Maße der zentralen Tendenz (aka Lagemaße)

… repräsentieren alle Einzelwerte der Verteilung zusammen

Gängige Maße der zentralen Tendenz: Modalwert (aka Modus), Arithmetisches Mittel (Mittelwert), Median

Modalwert (aka Modus)

- Am häufigsten vorkommender Wert einer Verteilung
- Berechnung erfordert lediglich Nominalskalenniveau
- Wert mit höchster absoluter Häufigkeit
- Graphen mit nur einem Modus heißen unimodal (aka eingipfelig)
- Zwei Modi = bimodale Verteilung
- Mehrere Maximalwerte nebeneinander = breitgipflig

Median

- Wert, der in der Mitte der Verteilung liegt (halbiert die Verteilung)
- Es liegen genauso viele Messwerte über wie unter dem Median
- Berechnung erfordert lediglich Ordinalskalenniveau

$$Md = \begin{cases} \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} & \text{falls } n \text{ gerade} \\ x_{(\frac{n+1}{2})} & \text{falls } n \text{ ungerade} \end{cases}$$

Vorteile Median gegenüber Mittelwert:

- Kann auch bei rangskalierten Merkmalen verwendet werden
- Ausreißer beeinflussen den Median kaum
- Schiefe Verteilungen: Median bildet zentrale Tendenz besser ab

Arithmetisches Mittel (Mittelwert)

- Gebräuchlichstes Maß der zentralen Tendenz
- Durchschnittswert einer Verteilung
- Nur für metrische Variablen sinnvoll (mind. Intervallskalenniveau)
- Der Mittelwert einer Variable x wird geschrieben als x̄

Summe aller Werte dividiert durch den Stichprobenumfang N

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Berechnung des Mittelwerts auf Grundlage der Häufigkeitstabelle

$$\bar{x} = \frac{\sum_{j=1}^{k} x'_j \cdot f_j}{n}$$

x'_j = mögliche Merkmalsausprägungen; f_j = absolute Häufigkeit der jeweili

Berechnung des Mittelwerts für zwei oder mehr Datensätze (Gruppen)

→ Gewogenes Arithmetisches Mittel (Weighted Mean)

$$\bar{\bar{x}} = \frac{n_1 \cdot \bar{x}_1 + n_2 \cdot \bar{x}_2}{n_1 + n_2}$$