

LU01a - Arten von Daten



Ziel: Du kannst Variablen korrekt als **qualitativ/quantitativ** einordnen und das **Messniveau** (nominal/ordinal/intervall/ratio) bestimmen. Falls quantitativ: **diskret/stetig**.

Warum ist das wichtig?

Bevor du Kennzahlen (Mittelwert, Median, Standardabweichung) berechnest oder Diagramme wählst, musst du wissen, **welche Art von Daten** du hast. Die Datenart bestimmt:

- welche Rechnungen **sinnvoll** sind (z.B. Mittelwert ja/nein)
- welche Darstellungen passen (Balken, Histogramm, Boxplot)
- welche Aussagen du machen darfst (z.B. „doppelt so viel“)

1) Zwei Hauptarten von Daten

1.1 Qualitative Daten (kategorisch)

Qualitative Daten sind **Kategorien/Labels**. Die Werte sind Namen/Typen und keine Messzahlen.

Informatik-Beispiele (qualitativ)

- Betriebssystem: Windows / macOS / Linux
- Team-Rolle: Dev / Ops / QA
- Build-Status: success / failed
- Browser: Chrome / Firefox / Safari

Typische Auswertungen:

- **Häufigkeiten** (wie oft kommt etwas vor?)
- **Modus** (welche Kategorie ist am häufigsten?)

1.2 Quantitative Daten (numerisch)

Quantitative Daten sind **Zahlenwerte**, die etwas **zählen oder messen**.

Informatik-Beispiele (quantitativ)

- Anzahl Commits pro Woche (gezählt)
- Anzahl offene Tickets im Sprint (gezählt)
- Ladezeit einer Webseite in ms (gemessen)

- Dateigrösse in MB (gemessen)
- CPU-Temperatur in °C (gemessen)

Typische Auswertungen:

- **Mittelwert, Median**
- Streuung (z.B. Standardabweichung, später)

Merksatz: **Qualitativ** = Kategorien/Labels. **Quantitativ** = Zahlen zum Zählen/Messen.

2) Messniveau (Skalenniveau): nominal / ordinal / intervall / ratio

Das **Messniveau** sagt dir, **welche Vergleiche** und **welche Rechnungen** sinnvoll sind.

2.1 Nominal (nur gleich/ungleich)

- **keine Reihenfolge**
- sinnvoll: Häufigkeiten, Modus, Balkendiagramm
- nicht sinnvoll: Median, Mittelwert

Informatik-Beispiele

- Betriebssystem: Windows / macOS / Linux
- Team-Rolle: Dev / Ops / QA
- HTTP-Methode: GET / POST / PUT
- HTTP-Statuscode: 200 / 404 / 500 (**als Kategorie**)

2.2 Ordinal (Reihenfolge, Abstände unklar)

- **Reihenfolge** ist vorhanden (höher/tiefer)
- Abstände sind nicht eindeutig gleich gross
- sinnvoll: Median, Rangvergleiche, (Häufigkeiten)
- Mittelwert: nur mit Vorsicht

Informatik-Beispiele

- Zufriedenheit 1-5 (sehr schlecht ... sehr gut)
- Priorität: low / medium / high
- Severity: minor / major / critical

2.3 Intervall (Abstände sinnvoll, Nullpunkt willkürlich)

- Differenzen sind sinnvoll (z.B. +10)
- Nullpunkt bedeutet nicht „nichts“
- sinnvoll: Mittelwert, Standardabweichung, Histogramm
- Verhältnisse („doppelt so viel“) sind **nicht** sinnvoll

Informatik-Beispiele

- CPU-Temperatur in °C (0°C ist nicht „keine Temperatur“)
- Uhrzeit als Tageszeit (z.B. 14:00) (Abstände ok, aber „doppelt so spät“ ergibt keinen Sinn)

2.4 Ratio (Abstände + echtes Null)

- Null bedeutet „nichts“
- Verhältnisse sind sinnvoll („doppelt so viele“, „halb so gross“)
- sinnvoll: praktisch alle Kennzahlen & Diagramme

Informatik-Beispiele

- Ladezeit in ms (0 ms = keine Zeit)
- Dateigrösse in MB (0 MB = keine Daten)
- Anzahl Commits/Tickets (0 = nichts)

Quick-Check: Wenn **0** wirklich „nichts“ bedeutet → meistens **Ratio**. Wenn Werte **nur Codes** sind → meistens **Nominal**.

3) Diskret vs. Stetig (nur bei quantitativen Daten)

3.1 Diskret (gezählt)

- Werte sind zählbar, typischerweise ganze Zahlen: 0, 1, 2, 3, ...
- entsteht durch **Zählen**

Informatik-Beispiele

- Anzahl Commits pro Woche
- Anzahl offene Tickets
- Anzahl Logins pro Tag

3.2 Stetig (gemessen)

- Werte können theoretisch beliebig fein sein (z.B. 12.3, 12.31, 12.312, ...)
- entsteht durch **Messen**

Informatik-Beispiele

- Ladezeit in ms
- Downloadrate in Mbit/s
- CPU-Temperatur

4) Typische Stolpersteine

Achtung: Nur weil etwas wie eine Zahl aussieht, ist es nicht automatisch quantitativ.

HTTP-Statuscodes (200/404/500) sind meistens **Kategorien** → **nominal**.

Weitere häufige Fehler:

- **Ordinal = nicht automatisch Mittelwert!** (Zufriedenheit 1–5: Median ist oft besser)
- **Intervall vs. Ratio** verwechseln: °C (Intervall) vs. ms/MB (Ratio)
- **Diskret vs. Stetig:** Zählen (diskret) vs. Messen (stetig)

6) Mini-Check

1. Nenne ein Beispiel für **nominale** Daten aus der Informatik.
2. Wieso ist **HTTP-Statuscode** meistens **nominal**?
3. Ist „Anzahl Commits“ eher **diskret** oder **stetig**? Begründe kurz.

M3-LU01



Kevin Maurizi

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

Permanent link:
<https://wiki.bzz.ch/modul/mathe/ma4/thema/statistik/aufgaben/klassifizieren?rev=1769417819>

Last update: **2026/01/26 09:56**

