Korrelationen

für ordinale & bivariat-normalverteilte Variablen

Samuel Merk

Inhalte dieses Erklärvideos

- Assoziation zweier (mindestens) ordinalskalierter Variablen (Kendal's τ_b)
- Assoziation (zweier) bivariat-normalverteilter Variablen (Pearson's r)

Assoziation zweier (mind.) ordinalskalierter Variablen

Zusammenh. von Bildungsabschl.

DERSTANDARD

Jubelangebot Abo Immosuche Jobsuche

Anmelden

Inland › Bildung International Wirtschaft Web Sport Panorama Kultur Etat Wissenschaft Lifestyle Diskurs Karriere mehr...

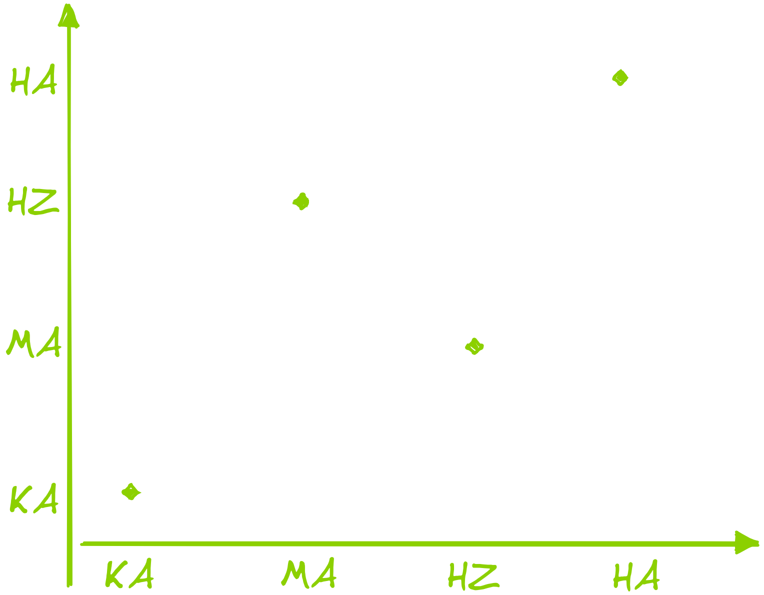
759 Postings

BILDUNGSSTATISTIK

Bildung wird weiterhin vererbt

Es gibt mehr Bildungsauf- und weniger -absteiger. Insgesamt dominieren aber immer noch jene, deren höchster Abschluss der "Familientradition" entspricht

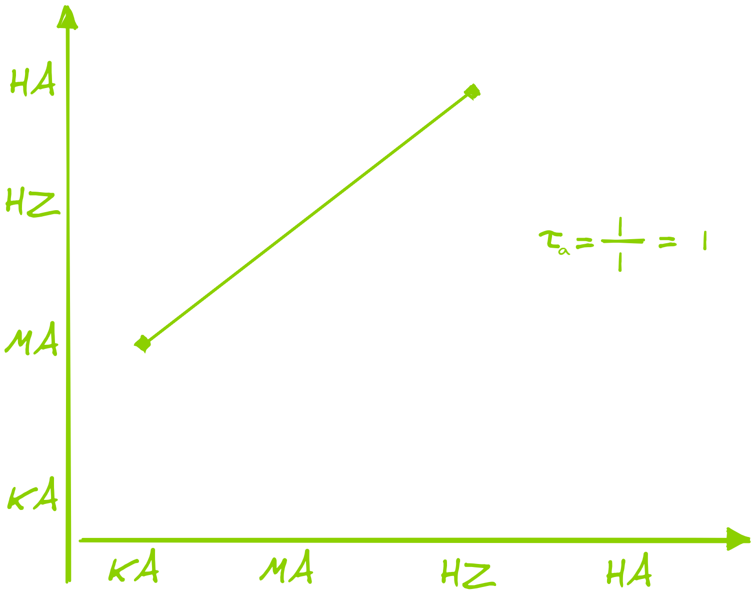
Definition Kendall's τ_a (1938)

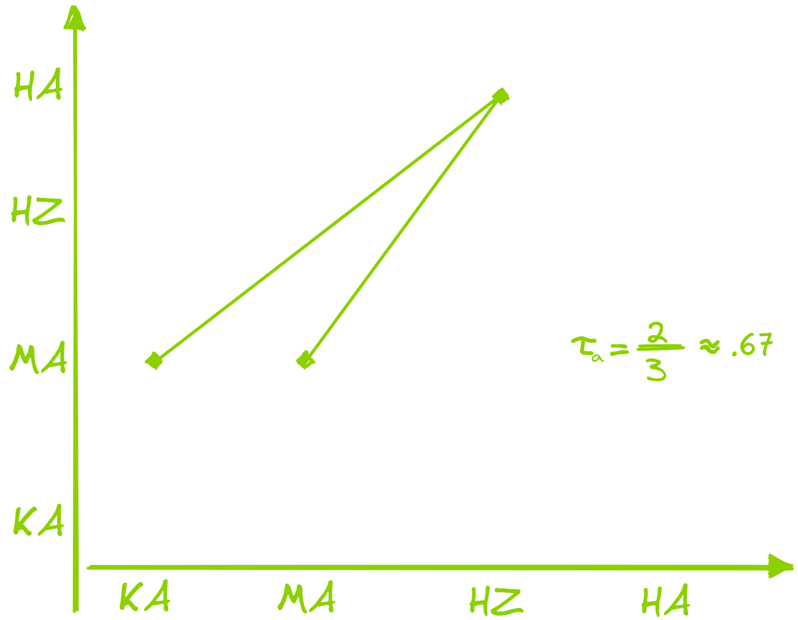


$$\text{Kendall's } \tau_A = \frac{\#(\text{konkordante Paare}) - \#(\text{diskordante Paare})}{\#(\text{alle Paare})}$$

$$= \frac{5 - 1}{6} = \frac{4}{6} =. \bar{6}$$

Problem Kendall's τ_a





Rangbindungen (engl. »ties«) nennt man Paarvergleiche, bei denen (mindestens) eine der beiden Variablen gleich ausgeprägt ist. Kendall's τ_A, verhält sich problematisch wenn Rangbindungen vorliegen.

Definition Kendall's τ_{b}

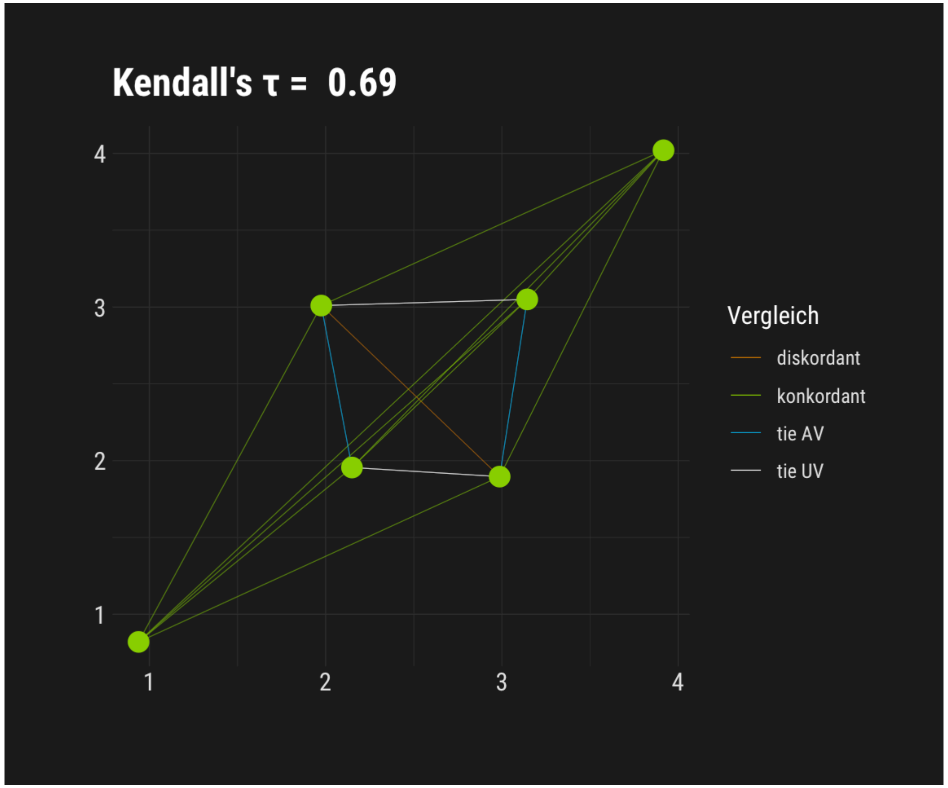
Gelten folgende Bezeichnungen

- n_{K} : Anzahl aller konkordanten Paare
- n_{D} : Anzahl aller diskordanten Paare
- $n_{\mathrm{B}(X)}$: Anzahl aller Rangbindungen in der Variable X
- $n_{\mathrm{B}(Y)}$: Anzahl aller Rangbindungen in der Variable Y

ist Kendall's tau_{b} wie folgt definiert

$$\tau_{\mathrm{b}}\left(X, Y\right) = \frac{n_{\mathrm{K}} - n_{\mathrm{D}}}{\sqrt{\left(n_{\mathrm{K}} + n_{\mathrm{D}} + n_{\mathrm{B}(X)}\right) \cdot \left(n_{\mathrm{K}} + n_{\mathrm{D}} + n_{\mathrm{B}(Y)}\right)}}$$

Visualisierung Kendall's τ_b



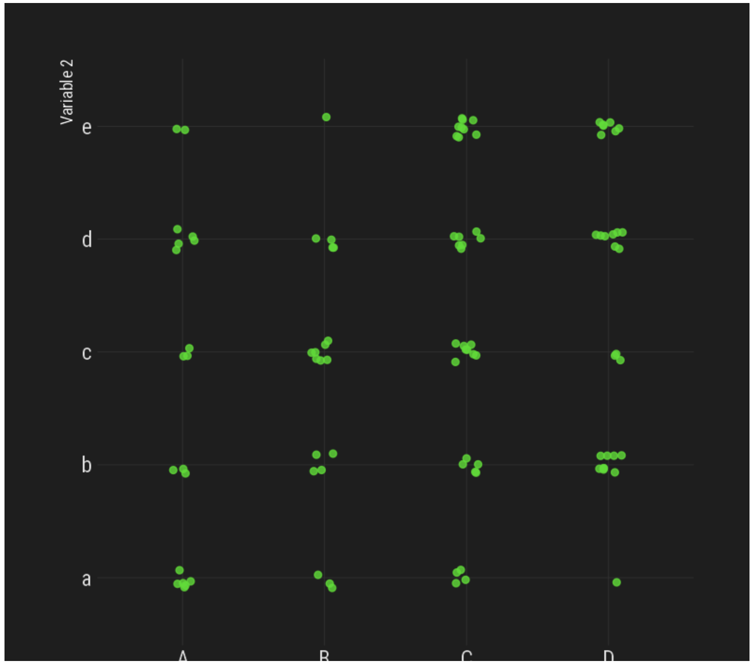
$$\tau_b\,(X, Y) = \frac{n_K - n_D}{\sqrt{(n_K + n_D + \quad\quad)\cdot(n_K + n_D + n_{B(Y)})}} = \frac{10-1}{\sqrt{(10+1+\quad)\cdot(10+1+2)}} \approx .69$$

Eigenschaften Kendall's τ_{b}

Visual Guessing Kendall's τ_b

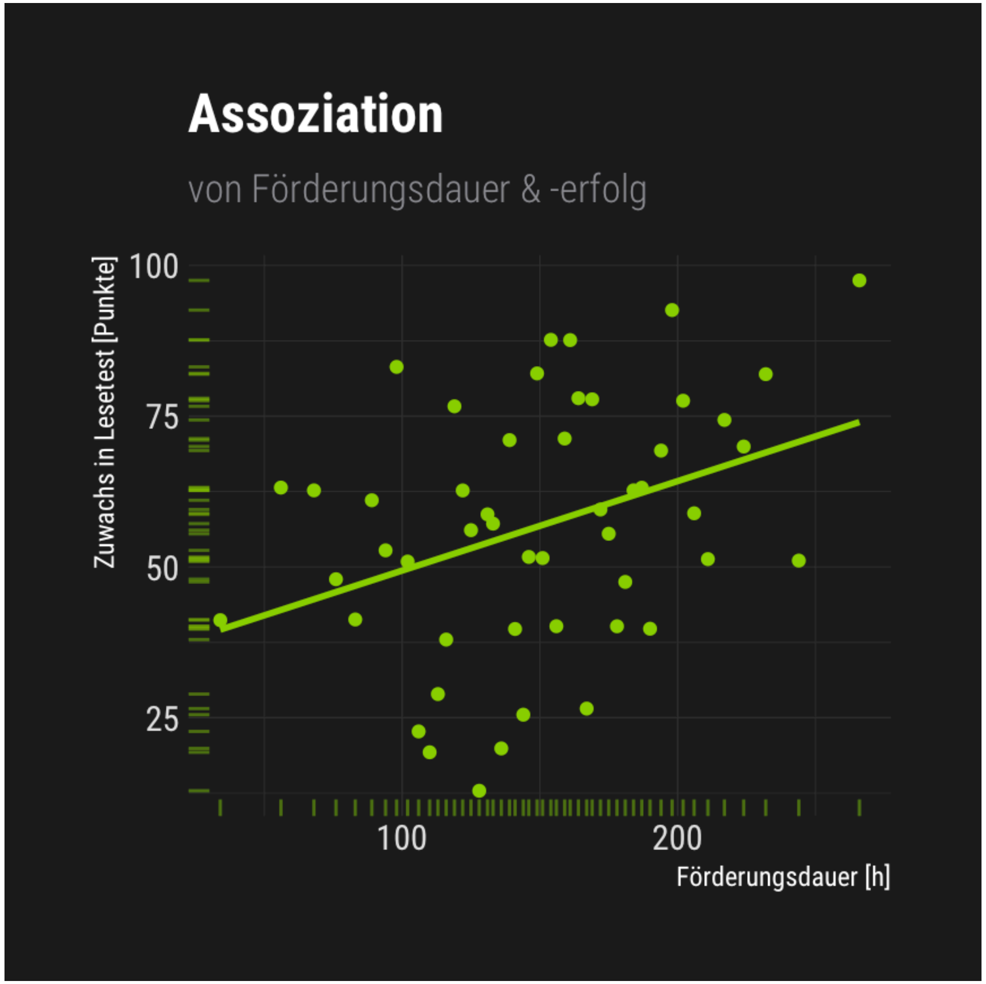
Aufgabe:

Schätzen Sie Kendall's τ^b aus den unten graphisch dargestellten Daten. Diese enthalten die ordinalen Variablen »Variable 1« (mit den Ausprägungen A < B < C < D) und »Variable 2« (mit den Ausprägungen a < b < c < d < e)

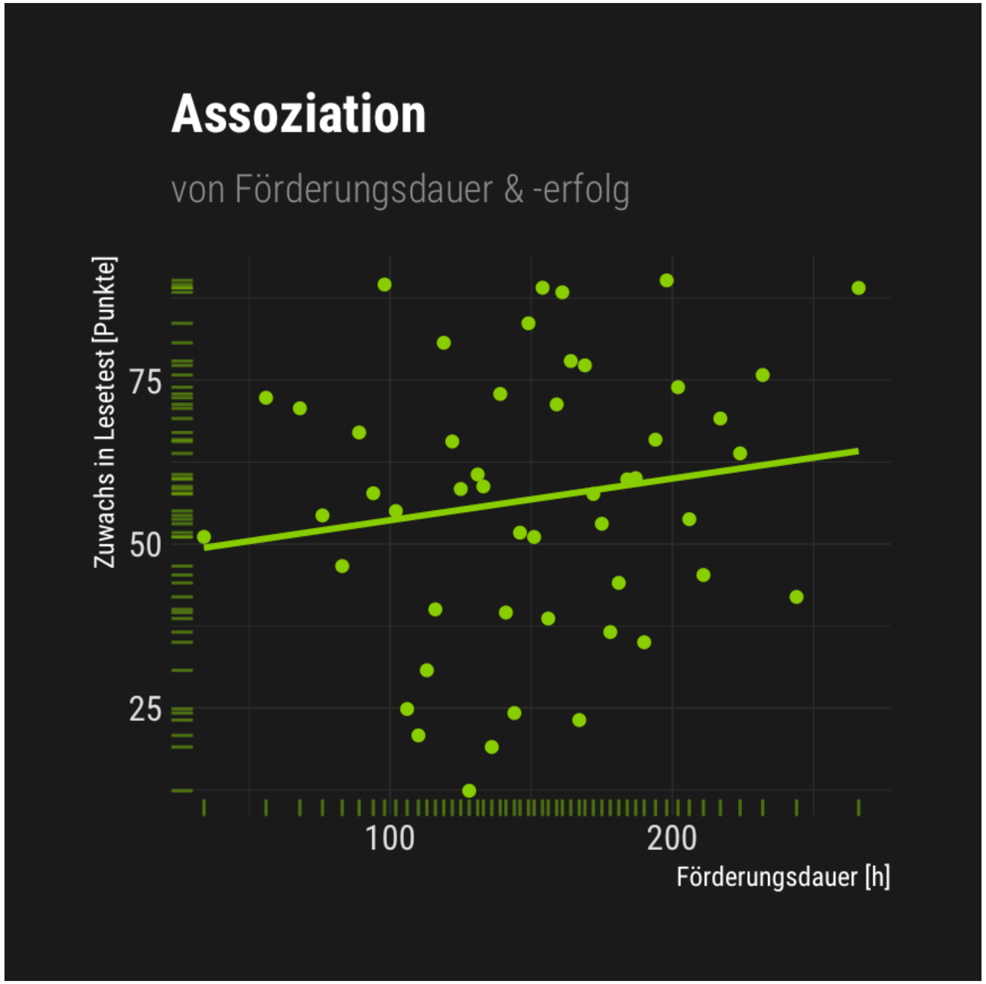


Assoziation zweier bivariat normalverteilter Variablen

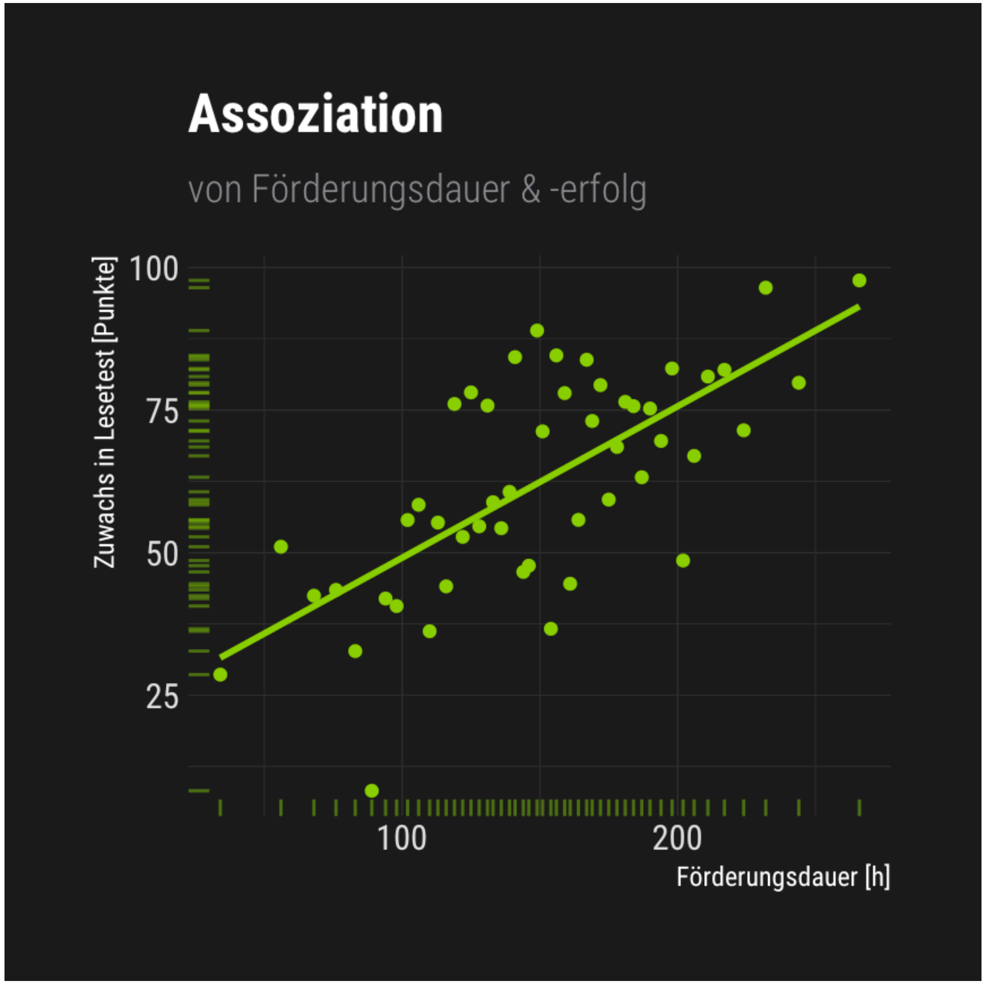
Förderungsumfang & -erfolg



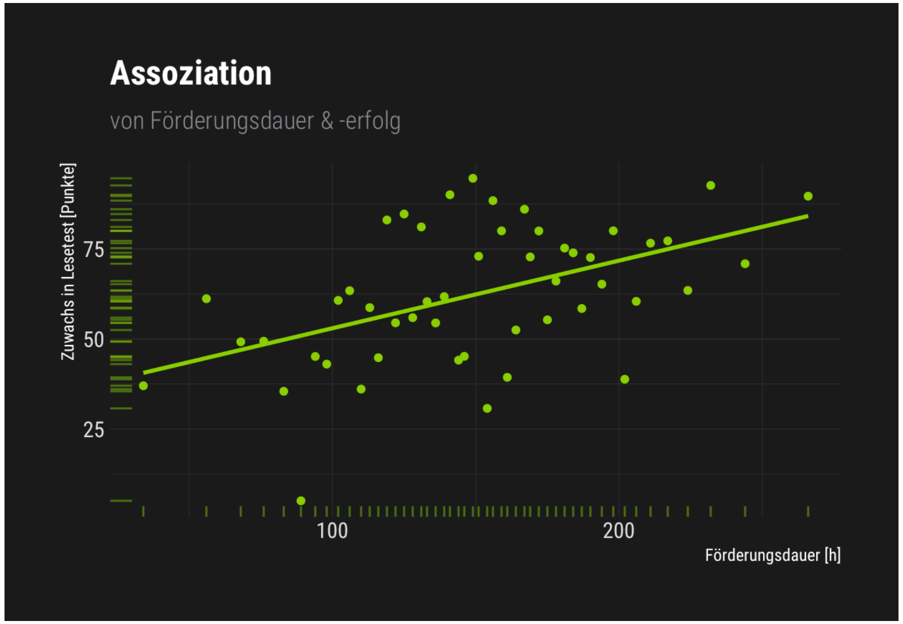
Förderungsumfang & -erfolg

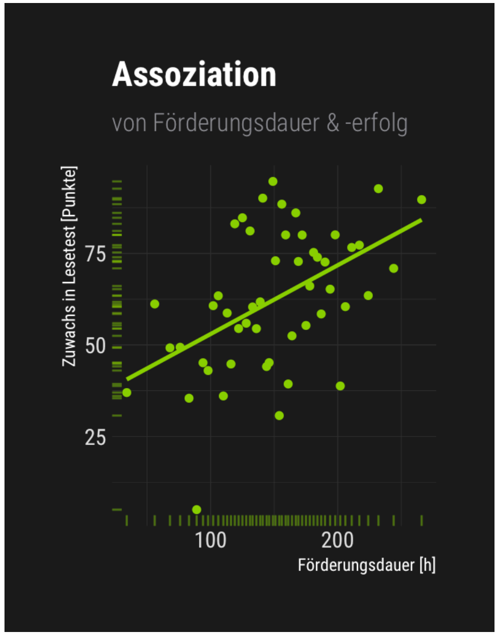


Förderungsumfang & -erfolg

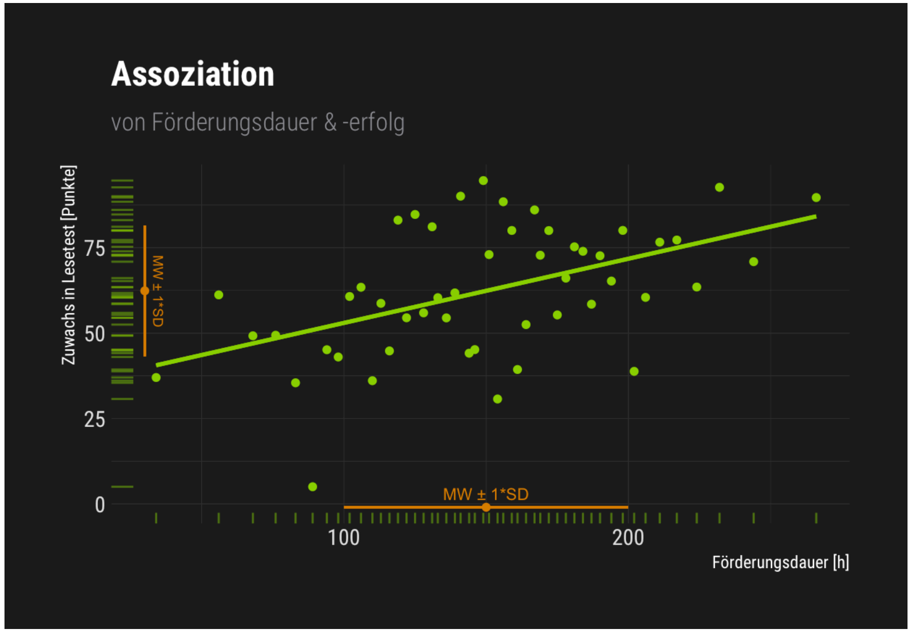


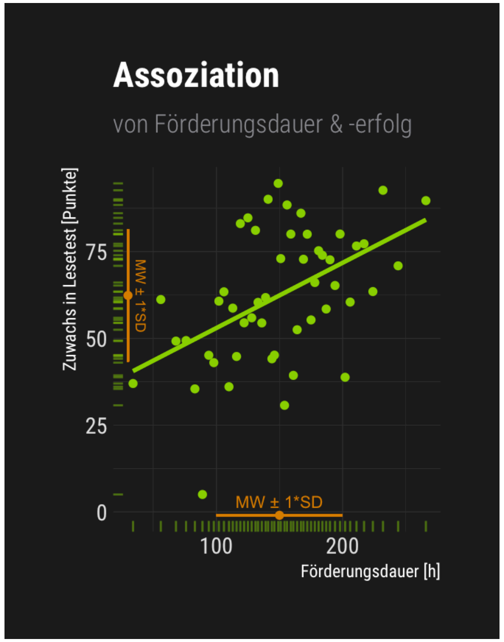
Gleiche Daten - untersch. Steigung?



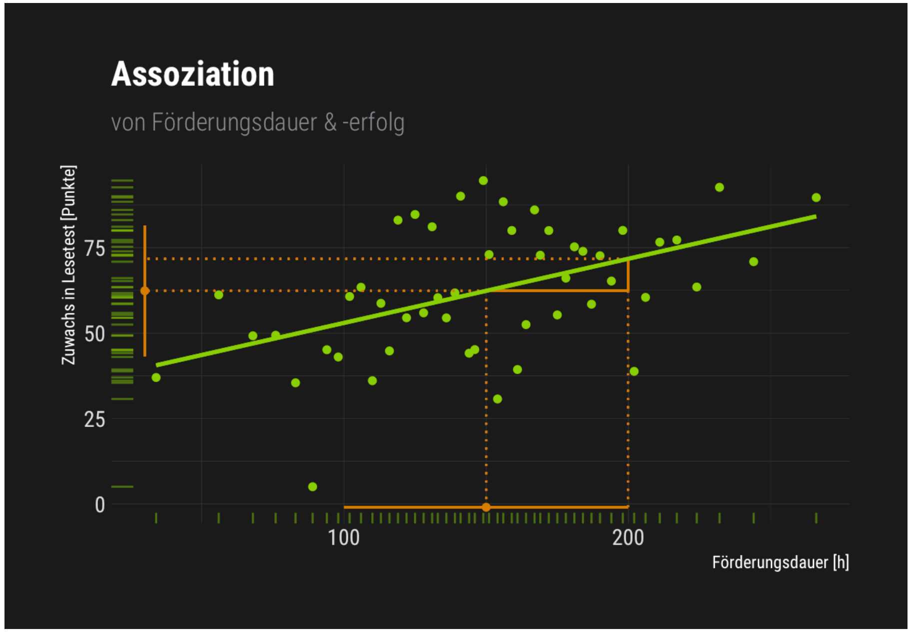


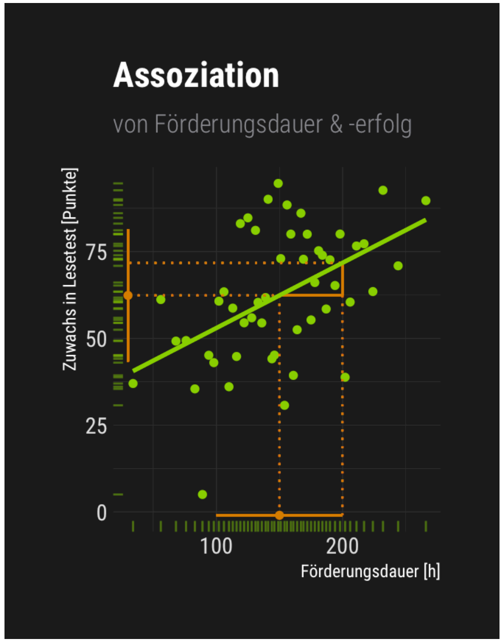
Stand. Steig. der Ausgleichsgerade



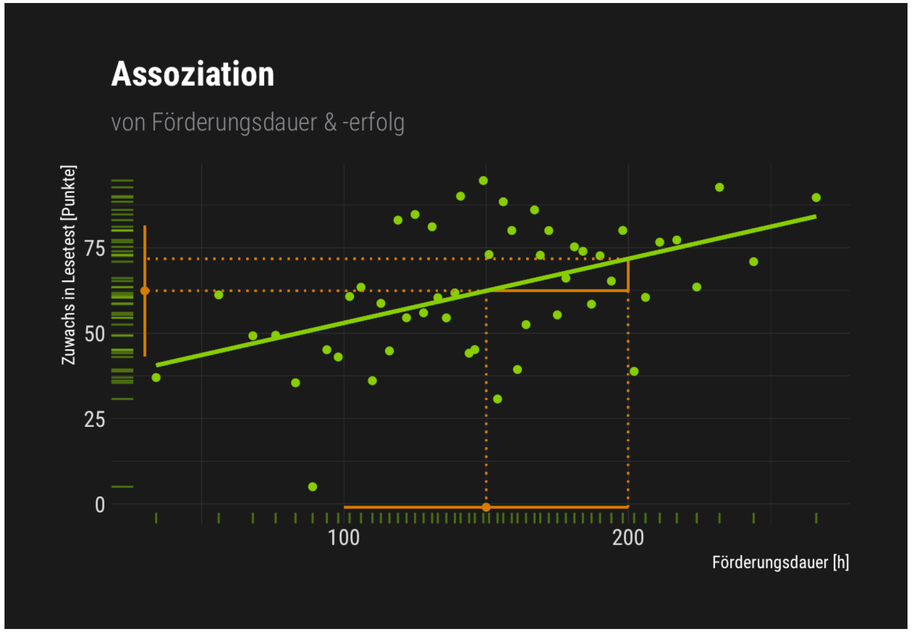


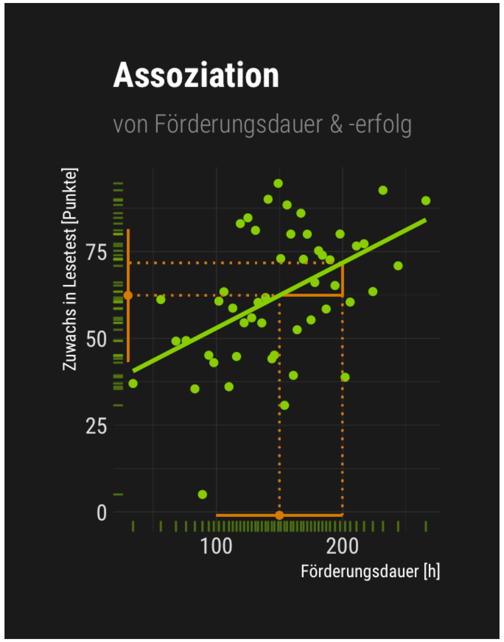
Stand. Steig. der Ausgleichsgerade





Stand. Steig. der Ausgleichsgerade





> Misst man die Steigung einer »Ausgleichsgerade« in Standardabweichungen, erhält man den Pearson Korrelationskoeffizienten r

Ableitung: Formel Pearsons's r

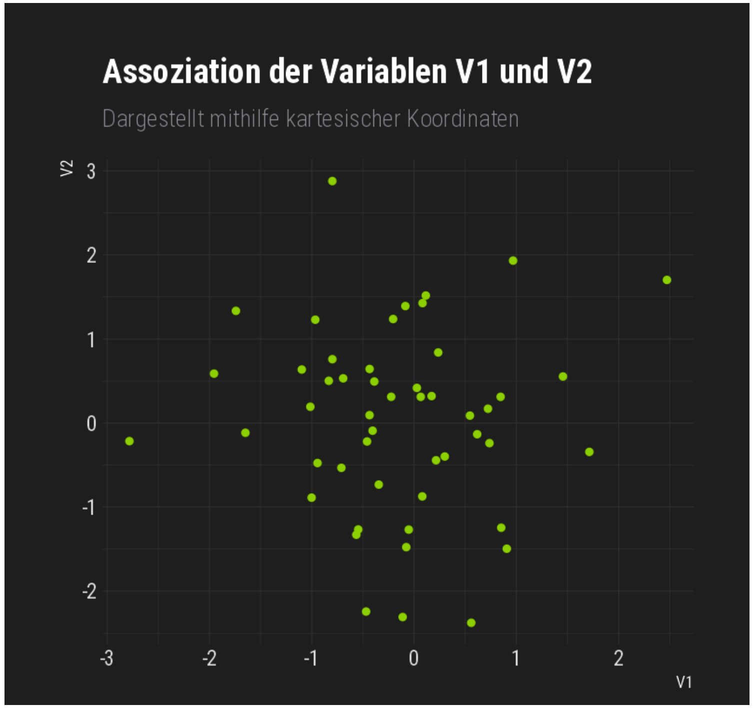
$$r_{x,y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Eigenschaften Pearson's r

Visual Guessing Pearson’s r

Aufgabe:

Schätzen Sie Pearsons's r aus den unten graphisch dargestellten Daten



Literatur

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Second). New Jersey: Lawrence Erlbaum.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, *30*(1-2), 81–93.