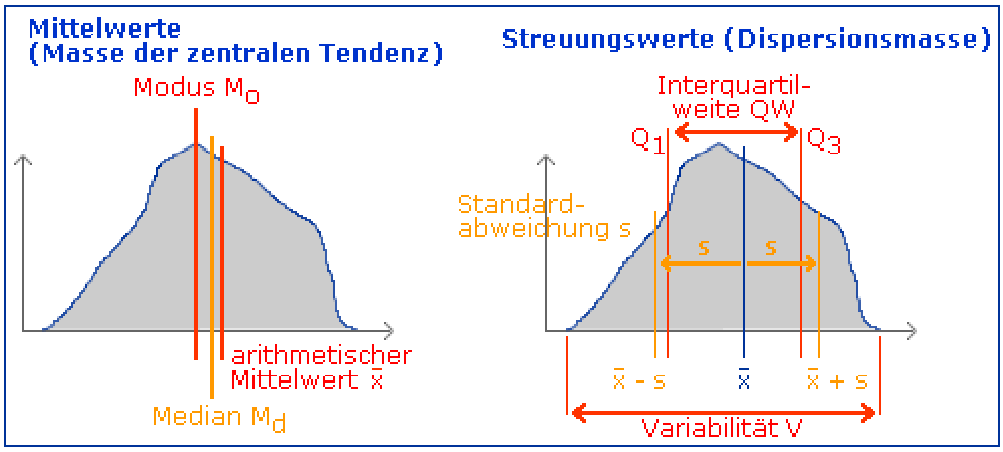
Zusammenfassung ‚univariate Verteilungen'

Verteilungsformen. Zur Charakterisierung der Form einer Häufigkeitsverteilung werden häufig folgende Begriffe verwendet: Ein- bzw. Mehrgipfligkeit (unimodal, bimodal und multimodal), Wölbung (schmalgipflig oder breitgipflig), Schiefe (symmetrisch oder asymmetrisch, linkssteil oder rechtssteil) oder auch spezielle Formen (u-förmig, abfallend oder glockenförmig).

Kennwerte zur Beschreibung von Häufigkeitsverteilungen
Eine Reihe von Kennzahlen, welche die Daten möglichst gut repräsentieren, können zur Beschreibung von Verteilungen verwendet werden. Dabei werden viele Einzelinformationen zu wenigen, aber aussagekräftigen Grössen verdichtet.

2 Gruppen von Kennwerten zur Beschreibung von Häufigkeitsverteilungen:



1. Masse der zentralen Tendenz (Mittelwerte)
Auf unterschiedliche Weise kann der „Schwerpunkt" oder „Mittelwert" einer Verteilung berechnet werden. Dies ermitteln die sogenannten Masse der zentralen Tendenz, auch Lagemasse genannt. Sie beschreiben den typischen, den zentralen oder durchschnittlichen Wert einer Verteilung.

- **Der arithmetische Mittelwert $\bar{x}$ [mean].** Die Summe aller Ausprägungswerte wird durch die Anzahl aller Beobachtungen dividiert.
- **Der Modalwert M_o [mode].** Der Modus ist derjenige Ausprägungsgrad der Merkmalsdimension, der am häufigsten beobachtet wurde.
- **Der Median M_d [median].** Dem Median entspricht der Grenzwert, unterhalb dessen 50% der Beobachtungen liegen.

2. Streuungswerte (Dispersionsmasse)
Häufig interessiert die Frage, wie dicht die einzelnen Daten beieinander liegen, oder wie stark sie streuen. Man spricht von einer grossen Streuung, wenn die Daten weit auseinander liegen. Diese Masszahlen charakterisieren somit die Variabilität oder Heterogenität der Beobachtungen.

- **Die Variabiliät V [range].** Die Variabilität entspricht der Differenz zwischen dem grössten und dem kleinsten beobachteten Merkmalswert (V = Maximaler Ausprägungsgrad – minimaler Ausprägungsgrad).
- **Die Interquartilweite QW [interquartil range] und Quartilabweichung Q.** Mit der Interquartilweite QW und der Quartilabweichung Q wird auf unterschiedliche Weise die „Breite" des Wertebereiches beschrieben, in dem die zentral gelegenen 50% der Merkmalsausprägungen liegen [$QW = Q_3 - Q_1$ bzw. $Q = QW/2 = (Q_3 - Q_1)/2$].
- **Die Varianz s^2 [variance] und die Standardabweichung s [standard deviation].** Der Varianz entspricht die Summe der quadrierten Abweichungen der Merkmalswerte vom arithmetischen Mittelwert, dividiert durch die Anzahl der Beobachtungen. Man erhält die Standardabweichung s, indem man die Quadratwurzel aus der Varianz s^2 berechnet.