:*

- Länge
- Punktzahlen
- Alter
- Körpergrösse

Bei der Verhältnisskala kommt zu den Operationen der Intervallskala ($=$, $\neq$, $<$, $>$, $-$) die Quotientenbildung ($x \div y$) hinzu.

$$\sum_{k=1}^{4}(2k-1)$$

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1}$$

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1} + \underbrace{(2\cdot 2-1)}_{k=2}$$

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1} + \underbrace{(2\cdot 2-1)}_{k=2} + \underbrace{(2\cdot 3-1)}_{k=3}$$

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1} + \underbrace{(2\cdot 2-1)}_{k=2} + \underbrace{(2\cdot 3-1)}_{k=3} + \underbrace{(2\cdot 4-1)}_{k=4}$$

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1} + \underbrace{(2\cdot 2-1)}_{k=2} + \underbrace{(2\cdot 3-1)}_{k=3} + \underbrace{(2\cdot 4-1)}_{k=4}$$

$$= 1$$

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1} + \underbrace{(2\cdot 2-1)}_{k=2} + \underbrace{(2\cdot 3-1)}_{k=3} + \underbrace{(2\cdot 4-1)}_{k=4}$$

$$= 1 + 3$$

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1} + \underbrace{(2\cdot 2-1)}_{k=2} + \underbrace{(2\cdot 3-1)}_{k=3} + \underbrace{(2\cdot 4-1)}_{k=4}$$

$$= 1 + 3 + 5$$

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1} + \underbrace{(2\cdot 2-1)}_{k=2} + \underbrace{(2\cdot 3-1)}_{k=3} + \underbrace{(2\cdot 4-1)}_{k=4}$$

$$= 1 + 3 + 5 + 7$$

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1} + \underbrace{(2\cdot 2-1)}_{k=2} + \underbrace{(2\cdot 3-1)}_{k=3} + \underbrace{(2\cdot 4-1)}_{k=4}$$

$$= 1 + 3 + 5 + 7 = 16$$

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1} + \underbrace{(2\cdot 2-1)}_{k=2} + \underbrace{(2\cdot 3-1)}_{k=3} + \underbrace{(2\cdot 4-1)}_{k=4}$$

$$= 1 + 3 + 5 + 7 = 16$$

$\sum$ Summenzeichen („Sigma“ = grosses griechisches S)

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1} + \underbrace{(2\cdot 2-1)}_{k=2} + \underbrace{(2\cdot 3-1)}_{k=3} + \underbrace{(2\cdot 4-1)}_{k=4}$$

$$= 1+3+5+7 = 16$$

$\sum$	Summenzeichen („Sigma“ $=$ grosses griechisches S)
k	Summationsindex (manchmal auch i, j, oder l)

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1} + \underbrace{(2\cdot 2-1)}_{k=2} + \underbrace{(2\cdot 3-1)}_{k=3} + \underbrace{(2\cdot 4-1)}_{k=4}$$

$$= 1+3+5+7 = 16$$

$\sum$	Summenzeichen („Sigma“ = grosses griechisches S)
k	Summationsindex (manchmal auch i, j, oder l)
$k=1$	Der Summationsindex k startet hier bei 1 ...

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1} + \underbrace{(2\cdot 2-1)}_{k=2} + \underbrace{(2\cdot 3-1)}_{k=3} + \underbrace{(2\cdot 4-1)}_{k=4}$$

$$= 1+3+5+7 = 16$$

$\sum$	Summenzeichen („Sigma“ = grosses griechisches S)
k	Summationsindex (manchmal auch i, j, oder l)
$k = 1$	Der Summationsindex k startet hier bei 1 ...
	... und läuft in Einerschritten bis zum ...

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1} + \underbrace{(2\cdot 2-1)}_{k=2} + \underbrace{(2\cdot 3-1)}_{k=3} + \underbrace{(2\cdot 4-1)}_{k=4}$$

$$= 1+3+5+7 = 16$$

$\sum$	Summenzeichen („Sigma“ = grosses griechisches S)
k	Summationsindex (manchmal auch i, j, oder l)
$k = 1$	Der Summationsindex k startet hier bei 1 ...
	... und läuft in Einerschritten bis zum ...
4	Summationsende $k = 4$

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1} + \underbrace{(2\cdot 2-1)}_{k=2} + \underbrace{(2\cdot 3-1)}_{k=3} + \underbrace{(2\cdot 4-1)}_{k=4}$$

$$= 1+3+5+7 = 16$$

$\sum$	Summenzeichen („Sigma“ $=$ grosses griechisches S)
k	Summationsindex (manchmal auch i, j, oder l)
$k=1$	Der Summationsindex k startet hier bei 1 ...
	... und läuft in Einerschritten bis zum ...
4	Summationsende $k=4$
$(2k-1)$	allgemeiner Summand

$$\sum_{k=1}^{4}(2k-1) = \underbrace{(2\cdot 1-1)}_{k=1} + \underbrace{(2\cdot 2-1)}_{k=2} + \underbrace{(2\cdot 3-1)}_{k=3} + \underbrace{(2\cdot 4-1)}_{k=4}$$

$$= 1+3+5+7 = 16$$

$\sum$	Summenzeichen („Sigma“ = grosses griechisches S)
k	Summationsindex (manchmal auch i, j, oder l)
$k = 1$	Der Summationsindex k startet hier bei 1 …
	… und läuft in Einerschritten bis zum …
4	Summationsende $k = 4$
$(2k-1)$	allgemeiner Summand

Sprich: „Die Summe über alle $(2k-1)$, wobei k von 1 bis 4 läuft.“

Übersicht

Liegen von einem Merkmal mehrere quantitative Merkmalsausprägungen vor, so lassen sich diese Daten im Hinblick auf die folgenden zwei Gesichtspunkte statistisch charakterisieren:

Übersicht

Liegen von einem Merkmal mehrere quantitative Merkmalsausprägungen vor, so lassen sich diese Daten im Hinblick auf die folgenden zwei Gesichtspunkte statistisch charakterisieren:

- Wo liegt die Mitte der Daten? (*zentrale Tendenz*)

Übersicht

Liegen von einem Merkmal mehrere quantitative Merkmalsausprägungen vor, so lassen sich diese Daten im Hinblick auf die folgenden zwei Gesichtspunkte statistisch charakterisieren:

- Wo liegt die Mitte der Daten? (*zentrale Tendenz*)
- Wie stark streuen die Daten? (*Variabilität*, *Streuung*)

In einer Umfrage wurden an einem Werktag zu unterschiedlichen Zeitpunkten 100 zufällig ausgewählte Bahnreisende nach ihrem Alter befragt.

51	9	44	76	38	43	13	41	82	84
74	2	31	46	45	38	80	58	27	65
7	25	46	18	29	20	17	53	6	56
17	21	59	20	43	63	45	43	17	11
21	3	27	59	58	27	17	62	51	53
47	10	17	37	18	14	35	28	14	39
10	59	29	11	20	13	32	1	17	55
16	65	64	15	2	28	47	27	50	18
80	61	17	66	35	46	32	53	25	51
81	76	41	16	30	27	33	19	43	62

Der Mittelwert (arithmetisches Mittel)

Der Mittelwert ist das gebräuchlichste Mass, um die zentrale Tendenz der Verteilung eines intervall- oder verhältnisskalierten Merkmals zu beschreiben.

Der Mittelwert (arithmetisches Mittel)

Der Mittelwert ist das gebräuchlichste Mass, um die zentrale Tendenz der Verteilung eines intervall- oder verhältnisskalierten Merkmals zu beschreiben.

Mittelwert einer Stichprobe $x_1, x_2, \ldots, x_n$:

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

Der Mittelwert (arithmetisches Mittel)

Der Mittelwert ist das gebräuchlichste Mass, um die zentrale Tendenz der Verteilung eines intervall- oder verhältnisskalierten Merkmals zu beschreiben.

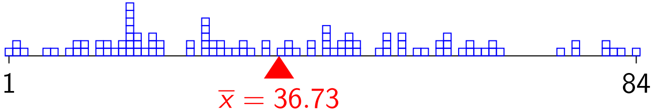
Mittelwert einer Stichprobe $x_1, x_2, \ldots, x_n$:

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

Mittelwert einer Grundgesamtheit $x_1, x_2, \ldots, x_n$:

$$\mu = \frac{x_1 + x_2 + \cdots + x_n}{n} = \cdots = \frac{1}{n}\sum_{i=1}^{n} x_i$$

In der Beispielstichprobe der Bahnreisenden gilt: $\overline{x} = 36.73$



Die Summe der Abweichungen vom Mittelwert ergibt immer Null:

Die Summe der Abweichungen vom Mittelwert ergibt immer Null:

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

Die Summe der Abweichungen vom Mittelwert ergibt immer Null:

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

Die Summe der Abweichungen vom Mittelwert ergibt immer Null:

$$\begin{aligned}\sum_{i=1}^{n}(x_i - \overline{x}) &= \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \overline{x} \\ &= \sum_{i=1}^{n} x_i - n \cdot \overline{x}\end{aligned}$$

Die Summe der Abweichungen vom Mittelwert ergibt immer Null:

$$\begin{aligned}
\sum_{i=1}^{n}(x_i - \overline{x}) &= \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \overline{x} \\
&= \sum_{i=1}^{n} x_i - n \cdot \overline{x} \\
&= \sum_{i=1}^{n} x_i - n \cdot \frac{1}{n} \sum_{i=1}^{n} x_i
\end{aligned}$$

Die Summe der Abweichungen vom Mittelwert ergibt immer Null:

$$
\begin{aligned}
\sum_{i=1}^{n}(x_i - \overline{x}) &= \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \overline{x} \\
&= \sum_{i=1}^{n} x_i - n \cdot \overline{x} \\
&= \sum_{i=1}^{n} x_i - n \cdot \frac{1}{n} \sum_{i=1}^{n} x_i \\
&= \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i
\end{aligned}
$$

Die Summe der Abweichungen vom Mittelwert ergibt immer Null:

$$
\begin{aligned}
\sum_{i=1}^{n}(x_i - \overline{x}) &= \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \overline{x} \\
&= \sum_{i=1}^{n} x_i - n \cdot \overline{x} \\
&= \sum_{i=1}^{n} x_i - n \cdot \frac{1}{n} \sum_{i=1}^{n} x_i \\
&= \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i \\
&= 0
\end{aligned}
$$

Der Median (Zentralwert)

Sind

$$x_1, x_2, \dots, x_n$$

die Werte einer Stichprobe oder einer Grundgesamtheit, so wird eine Zahl $\tilde{x}$, welche die *Ordnungsstatistik*, d. h. die Liste der sortierten Werte

$$x_{(1)}, x_{(2)}, \dots, x_{(n)}$$

in zwei gleich grosse Teile zerlegt, *Median* oder *Zentralwert* genannt.

Beispiel 3.1

Stichprobe:

$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6$

Ordnungsstatistik:

Beispiel 3.1

Stichprobe:

$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6$

Ordnungsstatistik:

$x_{(1)} = 2,$

Beispiel 3.1

Stichprobe:

$$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6$$

Ordnungsstatistik:

$$x_{(1)} = 2,\ x_{(2)} = 3,$$

Beispiel 3.1

Stichprobe:

$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6$

Ordnungsstatistik:

$x_{(1)} = 2,\ x_{(2)} = 3,\ x_{(3)} = 5,$

Beispiel 3.1

Stichprobe:

$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6$

Ordnungsstatistik:

$x_{(1)} = 2,\ x_{(2)} = 3,\ x_{(3)} = 5,\ x_{(4)} = 6,$

Beispiel 3.1

Stichprobe:

$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6$

Ordnungsstatistik:

$x_{(1)} = 2,\ x_{(2)} = 3,\ x_{(3)} = 5,\ x_{(4)} = 6,\ x_{(5)} = 9$

Beispiel 3.1

Stichprobe:

$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6$

Ordnungsstatistik:

$x_{(1)} = 2,\ x_{(2)} = 3,\ x_{(3)} = 5,\ x_{(4)} = 6,\ x_{(5)} = 9$

Median:

Beispiel 3.1

Stichprobe:

$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6$

Ordnungsstatistik:

$x_{(1)} = 2,\ x_{(2)} = 3,\ x_{(3)} = 5,\ x_{(4)} = 6,\ x_{(5)} = 9$

Median: $\tilde{x} = 5$

Beispiel 3.2

Stichprobe:

$$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6,\ x_6 = 2$$

Ordnungsstatistik:

Beispiel 3.2

Stichprobe:

$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6,\ x_6 = 2$

Ordnungsstatistik:

$x_{(1)} = 2,$

Beispiel 3.2

Stichprobe:

$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6,\ x_6 = 2$

Ordnungsstatistik:

$x_{(1)} = 2,\ x_{(2)} = 2,$

Beispiel 3.2

Stichprobe:

$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6,\ x_6 = 2$

Ordnungsstatistik:

$x_{(1)} = 2,\ x_{(2)} = 2,\ x_{(3)} = 3,$

Beispiel 3.2

Stichprobe:

$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6,\ x_6 = 2$

Ordnungsstatistik:

$x_{(1)} = 2,\ x_{(2)} = 2,\ x_{(3)} = 3,\ x_{(4)} = 5,$

Beispiel 3.2

Stichprobe:

$$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6,\ x_6 = 2$$

Ordnungsstatistik:

$$x_{(1)} = 2,\ x_{(2)} = 2,\ x_{(3)} = 3,\ x_{(4)} = 5,\ x_{(5)} = 6,$$

Beispiel 3.2

Stichprobe:

$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6,\ x_6 = 2$

Ordnungsstatistik:

$x_{(1)} = 2,\ x_{(2)} = 2,\ x_{(3)} = 3,\ x_{(4)} = 5,\ x_{(5)} = 6,\ x_{(6)} = 9$

Beispiel 3.2

Stichprobe:

$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6,\ x_6 = 2$

Ordnungsstatistik:

$x_{(1)} = 2,\ x_{(2)} = 2,\ x_{(3)} = 3,\ x_{(4)} = 5,\ x_{(5)} = 6,\ x_{(6)} = 9$

Median:

Beispiel 3.2

Stichprobe:

$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6,\ x_6 = 2$

Ordnungsstatistik:

$x_{(1)} = 2,\ x_{(2)} = 2,\ x_{(3)} = 3,\ x_{(4)} = 5,\ x_{(5)} = 6,\ x_{(6)} = 9$

Median: $\tilde{x} = 4$

Beispiel 3.2

Stichprobe:

$x_1 = 9,\ x_2 = 5,\ x_3 = 3,\ x_4 = 2,\ x_5 = 6,\ x_6 = 2$

Ordnungsstatistik:

$x_{(1)} = 2,\ x_{(2)} = 2,\ x_{(3)} = 3,\ x_{(4)} = 5,\ x_{(5)} = 6,\ x_{(6)} = 9$

Median: $\tilde{x} = 4$

Bei geradem Stichprobenumfang ist der Median das arithmetische Mittel der beiden Werte in der „Mitte“ der Ordnungsstatistik.

Wann soll man den Median verwenden?

(a) Wenn extrem grosse oder kleine Werte (*Ausreisser*) das arithmetische Mittel stark verzerren. *Beispiel:*

Stichprobe A: $x_{(1)} = 4$, $x_{(2)} = 5$, $x_{(3)} = 7$, $x_{(4)} = 8$

Stichprobe B: $x_{(1)} = 4$, $x_{(2)} = 5$, $x_{(3)} = 7$, $x_{(4)} = 80$

	Stichprobe A	Stichprobe B
$\overline{x}$		
$\tilde{x}$		

Wann soll man den Median verwenden?

(a) Wenn extrem grosse oder kleine Werte (*Ausreisser*) das arithmetische Mittel stark verzerren. *Beispiel:*

Stichprobe A: $x_{(1)} = 4$, $x_{(2)} = 5$, $x_{(3)} = 7$, $x_{(4)} = 8$

Stichprobe B: $x_{(1)} = 4$, $x_{(2)} = 5$, $x_{(3)} = 7$, $x_{(4)} = 80$

	Stichprobe A	Stichprobe B
$\bar{x}$	6	
$\tilde{x}$		

Wann soll man den Median verwenden?

(a) Wenn extrem grosse oder kleine Werte (*Ausreisser*) das arithmetische Mittel stark verzerren. *Beispiel:*

Stichprobe A: $x_{(1)} = 4$, $x_{(2)} = 5$, $x_{(3)} = 7$, $x_{(4)} = 8$

Stichprobe B: $x_{(1)} = 4$, $x_{(2)} = 5$, $x_{(3)} = 7$, $x_{(4)} = 80$

	Stichprobe A	Stichprobe B
$\overline{x}$	6	24
$\tilde{x}$		

Wann soll man den Median verwenden?

(a) Wenn extrem grosse oder kleine Werte (*Ausreisser*) das arithmetische Mittel stark verzerren. *Beispiel:*

Stichprobe A: $x_{(1)} = 4$, $x_{(2)} = 5$, $x_{(3)} = 7$, $x_{(4)} = 8$

Stichprobe B: $x_{(1)} = 4$, $x_{(2)} = 5$, $x_{(3)} = 7$, $x_{(4)} = 80$

	Stichprobe A	Stichprobe B
$\overline{x}$	6	24
$\tilde{x}$	6	

Wann soll man den Median verwenden?

(a) Wenn extrem grosse oder kleine Werte (*Ausreisser*) das arithmetische Mittel stark verzerren. *Beispiel:*

Stichprobe A: $x_{(1)} = 4$, $x_{(2)} = 5$, $x_{(3)} = 7$, $x_{(4)} = 8$

Stichprobe B: $x_{(1)} = 4$, $x_{(2)} = 5$, $x_{(3)} = 7$, $x_{(4)} = 80$

	Stichprobe A	Stichprobe B
$\bar{x}$	6	24
$\tilde{x}$	6	6

(b) Wenn man die zentrale Tendenz für ordinalskalierte Merkmalswerte bestimmen möchte. *Beispiel:*

$$\left.\begin{aligned} x_{(1)} &= \text{nie} \\ x_{(2)} &= \text{wenig} \\ x_{(3)} &= \text{wenig} \\ x_{(4)} &= \text{manchmal} \\ x_{(5)} &= \text{oft} \\ x_{(6)} &= \text{oft} \\ x_{(7)} &= \text{oft} \\ x_{(8)} &= \text{oft} \\ x_{(9)} &= \text{immer} \\ x_{(10)} &= \text{immer} \end{aligned}\right\} \quad \Rightarrow \quad \tilde{x} =$$

(b) Wenn man die zentrale Tendenz für ordinalskalierte Merkmalswerte bestimmen möchte. *Beispiel:*

$$\left.\begin{aligned} x_{(1)} &= \text{nie} \\ x_{(2)} &= \text{wenig} \\ x_{(3)} &= \text{wenig} \\ x_{(4)} &= \text{manchmal} \\ x_{(5)} &= \text{oft} \\ x_{(6)} &= \text{oft} \\ x_{(7)} &= \text{oft} \\ x_{(8)} &= \text{oft} \\ x_{(9)} &= \text{immer} \\ x_{(10)} &= \text{immer} \end{aligned}\right\} \quad \Rightarrow \quad \tilde{x} = \text{oft}$$

Der Modalwert (Modus)

Der *Modalwert* oder *Modus* ist der am häufigsten auftretende Merkmalswert.

Er kann grundsätzlich für Merkmalswerte auf allen Skalen berechnet werden.

- $x_1 = \text{ja}$, $x_2 = \text{nein}$, $x_3 = \text{nein}$, $x_4 = \text{ja}$, $x_5 = \text{nein}$

Der Modalwert (Modus)

Der *Modalwert* oder *Modus* ist der am häufigsten auftretende Merkmalswert.

Er kann grundsätzlich für Merkmalswerte auf allen Skalen berechnet werden.

- $x_1 = \text{ja}$, $x_2 = \text{nein}$, $x_3 = \text{nein}$, $x_4 = \text{ja}$, $x_5 = \text{nein}$

  Modus:

Der Modalwert (Modus)

Der *Modalwert* oder *Modus* ist der am häufigsten auftretende Merkmalswert.

Er kann grundsätzlich für Merkmalswerte auf allen Skalen berechnet werden.

- $x_1 = \text{ja}$, $x_2 = \text{nein}$, $x_3 = \text{nein}$, $x_4 = \text{ja}$, $x_5 = \text{nein}$

  Modus: nein

Der Modalwert (Modus)

Der *Modalwert* oder *Modus* ist der am häufigsten auftretende Merkmalswert.

Er kann grundsätzlich für Merkmalswerte auf allen Skalen berechnet werden.

- $x_1 = \text{ja}$, $x_2 = \text{nein}$, $x_3 = \text{nein}$, $x_4 = \text{ja}$, $x_5 = \text{nein}$

  Modus: nein

- $x_1 = \text{ja}$, $x_2 = \text{nein}$, $x_3 = \text{nein}$, $x_4 = \text{ja}$, $x_5 = \text{ja}$

Der Modalwert (Modus)

Der *Modalwert* oder *Modus* ist der am häufigsten auftretende Merkmalswert.

Er kann grundsätzlich für Merkmalswerte auf allen Skalen berechnet werden.

- $x_1 = \text{ja}$, $x_2 = \text{nein}$, $x_3 = \text{nein}$, $x_4 = \text{ja}$, $x_5 = \text{nein}$

  Modus: nein

- $x_1 = \text{ja}$, $x_2 = \text{nein}$, $x_3 = \text{nein}$, $x_4 = \text{ja}$, $x_5 = \text{ja}$

  Modus:

Der Modalwert (Modus)

Der *Modalwert* oder *Modus* ist der am häufigsten auftretende Merkmalswert.

Er kann grundsätzlich für Merkmalswerte auf allen Skalen berechnet werden.

- $x_1 = \text{ja},\ x_2 = \text{nein},\ x_3 = \text{nein},\ x_4 = \text{ja},\ x_5 = \text{nein}$

  Modus: nein

- $x_1 = \text{ja},\ x_2 = \text{nein},\ x_3 = \text{nein},\ x_4 = \text{ja},\ x_5 = \text{ja}$

  Modus: ja

Der Modalwert (Modus)

Der *Modalwert* oder *Modus* ist der am häufigsten auftretende Merkmalswert.

Er kann grundsätzlich für Merkmalswerte auf allen Skalen berechnet werden.

- $x_1 = \text{ja}$, $x_2 = \text{nein}$, $x_3 = \text{nein}$, $x_4 = \text{ja}$, $x_5 = \text{nein}$

  Modus: nein

- $x_1 = \text{ja}$, $x_2 = \text{nein}$, $x_3 = \text{nein}$, $x_4 = \text{ja}$, $x_5 = \text{ja}$

  Modus: ja

- $x_1 = 1.32$, $x_2 = 2.54$, $x_3 = 3.6$, $x_4 = 1.97$, $x_5 = 3.05$

Der Modalwert (Modus)

Der *Modalwert* oder *Modus* ist der am häufigsten auftretende Merkmalswert.

Er kann grundsätzlich für Merkmalswerte auf allen Skalen berechnet werden.

- $x_1 = \text{ja}$, $x_2 = \text{nein}$, $x_3 = \text{nein}$, $x_4 = \text{ja}$, $x_5 = \text{nein}$

  Modus: nein

- $x_1 = \text{ja}$, $x_2 = \text{nein}$, $x_3 = \text{nein}$, $x_4 = \text{ja}$, $x_5 = \text{ja}$

  Modus: ja

- $x_1 = 1.32$, $x_2 = 2.54$, $x_3 = 3.6$, $x_4 = 1.97$, $x_5 = 3.05$

  Modus:

Der Modalwert (Modus)

Der *Modalwert* oder *Modus* ist der am häufigsten auftretende Merkmalswert.

Er kann grundsätzlich für Merkmalswerte auf allen Skalen berechnet werden.

- $x_1 = \text{ja}$, $x_2 = \text{nein}$, $x_3 = \text{nein}$, $x_4 = \text{ja}$, $x_5 = \text{nein}$

  Modus: nein

- $x_1 = \text{ja}$, $x_2 = \text{nein}$, $x_3 = \text{nein}$, $x_4 = \text{ja}$, $x_5 = \text{ja}$

  Modus: ja

- $x_1 = 1.32$, $x_2 = 2.54$, $x_3 = 3.6$, $x_4 = 1.97$, $x_5 = 3.05$

  Modus: ?

Der Modalwert (Modus)

Der *Modalwert* oder *Modus* ist der am häufigsten auftretende Merkmalswert.

Er kann grundsätzlich für Merkmalswerte auf allen Skalen berechnet werden.

- $x_1 = \text{ja},\ x_2 = \text{nein},\ x_3 = \text{nein},\ x_4 = \text{ja},\ x_5 = \text{nein}$

  Modus: nein

- $x_1 = \text{ja},\ x_2 = \text{nein},\ x_3 = \text{nein},\ x_4 = \text{ja},\ x_5 = \text{ja}$

  Modus: ja

- $x_1 = 1.32,\ x_2 = 2.54,\ x_3 = 3.6,\ x_4 = 1.97,\ x_5 = 3.05$

  Modus: ? nicht definiert

Das geometrische Mittel

Für eine Familie haben sich die Krankenkassenprämien in den vergangenen vier Jahren wie folgt entwickelt.

Jahr	2011	2012	2013	2014
Anstieg	3%	4%	−5%	8%

(a) Um wie viel Prozent sind die Prämien in den letzten vier Jahren insgesamt gestiegen?

Das geometrische Mittel

Für eine Familie haben sich die Krankenkassenprämien in den vergangenen vier Jahren wie folgt entwickelt.

Jahr	2011	2012	2013	2014
Anstieg	3%	4%	−5%	8%

(a) Um wie viel Prozent sind die Prämien in den letzten vier Jahren insgesamt gestiegen?

$1.03 \cdot 1.04 \cdot 0.95 \cdot 1.08 =$

Das geometrische Mittel

Für eine Familie haben sich die Krankenkassenprämien in den vergangenen vier Jahren wie folgt entwickelt.

Jahr	2011	2012	2013	2014
Anstieg	3%	4%	-5%	8%

(a) Um wie viel Prozent sind die Prämien in den letzten vier Jahren insgesamt gestiegen?

$$1.03 \cdot 1.04 \cdot 0.95 \cdot 1.08 = 1.0990512$$

Das geometrische Mittel

Für eine Familie haben sich die Krankenkassenprämien in den vergangenen vier Jahren wie folgt entwickelt.

Jahr	2011	2012	2013	2014
Anstieg	3%	4%	−5%	8%

(a) Um wie viel Prozent sind die Prämien in den letzten vier Jahren insgesamt gestiegen?

$$1.03 \cdot 1.04 \cdot 0.95 \cdot 1.08 = 1.0990512 \quad \Rightarrow \quad +9.9\%$$

Das geometrische Mittel

Für eine Familie haben sich die Krankenkassenprämien in den vergangenen vier Jahren wie folgt entwickelt.

Jahr	2011	2012	2013	2014
Anstieg	3%	4%	−5%	8%

(a) Um wie viel Prozent sind die Prämien in den letzten vier Jahren insgesamt gestiegen?

$$1.03 \cdot 1.04 \cdot 0.95 \cdot 1.08 = 1.0990512 \quad \Rightarrow \quad +9.9\%$$

(b) Berechne den durchschnittlichen prozentualen Anstieg pro Jahr.

Das geometrische Mittel

Für eine Familie haben sich die Krankenkassenprämien in den vergangenen vier Jahren wie folgt entwickelt.

Jahr	2011	2012	2013	2014
Anstieg	3%	4%	−5%	8%

(a) Um wie viel Prozent sind die Prämien in den letzten vier Jahren insgesamt gestiegen?

$$1.03 \cdot 1.04 \cdot 0.95 \cdot 1.08 = 1.0990512 \quad \Rightarrow \quad +9.9\%$$

(b) Berechne den durchschnittlichen prozentualen Anstieg pro Jahr.

$$\sqrt[4]{1.0990512} \approx$$

Das geometrische Mittel

Für eine Familie haben sich die Krankenkassenprämien in den vergangenen vier Jahren wie folgt entwickelt.

Jahr	2011	2012	2013	2014
Anstieg	3%	4%	-5%	8%

(a) Um wie viel Prozent sind die Prämien in den letzten vier Jahren insgesamt gestiegen?

$$1.03 \cdot 1.04 \cdot 0.95 \cdot 1.08 = 1.0990512 \quad \Rightarrow \quad +9.9\%$$

(b) Berechne den durchschnittlichen prozentualen Anstieg pro Jahr.

$$\sqrt[4]{1.0990512} \approx 1.024$$

Das geometrische Mittel

Für eine Familie haben sich die Krankenkassenprämien in den vergangenen vier Jahren wie folgt entwickelt.

Jahr	2011	2012	2013	2014
Anstieg	3%	4%	−5%	8%

(a) Um wie viel Prozent sind die Prämien in den letzten vier Jahren insgesamt gestiegen?

$$1.03 \cdot 1.04 \cdot 0.95 \cdot 1.08 = 1.0990512 \quad \Rightarrow \quad +9.9\%$$

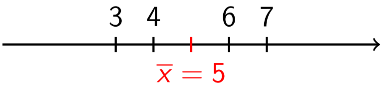
(b) Berechne den durchschnittlichen prozentualen Anstieg pro Jahr.

$$\sqrt[4]{1.0990512} \approx 1.024 \quad \Rightarrow \quad +2.4\%$$

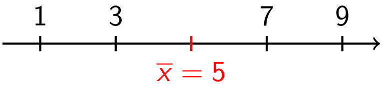
Das Konzept

Es ist möglich, dass zwei Grundgesamtheiten oder zwei Stichproben denselben Mittelwert haben; sich aber darin unterscheiden, wie stark die Daten um ihr Zentrum *streuen*.

Stichprobe A: $x_1 = 3$, $x_2 = 4$, $x_3 = 6$, $x_4 = 7$



Stichprobe B: $x_1 = 1$, $x_2 = 3$, $x_3 = 7$, $x_4 = 9$



Die Varianz einer Grundgesamtheit

Für eine Grundgesamtheit ist die *Varianz* σ^2 definiert als die mittlere quadratische Abweichung der Werte vom Mittelwert:

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

Durch die Quadrate werden Abweichungen, die kleiner als 1 sind noch kleiner gemacht und Abweichungen, die grösser als 1 sind, verstärkt. (Das ist in vielen Situation so „gewollt“.)

Die Varianz einer Stichprobe

Für eine Stichprobe ist die *empirische Varianz* s^2 definiert als

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2$$

Die Multiplikation mit $1/(n-1)$ ist nötig, damit man mit der empirischen Varianz die unbekannte Varianz der Grundgesamtheit σ^2 korrekt schätzen kann. Mit dem Faktor $1/n$ würde man systematisch zu tief liegen.

(Eine Erklärung für diese *Bessel-Korrektur* kann erst in der 6. Klasse gegeben werden.)

Die (empirische) Standardabweichung

In den Formeln für die Varianz treten die gegebenen Grössen im Quadrat auf. Um wieder mit den ursprünglichen Einheiten rechnen zu können, zieht man die Wurzel aus der (empirischen) Varianz und enthält so die *(empirische) Standardabweichung*.

- Standardabweichung: $\sigma = \sqrt{\sigma^2}$
- empirische Standardabweichung: $s = \sqrt{s^2}$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3,\ x_2 = 4,\ x_3 = 6,\ x_4 = 7$
- Stichprobe B: $x_1 = 1,\ x_2 = 3,\ x_3 = 7,\ x_4 = 9$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3$, $x_2 = 4$, $x_3 = 6$, $x_4 = 7$
- Stichprobe B: $x_1 = 1$, $x_2 = 3$, $x_3 = 7$, $x_4 = 9$

$\overline{x}_A =$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3$, $x_2 = 4$, $x_3 = 6$, $x_4 = 7$
- Stichprobe B: $x_1 = 1$, $x_2 = 3$, $x_3 = 7$, $x_4 = 9$

$\overline{x}_A = 5$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3,\ x_2 = 4,\ x_3 = 6,\ x_4 = 7$
- Stichprobe B: $x_1 = 1,\ x_2 = 3,\ x_3 = 7,\ x_4 = 9$

$\overline{x}_A = 5$

$s_A^2 =$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3,\ x_2 = 4,\ x_3 = 6,\ x_4 = 7$
- Stichprobe B: $x_1 = 1,\ x_2 = 3,\ x_3 = 7,\ x_4 = 9$

$\overline{x}_A = 5$

$$s_A^2 = \frac{(3-5)^2 + (4-5)^2 + (6-5)^2 + (7-5)^2}{3}$$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3,\ x_2 = 4,\ x_3 = 6,\ x_4 = 7$
- Stichprobe B: $x_1 = 1,\ x_2 = 3,\ x_3 = 7,\ x_4 = 9$

$$\overline{x}_A = 5$$

$$s_A^2 = \frac{(3-5)^2 + (4-5)^2 + (6-5)^2 + (7-5)^2}{3} = \frac{10}{3}$$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3,\ x_2 = 4,\ x_3 = 6,\ x_4 = 7$
- Stichprobe B: $x_1 = 1,\ x_2 = 3,\ x_3 = 7,\ x_4 = 9$

$\overline{x}_A = 5$

$$s_A^2 = \frac{(3-5)^2 + (4-5)^2 + (6-5)^2 + (7-5)^2}{3} = \frac{10}{3}$$

s_A

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3,\ x_2 = 4,\ x_3 = 6,\ x_4 = 7$
- Stichprobe B: $x_1 = 1,\ x_2 = 3,\ x_3 = 7,\ x_4 = 9$

$$\overline{x}_A = 5$$

$$s_A^2 = \frac{(3-5)^2 + (4-5)^2 + (6-5)^2 + (7-5)^2}{3} = \frac{10}{3}$$

$$s_A = \sqrt{\frac{10}{3}}$$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3,\ x_2 = 4,\ x_3 = 6,\ x_4 = 7$
- Stichprobe B: $x_1 = 1,\ x_2 = 3,\ x_3 = 7,\ x_4 = 9$

$$\overline{x}_A = 5$$

$$s_A^2 = \frac{(3-5)^2 + (4-5)^2 + (6-5)^2 + (7-5)^2}{3} = \frac{10}{3}$$

$$s_A = \sqrt{\frac{10}{3}}$$

$$\overline{x}_B =$$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3,\ x_2 = 4,\ x_3 = 6,\ x_4 = 7$
- Stichprobe B: $x_1 = 1,\ x_2 = 3,\ x_3 = 7,\ x_4 = 9$

$$\overline{x}_A = 5$$

$$s_A^2 = \frac{(3-5)^2 + (4-5)^2 + (6-5)^2 + (7-5)^2}{3} = \frac{10}{3}$$

$$s_A = \sqrt{\frac{10}{3}}$$

$$\overline{x}_B = 5$$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3,\ x_2 = 4,\ x_3 = 6,\ x_4 = 7$
- Stichprobe B: $x_1 = 1,\ x_2 = 3,\ x_3 = 7,\ x_4 = 9$

$\overline{x}_A = 5$

$$s_A^2 = \frac{(3-5)^2 + (4-5)^2 + (6-5)^2 + (7-5)^2}{3} = \frac{10}{3}$$

$$s_A = \sqrt{\frac{10}{3}}$$

$\overline{x}_B = 5$

$s_B^2 =$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3,\ x_2 = 4,\ x_3 = 6,\ x_4 = 7$
- Stichprobe B: $x_1 = 1,\ x_2 = 3,\ x_3 = 7,\ x_4 = 9$

$$\overline{x}_A = 5$$

$$s_A^2 = \frac{(3-5)^2 + (4-5)^2 + (6-5)^2 + (7-5)^2}{3} = \frac{10}{3}$$

$$s_A = \sqrt{\frac{10}{3}}$$

$$\overline{x}_B = 5$$

$$s_B^2 = \frac{(1-5)^2 + (3-5)^2 + (7-5)^2 + (9-5)^2}{3}$$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3,\ x_2 = 4,\ x_3 = 6,\ x_4 = 7$
- Stichprobe B: $x_1 = 1,\ x_2 = 3,\ x_3 = 7,\ x_4 = 9$

$$\overline{x}_A = 5$$

$$s_A^2 = \frac{(3-5)^2 + (4-5)^2 + (6-5)^2 + (7-5)^2}{3} = \frac{10}{3}$$

$$s_A = \sqrt{\frac{10}{3}}$$

$$\overline{x}_B = 5$$

$$s_B^2 = \frac{(1-5)^2 + (3-5)^2 + (7-5)^2 + (9-5)^2}{3} = \frac{40}{3}$$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3,\ x_2 = 4,\ x_3 = 6,\ x_4 = 7$
- Stichprobe B: $x_1 = 1,\ x_2 = 3,\ x_3 = 7,\ x_4 = 9$

$$\overline{x}_A = 5$$

$$s_A^2 = \frac{(3-5)^2 + (4-5)^2 + (6-5)^2 + (7-5)^2}{3} = \frac{10}{3}$$

$$s_A = \sqrt{\frac{10}{3}}$$

$$\overline{x}_B = 5$$

$$s_B^2 = \frac{(1-5)^2 + (3-5)^2 + (7-5)^2 + (9-5)^2}{3} = \frac{40}{3}$$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3,\ x_2 = 4,\ x_3 = 6,\ x_4 = 7$
- Stichprobe B: $x_1 = 1,\ x_2 = 3,\ x_3 = 7,\ x_4 = 9$

$$\overline{x}_A = 5$$

$$s_A^2 = \frac{(3-5)^2 + (4-5)^2 + (6-5)^2 + (7-5)^2}{3} = \frac{10}{3}$$

$$s_A = \sqrt{\frac{10}{3}}$$

$$\overline{x}_B = 5$$

$$s_B^2 = \frac{(1-5)^2 + (3-5)^2 + (7-5)^2 + (9-5)^2}{3} = \frac{40}{3}$$

$$\sqrt{40}$$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3,\ x_2 = 4,\ x_3 = 6,\ x_4 = 7$
- Stichprobe B: $x_1 = 1,\ x_2 = 3,\ x_3 = 7,\ x_4 = 9$

$$\overline{x}_A = 5$$

$$s_A^2 = \frac{(3-5)^2 + (4-5)^2 + (6-5)^2 + (7-5)^2}{3} = \frac{10}{3}$$

$$s_A = \sqrt{\frac{10}{3}}$$

$$\overline{x}_B = 5$$

$$s_B^2 = \frac{(1-5)^2 + (3-5)^2 + (7-5)^2 + (9-5)^2}{3} = \frac{40}{3}$$

$$\sqrt{40} \qquad \sqrt{10}$$

Beispiel

Berechne die empirische Varianz und Standardabweichung:

- Stichprobe A: $x_1 = 3,\ x_2 = 4,\ x_3 = 6,\ x_4 = 7$
- Stichprobe B: $x_1 = 1,\ x_2 = 3,\ x_3 = 7,\ x_4 = 9$

$$\overline{x}_A = 5$$

$$s_A^2 = \frac{(3-5)^2 + (4-5)^2 + (6-5)^2 + (7-5)^2}{3} = \frac{10}{3}$$

$$s_A = \sqrt{\frac{10}{3}}$$

$$\overline{x}_B = 5$$

$$s_B^2 = \frac{(1-5)^2 + (3-5)^2 + (7-5)^2 + (9-5)^2}{3} = \frac{40}{3}$$

Die Variationsbreite (Spannweite oder Range)

Hierbei handelt es sich um die leicht zu berechnende Grösse.

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

Quartile

- Das erste Quartil bezeichnet einen Wert $x_{0.25}$ mit der Eigenschaft, dass ein Viertel der Daten kleiner als $x_{0.25}$ sind. Unsere Berechnungsvorschrift* für das erste Quartil lautet: *Bestimme den Median der Werte unterhalb vom Median.*

Quartile

- Das erste Quartil bezeichnet einen Wert $x_{0.25}$ mit der Eigenschaft, dass ein Viertel der Daten kleiner als $x_{0.25}$ sind. Unsere Berechnungsvorschrift* für das erste Quartil lautet: *Bestimme den Median der Werte unterhalb vom Median.*
- Das dritte Quartil bezeichnet einen Wert $x_{0.75}$ mit der Eigenschaft, dass drei Viertel der Daten kleiner als $x_{0.75}$ sind. Die Berechnungsvorschrift* für das dritte Quartil lautet: *Bestimme den Median der Werte oberhalb vom Median.*

Quartile

- Das erste Quartil bezeichnet einen Wert $x_{0.25}$ mit der Eigenschaft, dass ein Viertel der Daten kleiner als $x_{0.25}$ sind. Unsere Berechnungsvorschrift* für das erste Quartil lautet: *Bestimme den Median der Werte unterhalb vom Median.*
- Das dritte Quartil bezeichnet einen Wert $x_{0.75}$ mit der Eigenschaft, dass drei Viertel der Daten kleiner als $x_{0.75}$ sind. Die Berechnungsvorschrift* für das dritte Quartil lautet: *Bestimme den Median der Werte oberhalb vom Median.*
- Das zweite Quartil $x_{0.5}$ entspricht dem Median $\tilde{x}$.

* Aufgrund der Definition sind die Quartile im Allgemeinen nicht eindeutig bestimmt. Die oben beschriebene Berechnungsmethode wird auch von den TI-84-Taschenrechnerfamilie angewendet.

Interquartilabstand

Der Interquartilabstand (*interquartile range*, IQR) ist die Differenz zwischen dem dritten und dem ersten Quartil:

$$\mathrm{IQR} = x_{0.75} - x_{0.25}$$

Im Gegensatz zur Standardabweichung und der Spannweite ist der Interquartilabstand ein Variabilitätsmass, das robust gegenüber Ausreissern ist.

Beispiel

Wie gross ist der IQR im letzten Beispiel?

Beispiel

Wie gross ist der IQR im letzten Beispiel?

IQR

Beispiel

Wie gross ist der IQR im letzten Beispiel?

$$\mathrm{IQR} = Q_{75\%} - Q_{25\%} =$$

Beispiel

Wie gross ist der IQR im letzten Beispiel?

$$\text{IQR} = Q_{75\%} - Q_{25\%} = 11 -$$

Beispiel

Wie gross ist der IQR im letzten Beispiel?

$$\mathrm{IQR} = Q_{75\%} - Q_{25\%} = 11 - 4.5$$

Beispiel

Wie gross ist der IQR im letzten Beispiel?

$$\mathrm{IQR} = Q_{75\%} - Q_{25\%} = 11 - 4.5 = 6.5$$

Beispiel

In einer Umfrage unter 100 Schülern einer Schule wurde gefragt, welches „Transportmittel“ hauptsächlich für den Schulweg genutzt wird.

Schulweg	**absolute Häufigkeit**	**relative Häufigkeit**	
zu Fuss	6	0.06	(6%)
mit Velo	32	0.32	(32%)
mit Bus	28	0.28	(28%)
mit Zug	19	0.19	(19%)
mit Mofa/Motorrad	14	0.14	(14%)
mit Auto	1	0.01	(1%)
Summe	100	1.00	(100%)

Daraus ergibt sich das einfache Stabdiagramm in Abbildung 1.

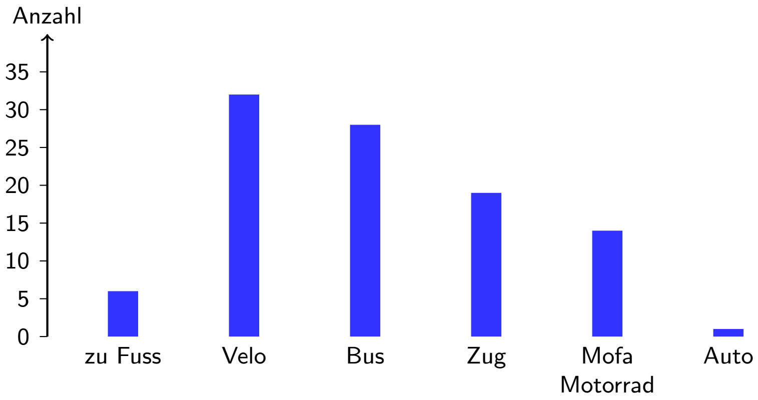


Abbildung 1: Primäres Transportmittel auf dem Schulweg

Die horizontale Darstellungsweise (Balkendiagramm) ist bei wenig Kategorien oder bei langen Kategoriennamen platzsparender (Abbildung 2).

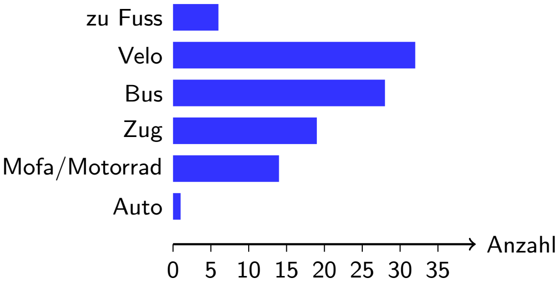


Abbildung 2: Primäres Transportmittel auf dem Schulweg

Wenn man ein Stabdiagramm nach Kategorien aufteilt, entsteht ein gruppiertes Stabdiagramm wie in Abbildung 3.

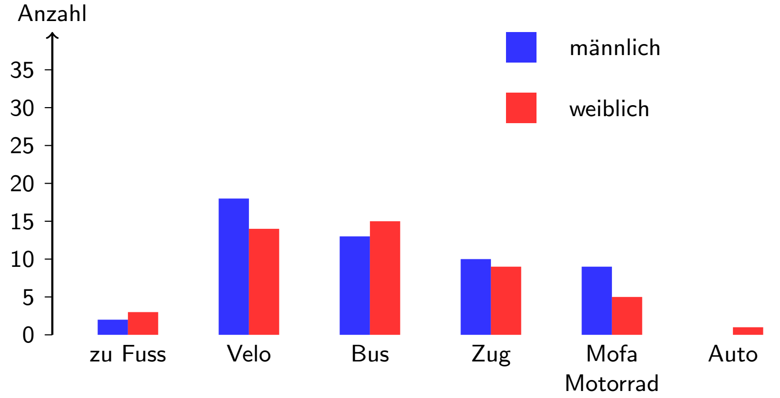


Abbildung 3: Primäres Transportmittel auf dem Schulweg (nach Geschlecht)

Kreisdiagramme wie in Abbildung 4 eignen sich nicht unbedingt für die Darstellung von Informationen, da wir Längenunterschiede besser erkennen können als Differenzen von Kreissektorflächen. Um Monotonie in der Wahl der Grafiken zu vermeiden, kann ein Kreisdiagramm gelegentlich sinnvoll sein.

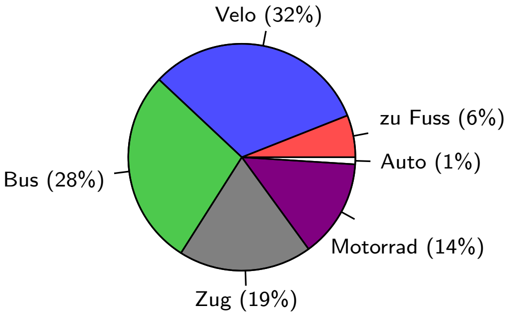
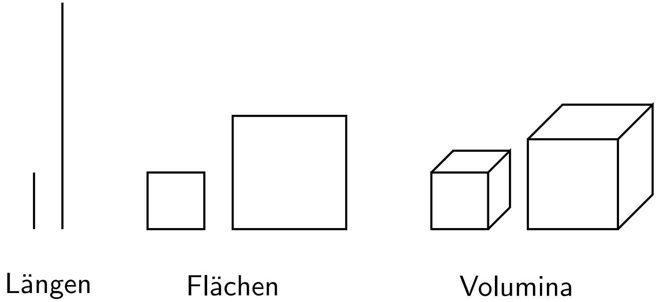


Abbildung 4: Primäres Transportmittel auf dem Schulweg

Finger weg von 3D-Darstellungen!

Das Verhältnis 1 : 4 in verschiedenen Dimensionen



Beispiel

Eine grosse Zahl metrisch skalierter Rohdaten ist intuitiv schlecht zu erfassen.

Anzahl Fehler im Diktat von zwei Klassen:

A: 4, 6, 21, 14, 0, 14, 7, 4, 15, 1, 13, 21, 17, 15, 21, 15

B: 2, 5, 20, 16, 20, 21, 21, 12, 2, 5, 4, 9, 10, 9, 24, 12, 19, 7

Beispiel

Eine grosse Zahl metrisch skalierter Rohdaten ist intuitiv schlecht zu erfassen.

Anzahl Fehler im Diktat von zwei Klassen:

A: 4, 6, 21, 14, 0, 14, 7, 4, 15, 1, 13, 21, 17, 15, 21, 15

B: 2, 5, 20, 16, 20, 21, 21, 12, 2, 5, 4, 9, 10, 9, 24, 12, 19, 7

Ordnungsstatistik:

Beispiel

Eine grosse Zahl metrisch skalierter Rohdaten ist intuitiv schlecht zu erfassen.

Anzahl Fehler im Diktat von zwei Klassen:

A: 4, 6, 21, 14, 0, 14, 7, 4, 15, 1, 13, 21, 17, 15, 21, 15

B: 2, 5, 20, 16, 20, 21, 21, 12, 2, 5, 4, 9, 10, 9, 24, 12, 19, 7

Ordnungsstatistik:

A: 0, 1, 4, 4, 6, 7, 13, 14, 14, 15, 15, 15, 17, 21, 21, 21

B: 2, 2, 4, 5, 5, 7, 9, 9, 10, 12, 12, 16, 19, 20, 20, 21, 21, 24

Um die Verteilungseigenschaften von metrisch skalierten Daten veranschaulichen zu können, werden sie in *Intervalle* eingeteilt.

Um die Verteilungseigenschaften von metrisch skalierten Daten veranschaulichen zu können, werden sie in *Intervalle* eingeteilt.

Dazu einige Faustregeln:

Um die Verteilungseigenschaften von metrisch skalierten Daten veranschaulichen zu können, werden sie in *Intervalle* eingeteilt.

Dazu einige Faustregeln:

- Alle Intervalle sollten die gleiche Breite haben.

Um die Verteilungseigenschaften von metrisch skalierten Daten veranschaulichen zu können, werden sie in *Intervalle* eingeteilt.

Dazu einige Faustregeln:

- Alle Intervalle sollten die gleiche Breite haben.
- Werte, die auf eine Intervallgrenze fallen, werden in der Regel zum darunterliegenden Intervall gezählt.

Um die Verteilungseigenschaften von metrisch skalierten Daten veranschaulichen zu können, werden sie in *Intervalle* eingeteilt.

Dazu einige Faustregeln:

- Alle Intervalle sollten die gleiche Breite haben.
- Werte, die auf eine Intervallgrenze fallen, werden in der Regel zum darunterliegenden Intervall gezählt.
- Maximal 20 Klassen

Tabellarische Darstellung

Die Häufigkeitsverteilung der Diktatfehler (gepoolt):

Intervall	absolute Häufigkeit	relative Häufigkeit
$0 \leq x \leq 5$		
$5 < x \leq 10$		
$10 < x \leq 15$		
$15 < x \leq 20$		
$20 < x \leq 25$		
Summe		

Tabellarische Darstellung

Die Häufigkeitsverteilung der Diktatfehler (gepoolt):

Intervall	**absolute Häufigkeit**	**relative Häufigkeit**
$0 \leq x \leq 5$	9	
$5 < x \leq 10$		
$10 < x \leq 15$		
$15 < x \leq 20$		
$20 < x \leq 25$		
Summe		

Tabellarische Darstellung

Die Häufigkeitsverteilung der Diktatfehler (gepoolt):

Intervall	**absolute Häufigkeit**	**relative Häufigkeit**
$0 \leq x \leq 5$	9	
$5 < x \leq 10$	6	
$10 < x \leq 15$		
$15 < x \leq 20$		
$20 < x \leq 25$		
Summe		

Tabellarische Darstellung

Die Häufigkeitsverteilung der Diktatfehler (gepoolt):

Intervall	absolute Häufigkeit	relative Häufigkeit
$0 \leq x \leq 5$	9	
$5 < x \leq 10$	6	
$10 < x \leq 15$	8	
$15 < x \leq 20$		
$20 < x \leq 25$		
Summe		

Tabellarische Darstellung

Die Häufigkeitsverteilung der Diktatfehler (gepoolt):

Intervall	absolute Häufigkeit	relative Häufigkeit
$0 \leq x \leq 5$	9	
$5 < x \leq 10$	6	
$10 < x \leq 15$	8	
$15 < x \leq 20$	5	
$20 < x \leq 25$		
Summe		

Tabellarische Darstellung

Die Häufigkeitsverteilung der Diktatfehler (gepoolt):

Intervall	**absolute Häufigkeit**	**relative Häufigkeit**
$0 \leq x \leq 5$	9	
$5 < x \leq 10$	6	
$10 < x \leq 15$	8	
$15 < x \leq 20$	5	
$20 < x \leq 25$	6	
Summe		

Tabellarische Darstellung

Die Häufigkeitsverteilung der Diktatfehler (gepoolt):

Intervall	absolute Häufigkeit	relative Häufigkeit
$0 \leq x \leq 5$	9	
$5 < x \leq 10$	6	
$10 < x \leq 15$	8	
$15 < x \leq 20$	5	
$20 < x \leq 25$	6	
Summe	34	

Tabellarische Darstellung

Die Häufigkeitsverteilung der Diktatfehler (gepoolt):

Intervall	absolute Häufigkeit	relative Häufigkeit	
$0 \leq x \leq 5$	9	0.265	(26.5%)
$5 < x \leq 10$	6		
$10 < x \leq 15$	8		
$15 < x \leq 20$	5		
$20 < x \leq 25$	6		
Summe	34		

Tabellarische Darstellung

Die Häufigkeitsverteilung der Diktatfehler (gepoolt):

Intervall	absolute Häufigkeit	relative Häufigkeit	
$0 \leq x \leq 5$	9	0.265	(26.5%)
$5 < x \leq 10$	6	0.176	(17.6%)
$10 < x \leq 15$	8		
$15 < x \leq 20$	5		
$20 < x \leq 25$	6		
Summe	34		

Tabellarische Darstellung

Die Häufigkeitsverteilung der Diktatfehler (gepoolt):

Intervall	absolute Häufigkeit	relative Häufigkeit	
$0 \leq x \leq 5$	9	0.265	(26.5%)
$5 < x \leq 10$	6	0.176	(17.6%)
$10 < x \leq 15$	8	0.235	(23.5%)
$15 < x \leq 20$	5		
$20 < x \leq 25$	6		
Summe	34		

Tabellarische Darstellung

Die Häufigkeitsverteilung der Diktatfehler (gepoolt):

Intervall	**absolute Häufigkeit**	**relative Häufigkeit**	
$0 \leq x \leq 5$	9	0.265	(26.5%)
$5 < x \leq 10$	6	0.176	(17.6%)
$10 < x \leq 15$	8	0.235	(23.5%)
$15 < x \leq 20$	5	0.147	(14.7%)
$20 < x \leq 25$	6		
Summe	34		

Tabellarische Darstellung

Die Häufigkeitsverteilung der Diktatfehler (gepoolt):

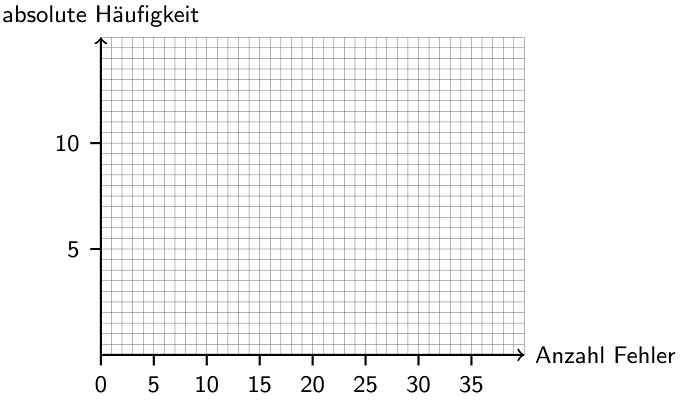
Intervall	**absolute Häufigkeit**	**relative Häufigkeit**	
$0 \leq x \leq 5$	9	0.265	(26.5%)
$5 < x \leq 10$	6	0.176	(17.6%)
$10 < x \leq 15$	8	0.235	(23.5%)
$15 < x \leq 20$	5	0.147	(14.7%)
$20 < x \leq 25$	6	0.176	(17.6%)
Summe	34		

Tabellarische Darstellung

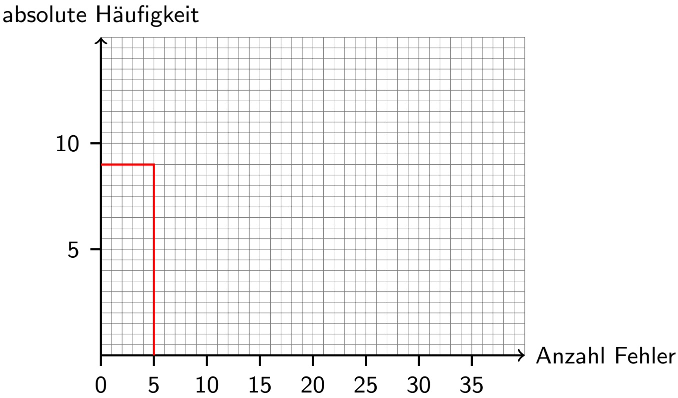
Die Häufigkeitsverteilung der Diktatfehler (gepoolt):

Intervall	absolute Häufigkeit	relative Häufigkeit	
$0 \leq x \leq 5$	9	0.265	(26.5%)
$5 < x \leq 10$	6	0.176	(17.6%)
$10 < x \leq 15$	8	0.235	(23.5%)
$15 < x \leq 20$	5	0.147	(14.7%)
$20 < x \leq 25$	6	0.176	(17.6%)
Summe	34	1.000	(100%)

Das Histogramm

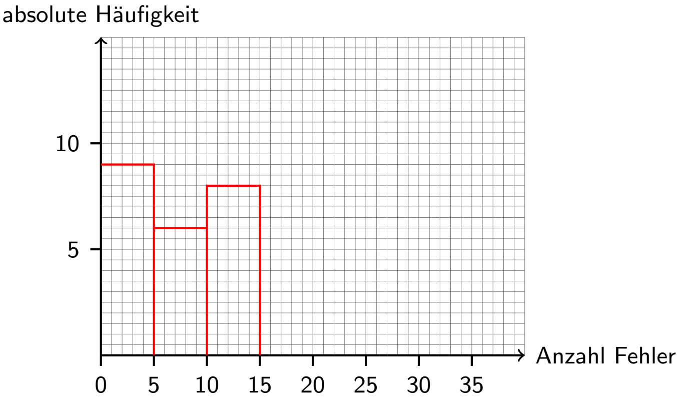


Das Histogramm

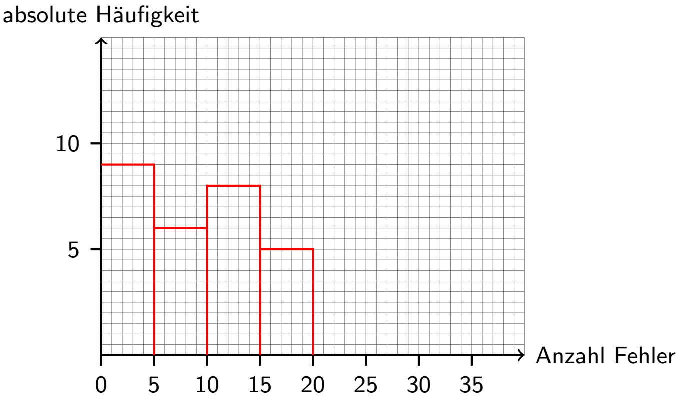


Das Histogramm

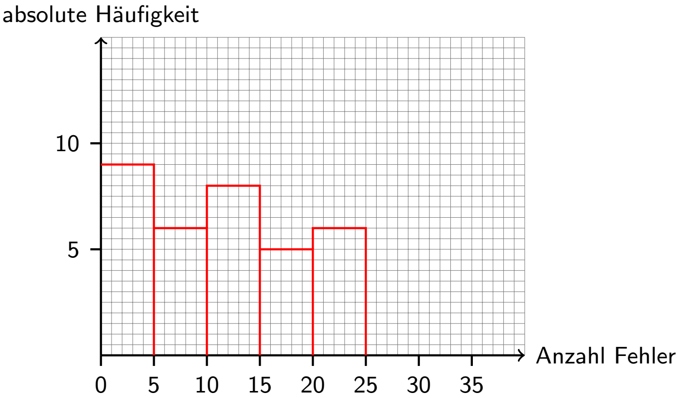
Das Histogramm



Das Histogramm



Das Histogramm



Median und Quartile

	x_{min}	$x_{0.25}$	$\tilde{x}$	$x_{0.75}$	x_{max}	IQR
Klasse A						
Klasse B						

Median und Quartile

	$x_{\min}$	$x_{0.25}$	$\tilde{x}$	$x_{0.75}$	$x_{\max}$	IQR
Klasse A	0					
Klasse B						

Median und Quartile

	$x_{\min}$	$x_{0.25}$	$\tilde{x}$	$x_{0.75}$	$x_{\max}$	IQR
Klasse A	0	5				
Klasse B						

Median und Quartile

	$x_{\min}$	$x_{0.25}$	$\tilde{x}$	$x_{0.75}$	$x_{\max}$	IQR
Klasse A	0	5	14			
Klasse B						

Median und Quartile

	$x_{\min}$	$x_{0.25}$	$\tilde{x}$	$x_{0.75}$	$x_{\max}$	IQR
Klasse A	0	5	14	16		
Klasse B						

Median und Quartile

	$x_{\min}$	$x_{0.25}$	$\tilde{x}$	$x_{0.75}$	$x_{\max}$	IQR
Klasse A	0	5	14	16	21	
Klasse B						

Median und Quartile

	x_{min}	$x_{0.25}$	$\tilde{x}$	$x_{0.75}$	x_{max}	IQR
Klasse A	0	5	14	16	21	11
Klasse B						

Median und Quartile

	$x_{\min}$	$x_{0.25}$	$\tilde{x}$	$x_{0.75}$	$x_{\max}$	IQR
Klasse A	0	5	14	16	21	11
Klasse B	2					

Median und Quartile

	$x_{\min}$	$x_{0.25}$	$\tilde{x}$	$x_{0.75}$	$x_{\max}$	IQR
Klasse A	0	5	14	16	21	11
Klasse B	2	5				

Median und Quartile

	$x_{\min}$	$x_{0.25}$	$\tilde{x}$	$x_{0.75}$	$x_{\max}$	IQR
Klasse A	0	5	14	16	21	11
Klasse B	2	5	11			

Median und Quartile

	$x_{\min}$	$x_{0.25}$	$\tilde{x}$	$x_{0.75}$	$x_{\max}$	IQR
Klasse A	0	5	14	16	21	11
Klasse B	2	5	11	20		

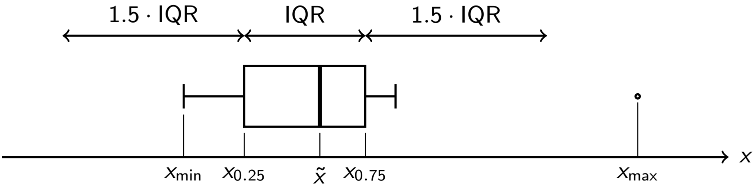
Median und Quartile

	$x_{\min}$	$x_{0.25}$	$\tilde{x}$	$x_{0.75}$	$x_{\max}$	IQR
Klasse A	0	5	14	16	21	11
Klasse B	2	5	11	20	24	

Median und Quartile

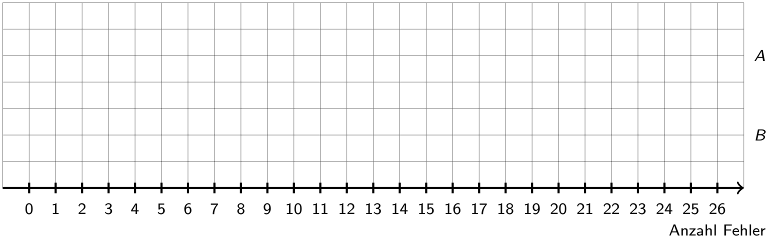
	x_{min}	$x_{0.25}$	$\tilde{x}$	$x_{0.75}$	x_{max}	IQR
Klasse A	0	5	14	16	21	11
Klasse B	2	5	11	20	24	15

Das Box-and-Whiskers Plot

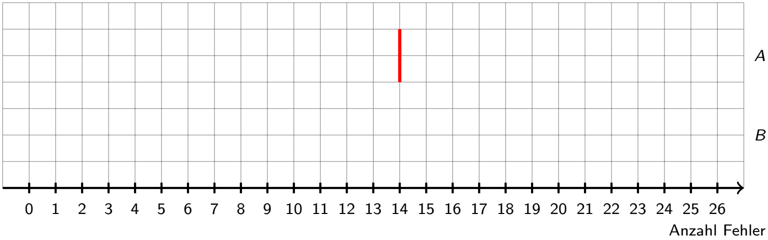


Werte, die kleiner als $x_{0.25} - 1.5 \cdot \mathrm{IQR}$ oder grösser als $x_{0.75} + 1.5 \cdot \mathrm{IQR}$ sind, werden als Ausreisser bezeichnet und als separate Punkte dargestellt.

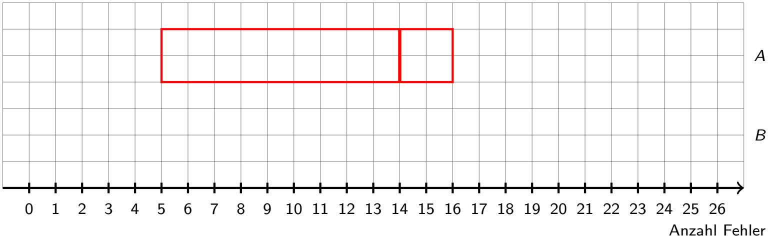
Das Box-and-Whiskers Plot der Beispieldaten



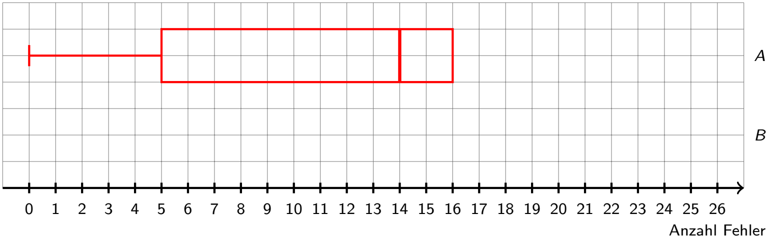
Das Box-and-Whiskers Plot der Beispieldaten



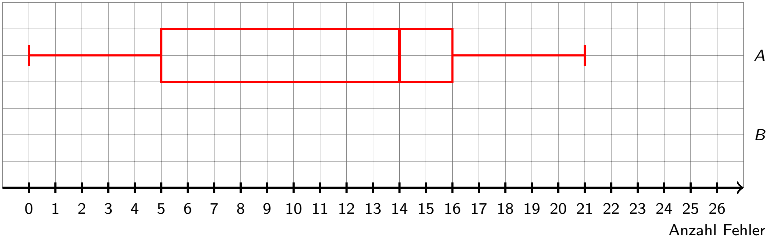
Das Box-and-Whiskers Plot der Beispieldaten



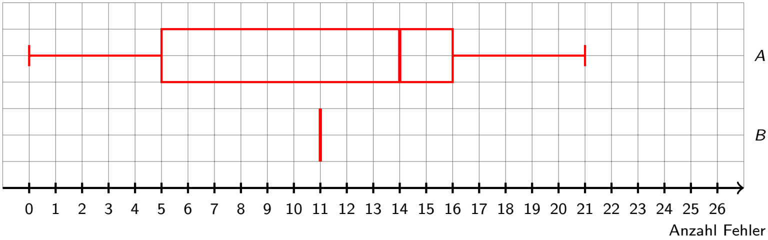
Das Box-and-Whiskers Plot der Beispieldaten



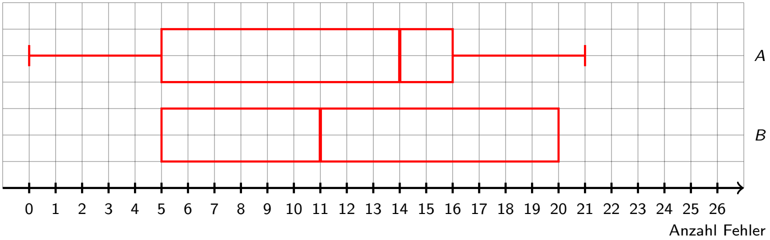
Das Box-and-Whiskers Plot der Beispieldaten



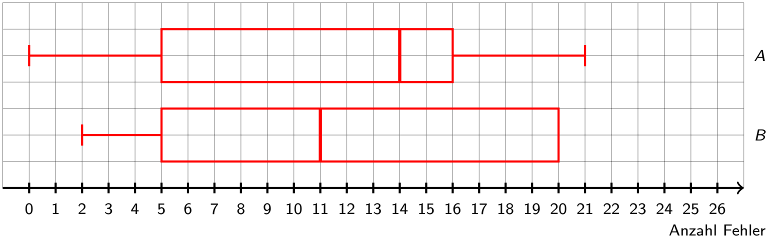
Das Box-and-Whiskers Plot der Beispieldaten



Das Box-and-Whiskers Plot der Beispieldaten



Das Box-and-Whiskers Plot der Beispieldaten



Das Box-and-Whiskers Plot der Beispieldaten