

LU01b - Stichprobe, Rangliste, Häufigkeitstabelle und Darstellung



Ziel: Du kannst Rohdaten (Urliste) **ordnen** und **verdichten** (geordnete Stichprobe, Strichliste, Häufigkeitstabelle). Du kannst eine **Rangliste** erstellen (inkl. Durchschnittsrang bei Mehrfachwerten) und den **Modus** bestimmen. Du kannst Daten mit **Säulen- und Kreisdiagramm** darstellen.

1) Begriffe: Urliste, Stichprobe, Verdichtung

Urliste (Rohdaten)

Eine **Urliste** enthält die erhobenen Werte **in der Reihenfolge**, in der sie gemessen/abgefragt wurden. Sie ist oft unübersichtlich – deshalb „verdichten“ wir die Daten.

Beispiel (Informatik): „Welchen Browser nutzt du am meisten?“ Urliste (24 Lernende): Chrome, Edge, Chrome, Firefox, Chrome, Safari, Chrome, Edge, Chrome, Firefox, Chrome, Edge, Chrome, Chrome, Firefox, Edge, Chrome, Safari, Edge, Chrome, Firefox, Edge, Chrome, Chrome

Fazit: So kann man schlecht erkennen, was häufig ist → wir ordnen/verdichten.

Stichprobe und Stichprobenumfang n

- **Stichprobe:** die Menge der erhobenen Daten (z.B. „unsere Klasse heute“)
- **Stichprobenumfang n:** Anzahl Datenpunkte (z.B. 24 Antworten)

2) Datenarten

In den Aufgaben wird oft so unterschieden:

- **nominal:** Kategorien **ohne** Reihenfolge (z.B. Nationalität, Zivilstand, Lieblingsgetränk)
- **ordinal:** Kategorien **mit** Reihenfolge (z.B. Energieklassen, Bewertungen „genügend-gut-sehr gut“)
- **kardinal:** Zahlenwerte (du kannst zählen/messen) → entspricht **quantitativ**

(dazu gehören z.B. Alter, Lohn, Länge, Gewicht, Besucherzahlen)

Achtung: Zahlen können auch **Codes** sein (z.B. HTTP-Statuscode 404). Dann ist es meist **nominal**, weil es ein Label ist.

3) Geordnete Stichprobe (sortieren)

Bei **numerischen** Daten sortierst du die Werte vom kleinsten zum grössten Wert. Dann schreibst du sie oft als: $x(1), x(2), x(3), \dots, x(n)$

Beispiel (Informatik): Antwortzeiten in ms (Urliste) 120 / 95 / 95 / 140 / 110 / 120

Geordnete Stichprobe: $x(1)=95, x(2)=95, x(3)=110, x(4)=120, x(5)=120, x(6)=140$

4) Rangliste (Rangwertfolge)

Eine **Rangliste** ordnet jedem Stichprobenwert seinen **Rang** zu (Position in der sortierten Liste).

Fall A: jeder Wert kommt nur einmal vor

Dann ist der Rang eindeutig:

- $x(1)$ hat Rang 1
- $x(2)$ hat Rang 2
- ...
- $x(n)$ hat Rang n

Fall B: Werte kommen mehrfach vor (Tie)

Dann bekommen alle gleichen Werte den **Durchschnittsrang**.

Regel (Durchschnittsrang): Wenn ein Wert in der sortierten Liste die Ränge von **k bis l** belegen würde, dann erhält jeder dieser gleichen Werte den Rang:

- $(k + l) / 2$

Beispiel (Antwortzeiten): Geordnet: 95, 95, 110, 120, 120, 140

Wert $x(i)$	„eigentliche“ Ränge	zugeteilter Rang
95	1 und 2	$(1+2)/2 = 1.5$
110	3	3
120	4 und 5	$(4+5)/2 = 4.5$
140	6	6

5) Strichliste (Zählen)

Eine **Strichliste** zählt, wie oft jede Ausprägung vorkommt (besonders nützlich bei Kategorien).

Beispiel (Browser):

- Chrome: ||||||||| (12)
- Edge: ||||| (6)
- Firefox: |||| (4)
- Safari: || (2)

6) Häufigkeitstabelle

Die Häufigkeitstabelle zeigt:

- n_i = absolute Häufigkeit (wie oft kommt Kategorie/Wert i vor?)
- h_i = relative Häufigkeit (Anteil an n)

Formeln:

- $h_i = n_i / n$
- Prozent = $h_i * 100$

Aufbau (typisch)

i	x_i (Ausprägung/Wert)	n_i (absolut)	h_i (relativ)	Prozent
1

Kontrolle:

- Summe aller $n_i = n$
- Summe aller $h_i = 1$ (oder 100%)

Beispiel (Browser, $n=24$)

i	x_i	n_i	h_i	Prozent
1	Chrome	12	$12/24 = 0.50$	50%
2	Edge	6	$6/24 = 0.25$	25%
3	Firefox	4	$4/24 = 0.1666...$	16.67%
4	Safari	2	$2/24 = 0.0833...$	8.33%
	Summe	24	1.00	100%

7) Modus

Der **Modus** ist die Ausprägung, die **am häufigsten** vorkommt.

Beispiel (Browser): Chrome hat $n_i = 12 \rightarrow \text{Modus} = \text{Chrome}$

Hinweis:

- Es kann auch **mehrere** Modi geben (wenn mehrere Kategorien gleich oft am häufigsten sind).

8) Grafische Darstellung

Kategorische Daten kann man häufig mit **Säulen-** oder **Kreisdiagramm** darstellen.

8.1 Säulendiagramm (Balken/Säulen)

- Für jede Ausprägung eine Säule
- Alle Säulen gleich breit
- Höhe = **absolute** oder **relative** Häufigkeit

Vorteil: Der Modus ist sofort sichtbar (höchste Säule).

8.2 Kreisdiagramm

- Für jede Ausprägung ein Kreissegment
- Segmentgröße ist proportional zur **relativen Häufigkeit**

Standard-Vorgehen:

- erstes Segment bei „12 Uhr“
- Segmente im Uhrzeigersinn

Winkelberechnung:

- $\text{Winkel}_i = h_i * 360^\circ$
- oder: $\text{Winkel}_i = \text{Prozent}_i * 3.6^\circ$

Beispiel (Browser):

- Chrome: $0.50 * 360^\circ = 180^\circ$
- Edge: $0.25 * 360^\circ = 90^\circ$
- Firefox: $0.1666... * 360^\circ = 60^\circ$
- Safari: $0.0833... * 360^\circ = 30^\circ$

Fazit: Für Vergleiche ist das **Säulendiagramm** meistens besser als das Kreisdiagramm, weil man Säulenhöhen schneller vergleichen kann als Flächen.

1. Was ist der Unterschied zwischen Urliste und geordneter Stichprobe?
2. Wie berechnest du h_i aus n_i und n ?
3. Was ist der Modus?
4. Wie berechnest du den Winkel im Kreisdiagramm?
5. Wie bestimmst du den Rang bei einem Wert, der mehrfach vorkommt?



M3-LU01



Kevin Maurizi

From:

<https://wiki.bzz.ch/> - **BZZ - Modulwiki**

Permanent link:

<https://wiki.bzz.ch/modul/mathe/ma4/thema/statistik/ordnenvondaten>

Last update: **2026/02/02 07:50**

