

Ein einfacher Workflow zur Datenvisualisierung

ChatGPT oder Claude

+

Skill für Statistik und Datenvisualisierung

+

Flourish Studio (Free Version)

als Plugin (ChatGPT)

oder

Konnektor (Claude)

Danny Walther

www.ki-recherchen.de

www.linkedin.com/in/ki-recherchen

1. Datensatz in Claude oder ChatGPT laden:

Hier:

Künstlicher Datensatz zum Thema KI-Nutzung

Inhalt:

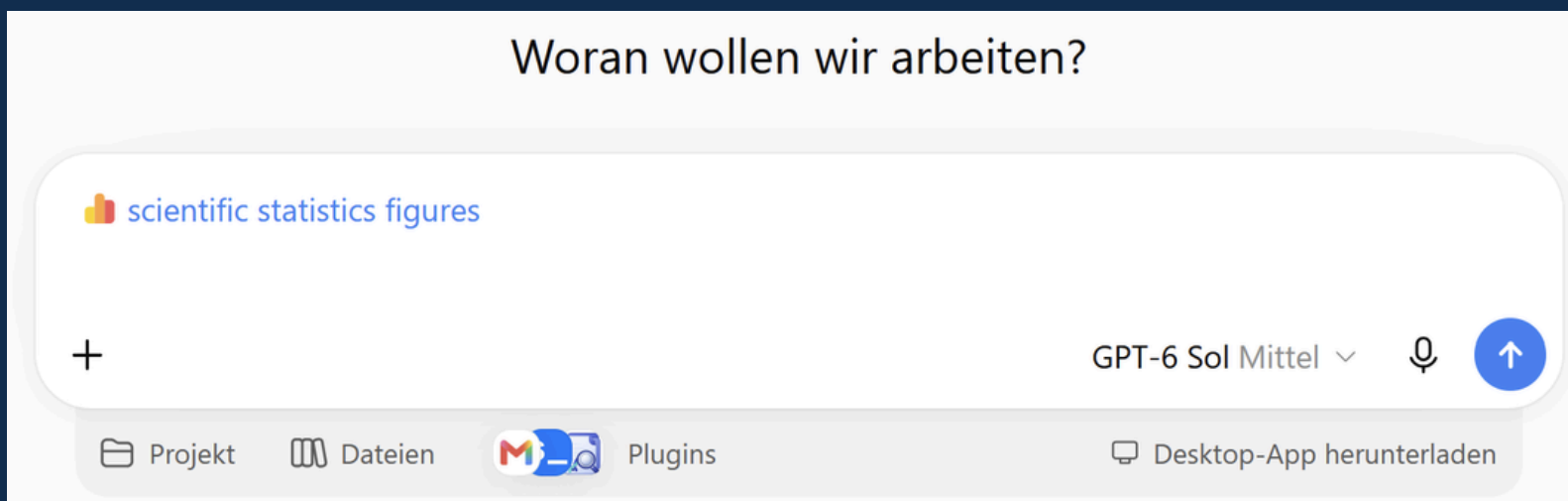
Fiktive Daten von : 300 Studierenden

120 Lehrenden

Mit jeweils bis zu 29 Variablen

J67																		
	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
107	106 Medieninformatik	2	29 m		nein	nein	ja	nein	nein	ja	täglich	1 ja	nein	nein	nein	nein	nein	nein
108	107 Medieninformatik	1	21 m		nein	nein	nein	nein	nein	ja	täglich	1 ja	ja	nein	ja	nein	nein	nein
109	108 Medieninformatik	5	28 w		nein	nein	ja	nein	nein	ja	täglich	5 ja	nein	nein	nein	nein	nein	nein
110	109 Medieninformatik	4	26 m		nein	nein	nein	nein	nein	ja	wöchentlich	1 ja	ja	nein	nein	nein	nein	nein
111	110 Medieninformatik	1	21 w		nein	nein	nein	nein	nein	ja	täglich	1 ja	nein	nein	nein	nein	nein	nein
112	111 Medieninformatik	4	21 m		nein	nein	nein	nein	nein	ja	täglich	5 ja	nein	nein	nein	nein	nein	nein
113	112 Medieninformatik	6	18 m		nein	nein	ja	nein	nein	ja	täglich	1 ja	nein	ja	nein	nein	nein	nein
114	113 Medieninformatik	6	23 w		nein	nein	ja	ja	ja	ja	täglich	1 ja	ja	nein	nein	nein	nein	nein
115	114 Medieninformatik	3	28 m		nein	nein	nein	nein	nein	ja	täglich	3 ja	nein	nein	nein	nein	nein	nein
116	115 Medieninformatik	7	20 m		nein	nein	nein	nein	ja	ja	wöchentlich	1 ja	nein	nein	nein	ja	nein	nein
117	116 Medieninformatik	1	28 w		nein	nein	ja	nein	nein	ja	täglich	1 ja	ja	nein	nein	nein	nein	nein
118	117 Medieninformatik	3	25 w		nein	nein	nein	nein	nein	ja	monatlich	1 nein	nein	nein	nein	nein	nein	nein
119	118 Medieninformatik	7	29 m		nein	nein	ja	nein	nein	ja	täglich	1 ja	ja	nein	nein	nein	nein	nein
120	119 Medieninformatik	6	28 m		nein	nein	nein	nein	ja	ja	täglich	2 ja	nein	nein	ja	nein	nein	nein
121	120 Medieninformatik	4	18 w		nein	nein	ja	nein	nein	ja	täglich	1 ja	nein	nein	nein	nein	nein	nein
122	121 Bauingenieurwesen	2	26 w		nein	nein	ja	nein	nein	ja	wöchentlich	1 ja	nein	nein	nein	nein	nein	nein
123	122 Bauingenieurwesen	4	20 w		nein	nein	ja	nein	nein	ja	täglich	1 ja	nein	nein	nein	nein	nein	nein
124	123 Bauingenieurwesen	5	18 m		nein	nein	nein	nein	nein	ja	täglich	3 ja	nein	nein	nein	nein	nein	nein
125	124 Bauingenieurwesen	5	22 m		nein	nein	nein	nein	nein	ja	täglich	3 ja	nein	nein	nein	nein	nein	nein
126	125 Bauingenieurwesen	6	18 m		nein	nein	ja	nein	nein	ja	täglich	1 ja	ja	nein	nein	nein	nein	nein
127	126 Bauingenieurwesen	7	26 w		nein	nein	ja	ja	nein	ja	täglich	1 ja	nein	nein	nein	nein	nein	nein
128	127 Bauingenieurwesen	5	28 m		nein	nein	nein	ja	nein	ja	täglich	1 ja	nein	nein	ja	nein	nein	nein
129	128 Bauingenieurwesen	2	29 m		nein	nein	ja	ja	nein	ja	monatlich	3 ja	nein	nein	nein	nein	nein	nein
130	129 Bauingenieurwesen	4	25 w		nein	nein	nein	nein	nein	ja	täglich	1 ja	nein	ja	nein	nein	nein	nein
131	130 Bauingenieurwesen	1	28 w		nein	nein	nein	nein	ja	ja	täglich	1 ja	nein	nein	nein	nein	nein	nein
132	131 Bauingenieurwesen	7	22 m		nein	nein	nein	nein	ja	ja	monatlich	5 ja	nein	nein	nein	nein	nein	nein
133	132 Bauingenieurwesen	6	26 w		nein	nein	nein	ja	ja	ja	täglich	2 nein	nein	ja	nein	nein	nein	nein
134	133 Bauingenieurwesen	1	18 w		nein	nein	nein	ja	nein	ja	täglich	1 ja	nein	nein	nein	nein	nein	nein
135	134 Bauingenieurwesen	7	19 w		nein	nein	ja	nein	nein	ja	täglich	3 ja	nein	nein	ja	nein	nein	nein
136	135 Bauingenieurwesen	7	21 w		nein	nein	nein	nein	nein	ja	täglich	4 nein	nein	nein	nein	nein	nein	nein
137	136 Bauingenieurwesen	6	26 w		nein	nein	ja	nein	nein	ja	täglich	1 ja	nein	ja	nein	nein	nein	nein

2. Skill “scientific statistic figures” aktivieren



Wissenschaftliche Statistik und Abbildungen

Ziel: **statistisch nachvollziehbare, reproduzierbare und publikationsreife** Ergebnisse. Statistik wird nicht als nachträgliche Dekoration eines Plots behandelt. Erst wird die analytische Frage sauber definiert und gerechnet; danach wird visualisiert.

Harte Grundregel

Bei statistischen Aufgaben immer diese Reihenfolge einhalten:

Datenaudit -> Forschungsfrage/Estimand -> Skalenniveau -> Methodenwahl ->
Voraussetzungen -> Missing-Data-Strategie -> Berechnung -> Verifikation ->
Unsicherheit/Multiplizität -> Interpretation -> Visualisierung

Keinen Schritt stillschweigend überspringen, wenn er das Ergebnis verändern kann.

3. Datenaudit durchführen lassen

Wichtige Audit-Befunde

- **IDs:** `ID` in beiden Blättern ist reiner Laufindex (1...n, 100 % eindeutig) – keine inhaltliche Variable, wird nicht analysiert.
- **Fehlende Werte:** Nur `Prüfungsnote_KI_Modul` hat Missings (253 von 300, 84 %) – das ist aber **kein Datenproblem**, sondern strukturell: die Note existiert nur für die 47 Studierenden, die `KI_Modul_belegt = ja` angegeben haben. Effektives n für diese Variable ist 47, nicht 300.
- **Near-constant Variablen** (>95 % gleicher Wert, liefern kaum Varianz): `Nutzt_KI`, `Elicit`, `Scite`, `SemanticScholarAI` bei Studierenden; `Claude`, `KI_Workshops_gehalten` bei Lehrenden.
- **Skalenniveaus:** Mehrere 1-5-Items (`KI_Selbsteinschätzung`, `Prompting_Kenntnisse`, `Verständnis_KI_Limits`, `KI_Rolle_im_Arbeitsprozess`) sind Ordinalskalen und nicht automatisch metrisch zu behandeln. Die Testscores (`KI_Testscore`, `Prompting_Testscore`, `Halluzinationen_Erkennen_Score`, `Quellenbewertung_Score` etc.) sind hingegen echte Summenscores mit 40+ Ausprägungen – metrisch vertretbar.
- **Tabelle1** ist im Grunde leer (nur „...“ statt Zahlen) und für Analysen nicht nutzbar.
- Keine Duplikate, keine konstanten Variablen, keine verdächtigen Missing-Codes (`-99`, `999` etc.) gefunden.

Damit die Statistik zur richtigen Frage passt (Schritt S1 des Workflows) – was soll ich als Nächstes tun?

4. Erste Berechnungen machen

Ergebnisse: Studierende vs. Lehrende

Design: unabhängige Gruppen, Studierende (n=300) vs. Lehrende (n=120). Missing-Data-Strategie: für die verglichenen Variablen keine fehlenden Werte in beiden Gruppen — effektives n = Gesamt-n. Multiplizität: zwei getrennte Testfamilien mit Holm-Korrektur — (1) die vier primären Vergleiche (Testscores, Nutzung, Häufigkeit), (2) die fünf Einzel-Tools (explorativ).

a) KI-Testscore — Welch-t(280.9) = -3.90, Differenz = -2.80 Punkte, 95%-KI [-4.21, -1.39], Hedges g = -0.38 [-0.59, -0.17], p_Holm < .001. Sensitivitätsanalyse (Mann-Whitney-U, wegen Varianzheterogenität Levene p=.017) bestätigt: p < .001. Lehrende erzielen höhere Testwerte, Effekt klein bis mittel.

b) Prompting-Testscore — Differenz = -1.81 Punkte, 95%-KI [-3.60, -0.03], Hedges g = -0.20 [-0.41, 0.02], p_Holm = .047. Das CI von g schließt 0 knapp ein — der Effekt ist real, aber klein und an der Signifikanzgrenze; keine starke Aussage rechtfertigen.

c) Nutzt_KI (binär) — 99.0 % der Studierenden vs. 66.7 % der Lehrenden nutzen KI, Risikodifferenz = 32.3 Prozentpunkte, 95%-KI [23.8, 40.8], p_Holm < .001.

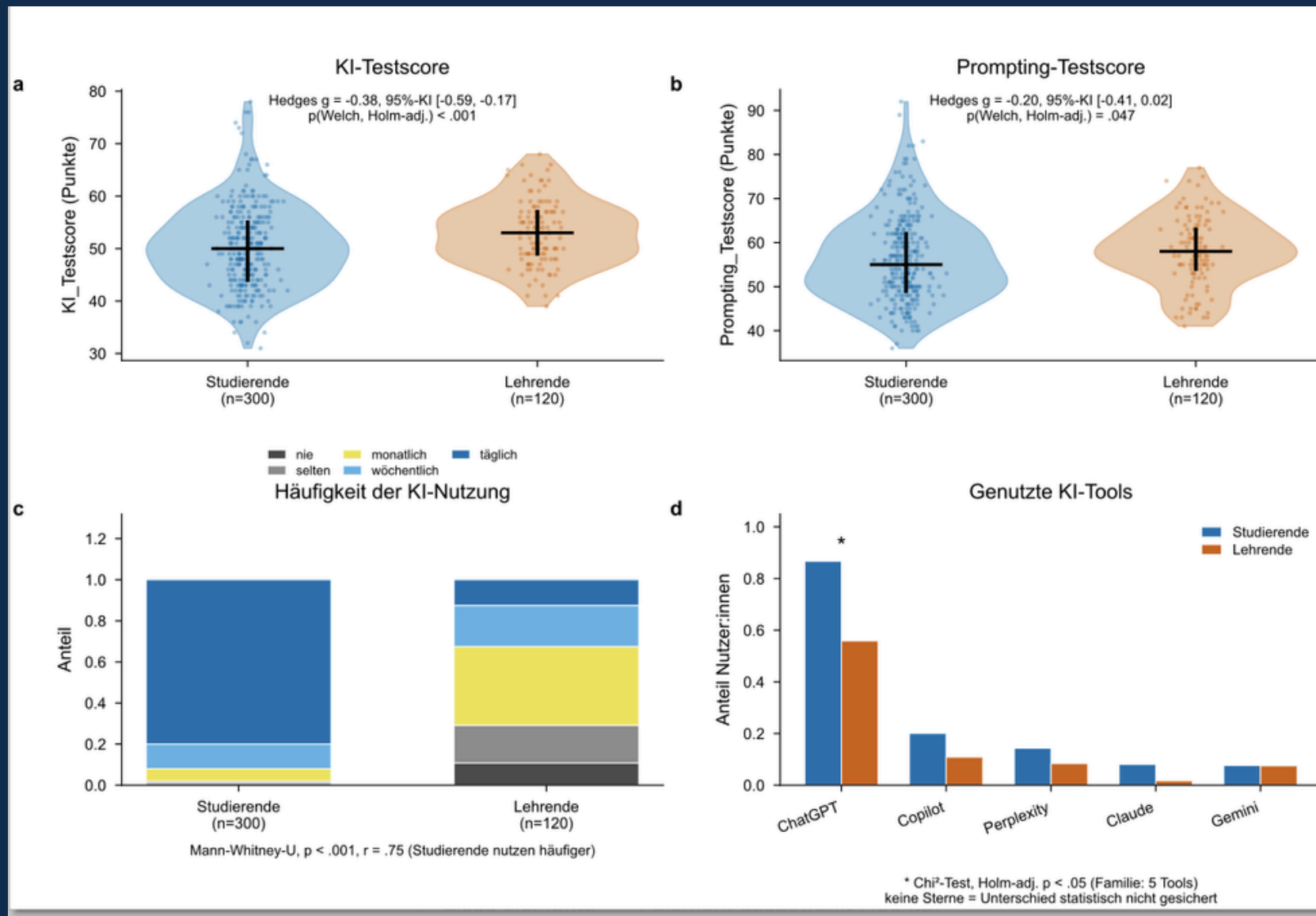
c) Häufigkeit der Nutzung (ordinal) — Mann-Whitney-U, p_Holm < .001, rangbiseriale r = .75: Studierende nutzen KI deutlich häufiger (80 % täglich) als Lehrende (nur 12,5 % täglich, Schwerpunkt monatlich/selten).

d) Einzelne Tools — nach Holm-Korrektur (Familie von 5) bleibt nur **ChatGPT** statistisch gesichert unterschiedlich (86,7 % vs. 55,8 %, p_Holm < .001). Copilot (p_Holm = .075) und Claude (p_Holm = .060) sind Trends, aber nicht gesichert; Perplexity und Gemini zeigen keinen belastbaren Unterschied.

Interpretation ohne Überinterpretation: Die Daten sind querschnittlich und korrelativ — kein Kausalschluss möglich (z. B. könnte Rollenunterschied, nicht "Kompetenz", die Testscore-Differenz erklären). Der auffälligste, robusteste Befund ist der **Nutzungsunterschied** (fast durchgängige vs. selektive Nutzung), nicht die Testscore-Differenz, die zwar signifikant, aber klein ist.

Alle Zahlen wurden direkt aus den Rohdaten berechnet (nicht aus der Grafik abgelesen) und mit einer manuellen Gegenrechnung des Welch-t verifiziert; das vollständige Ergebnis-JSON liegt bei.

5. Erste Visualisierungen erhalten, herunterladen und die SVG-Datei ggf. weiterbearbeiten



	Gruppenvergleich studierende... Dokument · PDF	Herunterladen 
	Gruppenvergleich studierende lehr... Bild · SVG	Herunterladen
	Gruppenvergleich studierende... Bild · PNG	Herunterladen 
	Group comparison results Code · JSON	Herunterladen 

6. Weitere Berechnungen durchführen

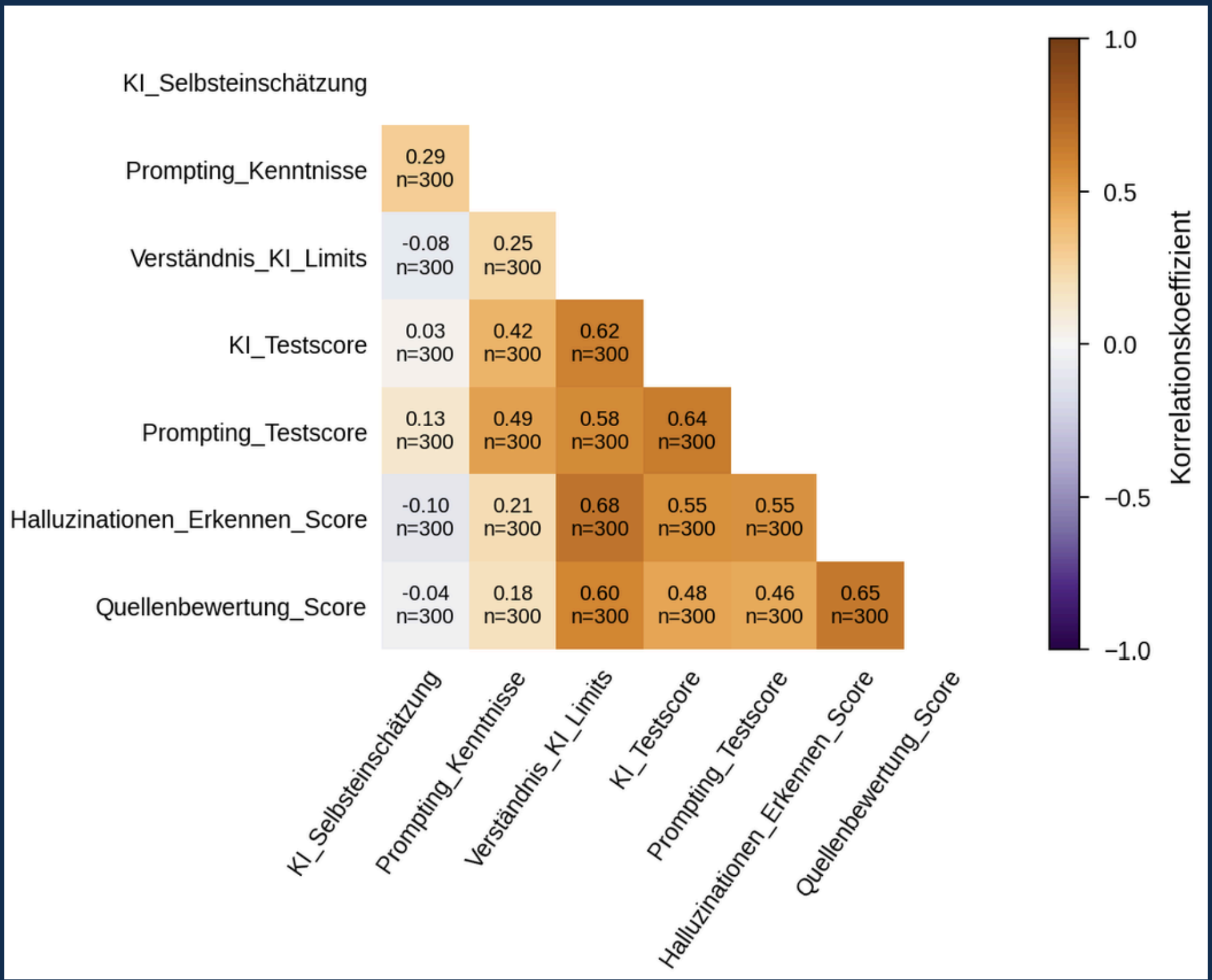
Verfahren: Pearson r und Spearman ρ , $n = 300$ (Studierende, effektives n konstant über alle Paare, keine Missings), Holm-Korrektur über die Familie von 21 einzigartigen Paarungen je Methode (siehe S7). Kernwerte direkt aus den Rohdaten gerechnet und mit einer unabhängigen scipy-Gegenrechnung verifiziert.

Befund — beide Methoden stimmen im Muster überein (Robustheit gegenüber der Pearson-vs-Spearman-Entscheidung bei den ordinalen 1-5-Items):

- **Starker positiver Cluster** zwischen `Verständnis_KI_Limits`, `KI_Testscore`, `Prompting_Testscore`, `Halluzinationen_Erkennen_Score` und `Quellenbewertung_Score`: r/ρ zwischen .42 und .68 — diese fünf Variablen hängen inhaltlich eng zusammen (alle nach Holm-Adjustierung signifikant).
- **KI_Selbsteinschätzung** korreliert nur schwach mit fast allem ($r = -.10$ bis $.29$) — subjektive Selbsteinschätzung sagt in diesen Daten wenig über objektive Testleistung aus.
- **Auffällig:** `Verständnis_KI_Limits` korreliert *negativ* mit `KI_Selbsteinschätzung` ($r = -.08$, $\rho = -.13$, beide $p > .05$, nicht signifikant) und `Halluzinationen_Erkennen_Score` ($r = -.10$, $\rho = -.15$), aber *stark positiv* mit den übrigen Testscores. Das ist inhaltlich auffällig genug, dass ich es nicht kommentarlos stehen lassen will: Das könnte auf eine **invers codierte Skala** hindeuten (z. B. 1 = "verstehe Limits sehr gut" statt "gar nicht") oder ein echtes, aber gegenläufiges Muster sein. Ohne Codebuch kann ich das nicht auflösen — bitte die Kodierung dieser Variable prüfen, bevor daraus inhaltliche Schlüsse gezogen werden.
- Alle Zusammenhänge sind **korrelativ**, keine Kausalaussage möglich.

Die vollständigen Matrizen (Koeffizienten + Holm-adjustierte p-Werte) liegen als CSV bei, die Heatmaps als editierbares SVG/PDF/PNG.

7. Darauf aufbauend weitere Visualisierungen erstellen



8. Ergebnisse in verschiedenen Formaten herunterladen



Corr pearson

Bild · PNG

Herunterladen

▼



Corr pearson

Dokument · PDF

Herunterladen

▼



Corr pearson

Bild · SVG

Herunterladen

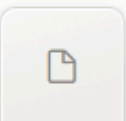


Corr spearman

Bild · PNG

Herunterladen

▼



Corr spearman

Dokument · PDF

Herunterladen

▼



Corr spearman

Bild · SVG

Herunterladen



Correlation pearson matrix

Tabelle · CSV

Herunterladen

▼



Correlation spearman matrix

Tabelle · CSV

Herunterladen

▼



Correlation pearson adjusted p...

Tabelle · CSV

Herunterladen

▼



Correlation spearman adjusted...

Tabelle · CSV

Herunterladen

▼

 Alles herunterladen



9. Mit Flourish verbinden

verbinde dich mit flourish

Geladene Tools, flourish-Integration verwendet >

Die Verbindung zu Flourish steht — ich bin mit deinem Account verbunden und sehe bereits 35 Visualisierungen dort, darunter einige, die offenbar schon zu genau diesem KI-Nutzungs-Datensatz gehören (z. B. "Nutzung von KI-Tools: Studierende und Lehrende", "Bubble Chart: KI-Tools bei Studierenden und Lehrenden").

Was möchtest du damit machen — eine neue Visualisierung aus unseren Ergebnissen erstellen (z. B. die Korrelationsmatrix oder den Gruppenvergleich als interaktive Flourish-Grafik), oder eine der bestehenden Visualisierungen ansehen/bearbeiten?

10. Visualisierungen für Flourish erzeugen

Geladene Tools, flourish-Integration verwendet >

Ich lade jetzt beide Matrizen (Pearson und Spearman) direkt in Flourish als interaktive Heatmaps.

5 Befehle ausgeführt, flourish-Integration verwendet, datei erstellt >

Fertig — die Daten liegen erfolgreich in Flourish. Du kannst die Visualisierung hier direkt weiterbearbeiten:

<https://app.flourish.studio/visualisation/30328530/edit>

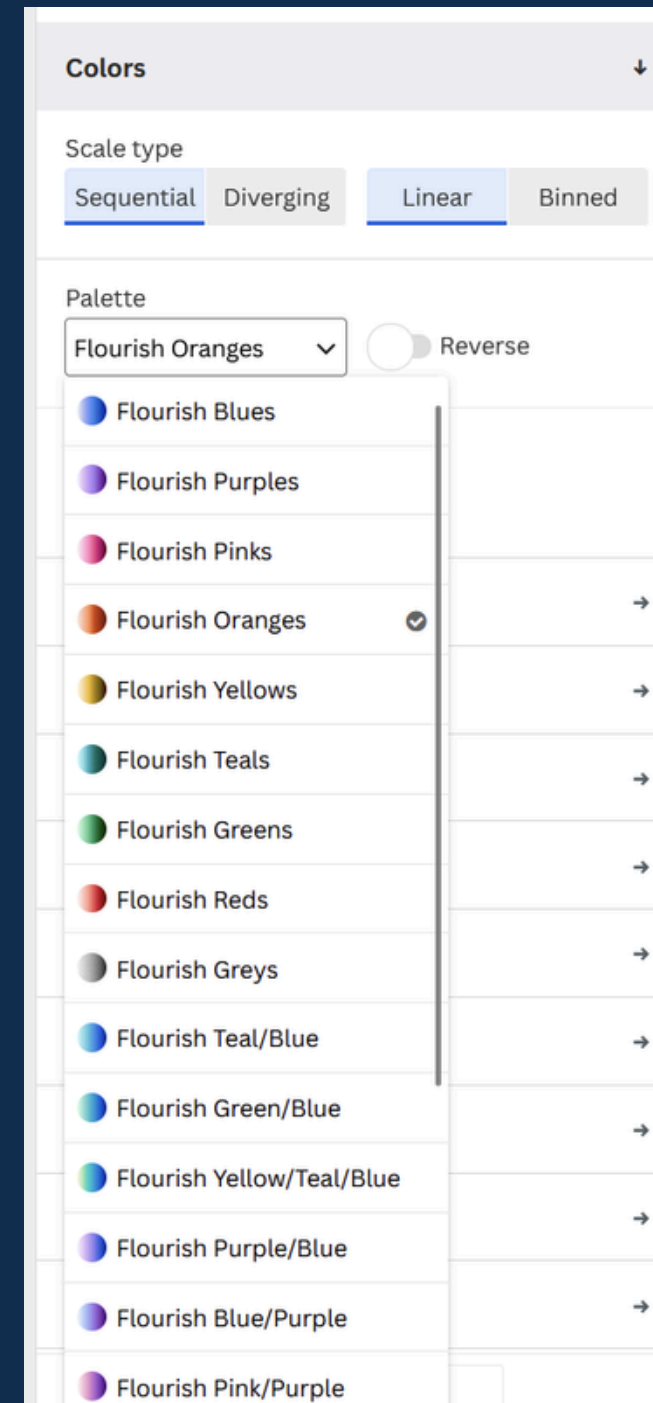
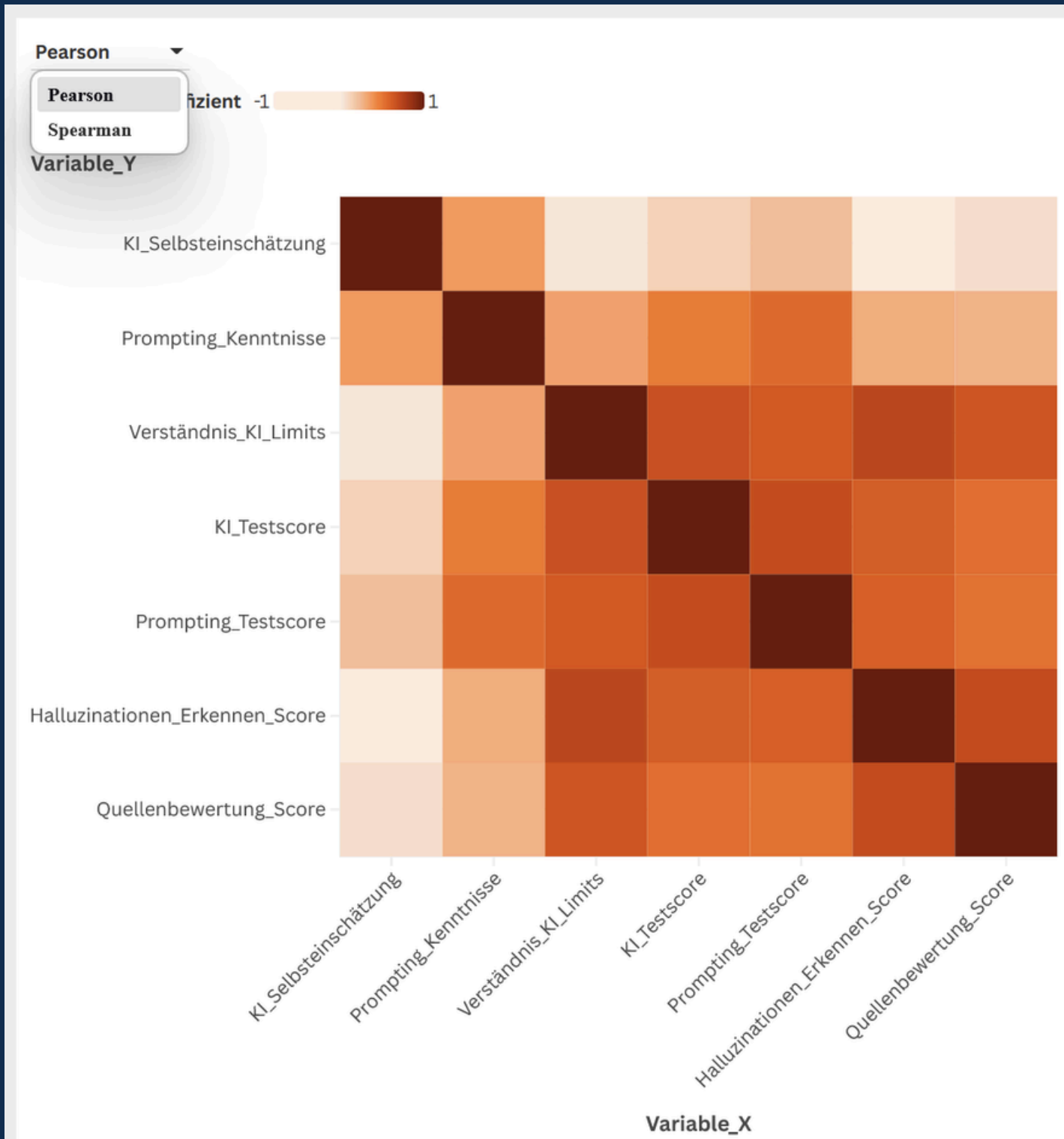
Was drin ist:

- Beide Matrizen (Pearson **und** Spearman) in einer Visualisierung, per **Filter „Methode“** umschaltbar — du kannst also direkt im Tool zwischen beiden Ansichten hin- und herklicken.
- X/Y-Achsen = die 7 Variablen, Farbwert = Korrelationskoeffizient.
- Im Popup jeder Zelle (Info-Feld) stehen zusätzlich n, der Holm-adjustierte p-Wert und ob das Ergebnis nach Korrektur signifikant ist.

Was du im Editor typischerweise noch anpassen willst:

- Farbskala auf eine divergierende Palette mit Mittelpunkt bei 0 setzen (Flourish nennt das im Reiter „Colour“ → "Diverging"), damit negative/positive Korrelationen klar unterschieden sind — Standard ist evtl. eine sequenzielle Skala.
- Wertebereich der Farbskala manuell auf **-1 bis +1** fixieren (nicht automatisch skalieren lassen), sonst wird der Kontrast verzerrt dargestellt.
- Zellen mit Zahlen beschriften (Reiter „Cell labels“ → an, **Korrelationskoeffizient** als Quelle).

11. Gewünschte Darstellungsform und Ergebnisse in Flourish auswählen und direkt weiterbearbeiten.



Vielen Dank für deine Aufmerksamkeit!

Ich bin Danny Walther.

Ich unterstütze Hochschulen, Bildungs- und Forschungseinrichtungen dabei, Künstliche Intelligenz praxisnah und wissenschaftlich fundiert einzusetzen.

Ich biete KI-Workshops, Webinare und Impulsvorträge.

Meine Schwerpunkte sind:
Daten- und Literaturrecherche mit KI,
Visualisierungen mit KI sowie
der Einsatz von KI in Lehre und Prüfungen.

Danny Walther
www.ki-recherchen.de
www.linkedin.com/in/ki-recherchen

