

LU01a - Arten von Daten

Überblick	Inhalt
Ziel	Du kannst Datenarten unterscheiden (qualitativ/quantitativ) und die passende Skala (nominal/ordinal/intervall/ratio) erklären.
Warum wichtig?	Die Datenart entscheidet, welche Kennzahlen (z.B. Mittelwert ja/nein) und welche Diagramme sinnvoll sind.
Kompetenzbezug	AG1/AF1 (Grundbegriffe verstehen & anwenden), Vorbereitung für Lagewerte/Verteilungen
Dauer	1 Doppellektion (90')

1. Einstieg: Was bedeutet „Daten“?

In der Datenanalyse beschreibt man reale Beobachtungen in strukturierter Form, um daraus Erkenntnisse abzuleiten.

Beispiele aus dem Informatik-Alltag

- Ladezeit einer Webseite in Millisekunden
- Anzahl Fehler in einem Log pro Stunde
- Betriebssystem der User (Windows/macOS/Linux)
- Zufriedenheit der User (Skala 1-5)
- HTTP-Statuscodes (200, 404, 500)

Wichtig: Bevor du Kennzahlen berechnest oder Diagramme zeichnest, musst du wissen: **Welche Art von Daten ist das?**

2. Grundbegriffe (kurz)

Begriff	Bedeutung	Beispiel
Merkmal / Variable	Eigenschaft, die beobachtet wird	„Browser-Typ“
Ausprägung	konkreter Wert einer Variable	„Firefox“
Beobachtung / Datensatz	eine Messung/Eintrag	„Firefox“ bei Person A
Stichprobe	Teilmenge der Daten	25 Lernende
Grundgesamtheit	alle möglichen Fälle	alle Lernenden der Schule

3. Zwei Hauptklassen von Daten

3.1 Qualitative Daten (kategorisch)

Qualitative Daten sind **Kategorien/Labels**. Man rechnet damit nicht „normal“ (kein Mittelwert).

Beispiele

- Betriebssystem: Windows / macOS / Linux
- Sprache: DE / EN / FR

- Bug-Typ: UI / Backend / Security

3.2 Quantitative Daten (numerisch)

Quantitative Daten sind **Zahlenwerte**, die etwas **messen oder zählen**. Hier sind mathematische Operationen sinnvoll (je nach Skala).

Beispiele

- Ladezeit: 123 ms
- RAM-Verbrauch: 2.4 GB
- Anzahl Commits: 18

Merksatz: **Qualitativ** = Kategorien/Labels. **Quantitativ** = Zahlenwerte (zählen oder messen).

4. Skalentypen (Messniveaus)

Die Skala sagt dir, **was du mit den Daten machen darfst**.

4.1 Nominalskala (nur „gleich/ungleich“)

- **Keine Reihenfolge**
- Erlaubt: Häufigkeiten, Modus, Balkendiagramm, Kreisdiagramm
- Nicht sinnvoll: Median, Mittelwert

Beispiele (Informatik)

- Browser: Chrome / Firefox / Safari
- HTTP-Methode: GET / POST / PUT
- Statuscode **als Kategorie**: 200 / 404 / 500

→ Achtung: Das sind Zahlen, aber sie funktionieren hier als **Labels**, nicht als Messwert!

4.2 Ordinalskala (Reihenfolge, aber Abstände unbekannt)

- **Rangfolge** ist möglich
- Abstände zwischen Stufen sind nicht zwingend gleich
- Erlaubt: Median, Modus, Rangvergleiche, Boxplot (oft ok), Balkendiagramm
- Mittelwert: meist **nicht** sinnvoll (nur mit Vorsicht)

Beispiele

- Zufriedenheit: 1-5 (sehr schlecht ... sehr gut)
- Priorität: low / medium / high
- Schulnoten (je nach System): Rangordnung vorhanden

4.3 Intervallskala (Abstände sinnvoll, aber kein „echtes Null“)

- Differenzen sind sinnvoll (z.B. +10)
- Nullpunkt ist **willkürlich**
- Erlaubt: Mittelwert, Standardabweichung, Histogramm, Linie
- Verhältnis („doppelt so viel“) ist **nicht** sinnvoll

Beispiele

- Temperatur in °C (0°C ist nicht „keine Temperatur“)
- Kalenderjahr (2026 ist nicht „doppelt so viel“ wie 1013)

4.4 Ratioskala / Verhältnisskala (Abstände + echtes Null)

- Abstände sinnvoll **und** Null bedeutet „nichts“
- Verhältnis ist sinnvoll („doppelt so viel“)
- Erlaubt: alle üblichen Kennzahlen + viele Diagramme

Beispiele (Informatik)

- Dateigröße (0 MB = keine Daten)
- Laufzeit/Antwortzeit (0 ms = keine Zeit)
- Anzahl Requests (0 = kein Request)

Quick-Check: Wenn „0“ wirklich „nichts“ bedeutet → meist **Ratioskala**. Wenn Zahlen nur Codes sind (z.B. 404) → meist **Nominalskala**.

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

5.1 Diskrete Daten (zählbar, ganze Werte)

- entstehen durch **Zählen**
- Werte sind typischerweise ganze Zahlen

Beispiele

- Anzahl Commits pro Woche
- Anzahl Fehler pro Build
- Anzahl User in einem Chatraum

5.2 Stetige Daten (messbar, beliebig fein)

- entstehen durch **Messen**
- theoretisch unendlich viele Zwischenwerte möglich

Beispiele

- Ladezeit in ms (in der Realität gemessen, kann sehr fein sein)
- CPU-Temperatur
- Netzwerk-Latenz

6. Welche Kennzahlen und Diagramme passen?

Datenart / Skala	Sinnvolle Kennzahlen	Typische Diagramme
Nominal (qualitativ)	absolute/relative Häufigkeit, Modus	Balken, Kreis
Ordinal (qualitativ)	Median, Modus, Quartile (oft ok)	Balken, (Boxplot oft ok)
Intervall (quantitativ)	Mittelwert, Median, Stdabw.	Histogramm, Linie, Boxplot
Ratio (quantitativ)	Mittelwert, Median, Stdabw., Verhältnisse	Histogramm, Linie, Boxplot
Diskret	Mittelwert/Median möglich (je nach Skala)	Balken, Histogramm (mit ganzen Klassen)
Stetig	Mittelwert/Median/Stdabw.	Histogramm, Dichtekurve, Boxplot

7. Beispiele (Informatik-Kontext)

Beispiel 1: HTTP-Statuscodes

Frage: Ist 500 „mehr“ als 200? **Antwort:** Als Zahl ja – aber als Bedeutung nicht. Es sind **Kategorien**.

- Datenklasse: **qualitativ**
- Skala: **nominal**
- Sinnvolle Auswertung: Häufigkeiten pro Code, Top-3 Codes, Balkendiagramm

Beispiel 2: Zufriedenheit 1-5

- Datenklasse: **qualitativ** (Ordinal)
- Skala: **ordinal**
- Sinnvolle Auswertung: Median, Häufigkeiten, Verteilung als Balken

Interpretation: Median = 4 bedeutet „typisch eher gut“, aber „4 ist doppelt so gut wie 2“ ist nicht zwingend korrekt.

Beispiel 3: Ladezeit (ms)

- Datenklasse: **quantitativ**
- Skala: **Ratio**
- Typ: **stetig**
- Sinnvolle Auswertung: Mittelwert/Median/Stdabw., Boxplot (Ausreisser!), Histogramm

8. Übung: Datenarten klassifizieren

Auftrag: Bestimme für jede Variable:

1. qualitativ oder quantitativ?
2. Skala (nominal/ordinal/intervall/ratio)?
3. falls quantitativ: diskret oder stetig?

Variable	qualit./quant.	Skala	diskret/stetig
Betriebssystem (Windows/macOS/Linux)			
Anzahl Commits pro Woche			
Ladezeit einer Seite (ms)			
HTTP-Statuscode (200/404/500)			
Zufriedenheit (1-5)			
CPU-Temperatur in °C			
Dateigrösse in MB			

Musterlösung (kurz)

- Betriebssystem → qualitativ, nominal
- Commits → quantitativ, ratio, diskret
- Ladezeit → quantitativ, ratio, stetig
- Statuscode → qualitativ, nominal
- Zufriedenheit → qualitativ, ordinal
- CPU-Temperatur °C → quantitativ, intervall, stetig
- Dateigrösse MB → quantitativ, ratio, stetig

9. Häufige Fehler

- „Zahlen = automatisch quantitativ“ → falsch. Codes (z.B. 404) sind oft **nominal**.
- **Mittelwert bei Ordinaldaten** (z.B. Zufriedenheit) → kann täuschen.
- **Diskret/stetig** verwechseln: Zählen = diskret, Messen = stetig.
- Diagramm falsch gewählt: Kreisdiagramm für sehr viele Kategorien ist unübersichtlich → besser Balken.

10. Mini-Check (Exit Ticket)

1. Nenne ein Beispiel für **nominale** Daten aus der Informatik.
2. Warum ist der **HTTP-Statuscode** meistens **nicht** quantitativ?
3. Was ist der Unterschied zwischen **diskret** und **stetig**?

M3-LU01



Kevin Maurizi

From:

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

Permanent link:

<https://wiki.bzz.ch/modul/mathe/ma4/thema/statistik/artenvondaten?rev=1769417092>

Last update: **2026/01/26 09:44**

