

**Darum geht es:**

Daten sind Informationen aus der Vergangenheit und Gegenwart, die durch Beobachtungen, Befragungen, Recherchen oder als Ergebnis von Messungen gewonnen werden. Dabei wird eine bestimmte Menge von Objekten und Individuen (die *Grundgesamtheit*) bezüglich verschiedener Merkmale untersucht. Vor Beginn der Datenerhebung ist zu klären, welche Elemente (oder *Merkmalsträger*) zur Grundgesamtheit gehören, welches Merkmal untersucht werden soll und welche Fragen geeignet und auswertbar sind.

Merkmale und ihre Ausprägungen unterscheiden sich in der Art ihrer Skalierung. Je nach Art der Merkmale stehen Nominalskalen, Ordinalskalen oder Kardinalskalen zur Verfügung.

Skala	Merkmalsausprägungen können...	Beispiel für Merkmale
Nominalskala	unterschieden werden	Lieblingssportart, Farbe, Automarken
Ordinalskala	geordnet werden	Schulnoten, Schwierigkeitsgrade, Platzierungen
Kardinalskala	im Abstand oder Verhältnis verglichen werden	Körpergröße in cm, gemessene Größen

Merkmalsausprägungen können stets in (ungeordneten) Urlisten und als Strichlisten erfasst werden.

Liegt ein ordinal oder kardinal skaliertes Merkmal vor, so kann die Urliste auch zu einer Rangliste geordnet werden. Daraus lassen sich leicht Häufigkeitstabellen erstellen, die die absoluten Häufigkeiten von Merkmalsausprägungen erfassen und mit deren Hilfe man die relativen Häufigkeiten ermitteln kann. Beim Übergang von der Urliste zur Strichliste bzw. zur Häufigkeitstabelle wird eine Reduktion der Informationen vorgenommen. Die Zuordnung der einzelnen Merkmalsausprägungen zu den einzelnen Merkmalsträgern ist nicht mehr erkennbar.

Zur Auswertung von Daten ist es häufig notwendig, verschiedene Merkmalsausprägungen zusammenzufassen. Diese Klassenbildung ist ein wichtiger Prozess, der die Übersichtlichkeit und Aussagekraft der Darstellung bestimmt.

Die relative Häufigkeit wird dazu genutzt, verschieden große Datenmengen oder Stichproben miteinander zu vergleichen. Sie ist zudem die Grundlage für die Ermittlung von statistischen Wahrscheinlichkeiten.

Für die Veranschaulichung von Daten steht eine Vielzahl verschiedener Diagrammtypen zur Verfügung, z. B. Streifen-, Säulen-, Balken-, Kreisdiagramme oder Boxplots. Die Art der gesammelten Daten hat Einfluss auf die Auswahl eines geeigneten Diagrammtyps und ggf. die Einteilung der Achsen.

Die Reduktion, Darstellung und Auswertung von Daten wird durch die Bestimmung statistischer Kennwerte erleichtert. Je nach Skalierung können verschiedene Kennwerte bestimmt werden:

- Nominalskala – seltenster Wert, häufigster Wert (Modalwert)
- Ordinalskala – kleinstes Beobachtungsergebnis (Minimum), größtes Beobachtungsergebnis (Maximum), Zentralwert (Median), Viertelwert (Quartil)
- Kardinalskala – Durchschnitt (arithmetisches Mittel), Spannweite (Differenz von Maximum und Minimum), Quartilsabstand

Mithilfe ausgewählter Kennwerte können vergleichende oder beurteilende Aussagen zu Datenerhebungen getroffen werden. Dazu gehören auch die Prüfung der Plausibilität von Daten und die besondere Beachtung von Ausreißern.

Das arithmetische Mittel wird oft auch als Durchschnitt bezeichnet. Für die Ausbildung eines inhaltlichen Verständnisses und den Aufbau von Grundvorstellungen zum arithmetischen Mittel nutzt man die Ausgleichs-, die Gleichverteilungs- und die Schwerpunkteigenschaft.

Die Boxplotdarstellung (Kastenschaubild) ist ein gutes Hilfsmittel für den Vergleich von Verteilungen von Daten zu einem bestimmten Merkmal. Dabei werden die ermittelten Kennwerte Minimum, Maximum, Median und Quartil genutzt.

Ein kompetenter und kritischer Umgang mit bereits erfassten Daten und ihren Darstellungen, aber auch mit verschiedenen Formen der Datenerfassung, insbesondere personenbezogener Daten, stellt im zunehmenden Maße eine grundlegende Alltagskompetenz dar.