Lineare Transformation der Daten:

$$\text{Sei} \quad y_1 = a \cdot x_1 + b, y_2 = a \cdot x_2 + b, ..., y_n = a \cdot x_n + b, a \neq 0\,. \text{ Dann gilt:}$$

$$\begin{aligned} \overline{y} &= a \cdot \overline{x} + b \text{ und } y_p = a \cdot x_p + b \text{ und } QKS_y = QKS_x \\ s_y^2 &= a^2 \cdot s_x^2 \text{ und } s_y = |a| \cdot s_x \text{ und falls } b = 0: \; V_y = V_x\,. \end{aligned}$$

geordnete Stichprobe: $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$.
Konzentrationsrate K_m gibt den Anteil der m größten Werte an der Gesamtsumme an:

$$K_m = \frac{x_{(n)} + x_{(n-1)} + \ldots + x_{(n-m+1)}}{x_1 + \ldots + x_n}, m = 1, \ldots, n\,.$$

Der Graph der Abbildung $m \to K_m$, wobei zwischen den Punkten (m, K_m) linear interpoliert wird, heißt **Konzentrationskurve**.
Lorenzkurve $\frac{m}{n} \to L_m$ gibt jeweils den Anteil der m kleinsten Werte an der Gesamtsumme an:

$$L_m = \frac{x_{(1)} + x_{(2)} + \ldots + x_{(m)}}{x_1 + \ldots + x_n}, m = 1, \ldots, n\,.$$

Gini-Koeffizient: Maßzahl für den Grad der relativen Konzentration

$$G = 2 \cdot \frac{1 \cdot x_{(1)} + 2 \cdot x_{(2)} + \ldots + n \cdot x_{(n)}}{n \cdot (x_1 + \ldots + x_n)} - \frac{n+1}{n} = 1 - \frac{1}{n} \sum_{i=1}^{n} (L_i + L_{i-1})\,.$$

Lorenzkurve gruppierter Daten: $\left(\sum_{j=1}^{i} r_j, L_i^*\right)$, $i = 1, \ldots, K$, zwischen diesen Punkten wird linear interpoliert.
Gini-Koeffizient bei gruppierten Daten: Maßzahl für den Grad der relativen Konzentration

$$G = 1 - \sum_{i=1}^{K} r_i \cdot (L_i^* + L_{i-1}^*)\,.$$

Herfindahl-Index

$$H = \sum_{i=1}^{n} \left(\frac{x_i}{n\bar{x}}\right)^2 = \frac{1}{n} \left(\frac{n-1}{n} V_x^2 + 1\right)\,.$$

Rangkorrelationskoeffizient nach Spearman: Ohne Bindungen (d.h. $x_i \neq x_j$ und $y_i \neq y_j$ für $i \neq j$) gilt:

$$r_{xy}^{Rang} = 1 - \frac{6 \cdot \sum_{i=1}^{n} (R_i - R_i')^2}{n \cdot (n^2 - 1)} \quad \text{mit } R_i = j \Leftrightarrow x_i = x_{(j)} \text{ und } R_i' = k \Leftrightarrow y_i = y_{(k)}.$$

Empirische Kovarianz:

$$s_{xy} = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y}) = \frac{1}{n-1} \left(\sum_{i=1}^{n} x_i \cdot y_i - n \cdot \overline{x} \cdot \overline{y}\right)\,.$$

Empirischer Korrelationskoeffizient (Korrelationskoeffizient nach Pearson):

$$\begin{aligned} r_{xy} &= \frac{s_{xy}}{s_x \cdot s_y} = \frac{\sum_{i=1}^{n} (x_i - \overline{x}) \cdot (y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2 \cdot \sum_{i=1}^{n} (y_i - \overline{y})^2}} \\ &= \frac{\sum_{i=1}^{n} x_i \cdot y_i - n \cdot \overline{x} \cdot \overline{y}}{\sqrt{\left(\sum_{i=1}^{n} x_i^2 - n \cdot (\overline{x})^2\right) \cdot \left(\sum_{i=1}^{n} y_i^2 - n \cdot (\overline{y})^2\right)}}\,. \end{aligned}$$

$$\begin{aligned} \widehat{m} &= \frac{s_{xy}}{s_x^2} = \frac{\sum_{i=1}^{n} (x_i - \overline{x}) \cdot (y_i - \overline{y})}{\sum_{i=1}^{n} (x_i - \overline{x})^2} = \frac{\sum_{i=1}^{n} x_i \cdot y_i - n \cdot \overline{x} \cdot \overline{y}}{\sum_{i=1}^{n} x_i^2 - n \cdot (\overline{x})^2} \\ \widehat{b} &= \overline{y} - \widehat{m} \cdot \overline{x}\,. \end{aligned}$$