



Crashkurs deskriptive Statistik – Teil 2

Lern' die wichtigsten Grundlagen in kürzester Zeit!

Ablauf Nachmittag

Zusammenhangsmaße:

- 14:00 – 14:50 **Chi-Quadrat-Koeffizient, Rangkorrelation Spearman**
- 14:50 – 15:00 Pause
- 15:00 – 15:50 **Kovarianz & Korrelationskoeffizient r (Bravais-Pearson)**
- 15:50 – 16:10 Pause
- 16:10 – 17:00 **SPSS**

Zusammenhangsmaße

Übersicht

DESKRIPTIVE STATISTIK
Beschreibung und Charakterisierung eines Datensatzes

LAGEMAßE
Wo ist das Zentrum des Datensatzes?
= Was ist der "typische" Wert?

MEDIAN
Unterteilt den Datensatz in die 50% kleineren und 50% größeren Werte
Ab Ordinalskala

MITTELWERT
Der "Durchschnitt"
Alle Skalen

QUANTILE
Ein bestimmter Anteil der Werte ist kleiner als das Quantil, der Rest größer
z. B. Q1 = 25%-Quantil

MODUS / MODALWERT
Häufigste Merkmalsausprägung
Alle Skalen

ZUSAMMENHANGSMAßE
Messen den Zusammenhang zwischen zwei Variablen

CHI-QUADRAT-KOEFFIZIENT
Zusammenhang zwischen nominalskalierten Variablen

r SP
Zusammenhang zwischen ordinalskalierten Variablen

KOVARIANZ
Zusammenhang zwischen metrischen Variablen
UNSTANDARDISIERTES Maß

r KORRELATIONSKOEFFIZIENT NACH BRAVAIS-PEARSON
Zusammenhang zwischen metrischen Variablen
STANDARDISIERTES Maß

STREUUNGSMAßE
Wie weit streuen die Daten?
= Wie sehr unterscheiden sich die Versuchspersonen?

SPANNWEITE
Metrische Skalen

VARIANZ
Wie sehr streuen die Werte?
Quadratisches Streuungsmaß
"Mutter" der Standardabweichung
Nicht zur Interpretation geeignet
Metrische Skalen

INTERQUARTILBEREICH (IQR)
Zwischen Q1 und Q3 liegt die Hälfte aller beobachteten Werte

STANDARDABWEICHUNG
Was ist die typische Abweichung vom Mittelwert?
Metrische Skalen

DIVERSES
v. a. grafische Darstellungen

ABSOLUTE UND KUMULIERTE HÄUFIGKEITEN
ABSOLUT: Wieviele Werte einer Merkmalsausprägung liegen vor?
KUMULIERT: Wieviele Werte haben sich bis zu einem bestimmten Punkt angesammelt?

BALKEN-, SÄULEN- UND KREISDIAGRAMM
Alle Skalen

HISTOGRAMM
Grafische Darstellung der Häufigkeiten bei metrischen Merkmalen

BOXPLOT
Grafische Darstellung von Schwerpunkt und Streuung in einem Mindestens Ordinalskala

Zusammenhangsmaße

Übersicht

- **Chi-Quadrat-Koeffizient:**
 - Zusammenhang zwischen zwei **nominalskalierten** Variablen
- **Rangkorrelations-Koeffizient:**
 - Zusammenhang zwischen zwei **ordinalskalierten** Variablen

Zusammenhangsmaße

Übersicht

- **Kovarianz:**

- Zusammenhang zwischen zwei metrischen Variablen – nicht normiert

- **Korrelationskoeffizient r :**

- Zusammenhang zwischen zwei metrischen Variablen – normiert!

Korrelation

Was? Wofür?

Korrelation

Was?

Zusammenhang zwischen
zwei Merkmalen

Synonyme:

Kovariation, Abhängigkeit,
Assoziation

Korrelation

Wofür?

- Um zu messen, ob zwischen zwei Variablen ein Zusammenhang vorliegt & wie stark dieser ist
- Verwendung in deskriptiver & schließender Statistik

Zusammenhangsmaße

Wann nehme ich welches?

- Viele verschiedene Zusammenhangsmaße
 - Nimm' das, was zum Skalenniveau deiner erhobenen Variablen passt!

Zusammenhangsmaße *und Skalenniveaus*

Nominalskala	Chi-Quadrat-Koeffizient Phi-Koeffizient Cramér's V
Ordinalskala	Rangkorrelationskoeffizient nach Spearman Kendalls Tau
Metrische Skalen	Kovarianz Korrelationskoeffizient r nach Bravais-Pearson

Korrelation & Kausalität

Wie hängen sie zusammen?

Korrelation & Kausalität

Beispiel

Negativer Zusammenhang zwischen Perfektionismus
und Selbstwertgefühl

Folgendes könnte zutreffen:

- Der Perfektionismus beeinflusst das Selbstwertgefühl
- Das Selbstwertgefühl beeinflusst den Perfektionismus
- Beide werden von einer dritten Variable beeinflusst (Scheinkorrelation!)
- Beide beeinflussen sich gegenseitig kausal

Korrelation & Kausalität

- Korrelationskoeffizienten zeigen nicht an, welche der Möglichkeiten die Richtige ist!
 - Korrelationen dürfen ohne Zusatzinformation NICHT kausal interpretiert werden!
- Methode zur Überprüfung von Kausalität:
 - Experiment

Kausalität

Bedingungen

- Variable A und Variable B **kovariieren**
- Variable **A tritt zeitlich vor B auf**
- **Alternative Erklärungen können ausgeschlossen werden**

Zusammenhänge

Die Maße im Einzelnen



Chi-Quadrat-Koeffizient x^2

Was? Wofür?

χ^2

Was?

Zusammenhangsmaß für zwei
nominalskalierte Merkmale

➤ Kann auch bei höher
skalierten Merkmalen
verwendet werden – diese
müssen dann in Kategorien
verpackt werden

χ^2

Wofür?

- Um festzustellen, ob zwischen zwei nominalskalierten Merkmalen ein Zusammenhang besteht
 - Schwierig zu interpretieren, da nicht normiert

Chi-Quadrat-Koeffizient

Beispiel

Chi-Quadrat-Koeffizient

Beispiel

Zusammenhang zwischen dem Geschlecht
& der Wahl einer Dating-App

Stichprobe:

- 20 Frauen & 20 Männer

Dating-Apps/Singlebörsen:

- Parship & Tinder

Chi-Quadrat-Koeffizient

Beispiel

Kreuz- oder
Kontingenz-
tabelle

	Weiblich	Männlich	Zeilensummen:
Parship	14	3	17
Tinder	6	17	23
Spaltensummen:	20	20	N = 40

Chi-Quadrat-Koeffizient

Berechnung

Chi-Quadrat-Koeffizient

Grundprinzip

Vergleich der

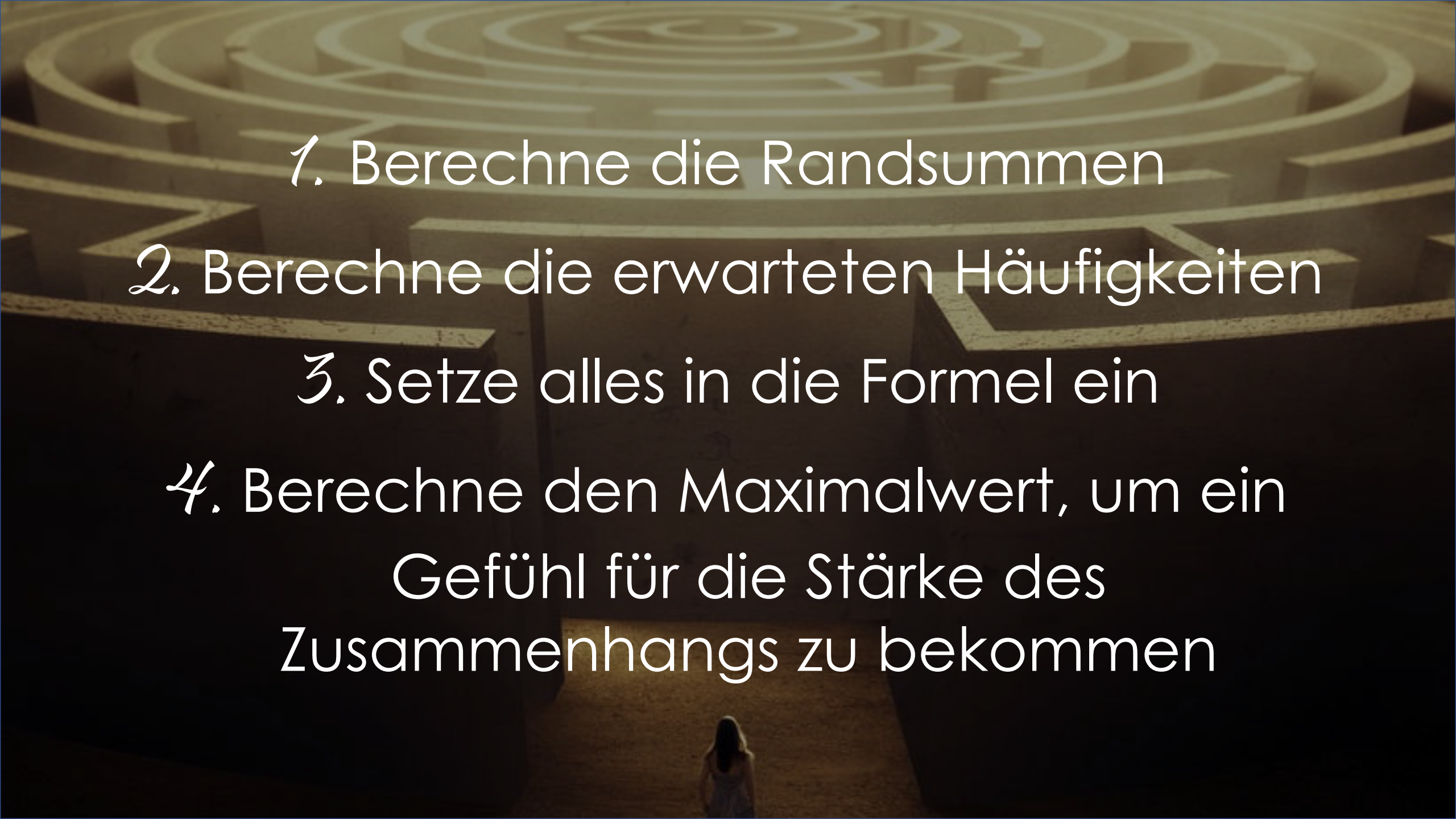
- tatsächlich beobachteten Häufigkeiten
mit den
- bei Unabhängigkeit erwarteten Häufigkeiten
 - Unabhängigkeit = kein Zusammenhang

Chi-Quadrat-Koeffizient

Formel

$$x^2 = \sum_{\text{alle Zellen}} \frac{(\text{beobachtete Häufigkeit} - \text{erwartete Häufigkeit})^2}{\text{erwartete Häufigkeit}}$$

Wenn es (signifikante) Abweichungen zwischen den tatsächlich beobachteten & den theoretisch erwarteten Häufigkeiten gibt, liegt ein Zusammenhang vor

- 
- A person is seen from behind, walking through a large, circular maze. The maze is constructed from light-colored concrete walls on a dark floor. The person is in the center of the maze, and the path leads towards a bright opening at the far end. The lighting is dramatic, with strong shadows and highlights.
1. Berechne die Randsummen
 2. Berechne die erwarteten Häufigkeiten
 3. Setze alles in die Formel ein
 4. Berechne den Maximalwert, um ein Gefühl für die Stärke des Zusammenhangs zu bekommen

A large, circular maze with a person standing in the center, looking out at the complex paths. The maze is constructed from light-colored concrete or stone walls, creating a series of concentric and irregular paths. The person is a small figure in the distance, wearing a light-colored dress, standing in a narrow path that leads towards the center of the maze. The lighting is dramatic, with strong shadows and highlights, creating a sense of depth and mystery.

1. Schritt:

Berechne die
Randsummen

Randsummen

	Weiblich	Männlich	Zeilensummen:
Parship	14	3	17
Tinder	6	17	23
Spaltensummen:	20	20	N = 40

A large, complex maze with a person standing in the center, looking out. The maze is made of light-colored stone or concrete walls, and the person is a small figure in the distance, wearing a white dress. The lighting is dramatic, with strong shadows and highlights, creating a sense of depth and mystery.

2. Schritt:

Berechne die
erwarteten Häufigkeiten

Erwartete Häufigkeiten

Formel

Erwartete Häufigkeit \rightarrow

$$m_{ij} = \frac{n_{i.} \cdot n_{.j}}{N}$$

Zeilensumme \rightarrow $n_{i.}$ $n_{.j}$ \leftarrow Spaltensumme

\uparrow
Gesamtstichprobe

Erwartete Häufigkeiten

Formel

$$m_{ij} = \frac{n_{i.} \cdot n_{.j}}{N}$$

Übersetzt:

- Multipliziere für jede Zelle die dazugehörigen Randhäufigkeiten miteinander
- Teile anschließend durch die Gesamtstichprobe
 - Weichen die erwarteten von den beobachteten Häufigkeiten ab, liegt ein Zusammenhang vor!

Erwartete Häufigkeiten

Berechnung

$$m_{ij} = \frac{n_{i.} \cdot n_{.j}}{N} = \frac{20 \cdot 17}{40} = 8.5$$

	Weiblich	Männlich	Zeilensummen:
Parship	14	3	17
Tinder	6	17	23
Spaltensummen:	20	20	N = 40

Erwartete Häufigkeit sind 8.5, es wurden jedoch 14 beobachtet

A large, circular maze with a person standing in the center, looking out at the complex paths. The maze is constructed from light-colored stone or concrete walls, creating a series of concentric and irregular paths. The lighting is dramatic, with the center of the maze being brightly lit, while the outer edges are in deep shadow. A person with long dark hair, wearing a light-colored dress, stands in the center of the maze, looking out at the complex paths. The overall atmosphere is one of mystery and complexity.

3. Schritt:

Setze alles in die
Formel ein

Chi-Quadrat-Koeffizient

Berechnung

$$\chi^2 = \frac{(14 - 8.5)^2}{8.5} + \frac{(6 - 11.5)^2}{11.5} + \frac{(3 - 8.5)^2}{8.5} + \frac{(17 - 11.5)^2}{11.5} = 12.377$$

Was sagt uns das?

- Nicht wirklich viel...
 - **Maximalwert berechnen**, um einen Anhaltspunkt für die Größe zu bekommen

A large, circular maze with a person standing in the center, looking out at the complex paths. The maze is constructed from light-colored concrete or stone walls, creating a complex network of paths and dead ends. The person is a small figure in the distance, wearing a light-colored dress, standing in a wide, open path that leads towards the center of the maze. The lighting is warm and golden, suggesting a sunset or sunrise, with long shadows cast across the maze's walls.

4. Schritt:

Berechne den
Maximalwert

Chi-Quadrat-Koeffizient

Maximalbereich

$$0 \leq x^2 \leq x_{max}^2 = N \cdot (M - 1)$$

M = der kleinere Wert
der Anzahl der Zeilen und Spalten
bzw. der Ausprägungen der Variablen

Chi-Quadrat-Koeffizient

Maximalbereich

$$N \cdot (M - 1)$$

$$= 40 \cdot (2 - 1) = 40$$

Maximalwert ist hier 40

- Unser Ergebnis war 12.377
 - Also ein eher schwacher Zusammenhang



Interpretation:

Wie sag' ich's Tante Erna?



Interpretation

Chi-Quadrat

„Es gibt einen schwachen Zusammenhang zwischen dem Geschlecht und der Nutzung einer Dating-App.“

„Frauen neigen eher zu Parship, Männer eher zu Tinder.“

➤ Letzteres lässt sich aus der Tabelle herauslesen!

Rangkorrelations-Koeffizient R_{SP}

Was? Wofür?

R_{SP}

Was?

Normiertes
Zusammenhangsmaß für zwei
mindestens ordinalskalierte
Merkmale

- Kann nur Werte zwischen
-1 und +1 annehmen

R_{SP}

*Stärke des
Zusammenhangs*

0.1 schwach

0.3 mittel

0.5 stark

➤ Betrag: gilt für positive oder negative Werte

R_{SP} Wofür?

- Um **lineare Zusammenhänge** zwischen **Rangplätzen** zu quantifizieren
- Auch für metrische Merkmale einsetzbar
 - Robuster bei Ausreißern

Rangkorrelations-Koeffizient

Beispiel

Rangkorrelations-Koeffizient

Beispiel

Beurteilung der Arbeitsmotivation durch zwei Teamleiter

Skala: 1 (kaum) – 9 (sehr stark)

Stichprobe:

- 4 Mitarbeiterinnen in einer Kurklinik

Rangkorrelations-Koeffizient

Beispiel

Versuchs- person Nr.	Beurteilung Herr Huber		Beurteilung Herr Müller	
1: Berta	3		4	
2: Susi	9		7	
3: Chantal	6		6	
4: Alexa	7		5	

Rangkorrelations-Koeffizient

Berechnung



1. Schritt:

Ränge zuordnen

2. Schritt:

Alles in die Formel
einsetzen

Rangkorrelations-Koeffizient

Ränge zuordnen

Versuchsperson Nr.	Beurteilung Herr Huber	Rangplätze	Beurteilung Herr Müller	Rangplätze
1: Berta	3	4	4	4
2: Susi	9	1	7	1
3: Chantal	6	3	6	2
4: Alexa	7	2	5	3

Rangkorrelations-Koeffizient

Formel

Rangkorrelations-Koeffizient

Formel

$$r_{SP} = \frac{\sum_{i=1}^n (rg(x_i) - \overline{rg_x})(rg(y_i) - \overline{rg_y})}{\sqrt{\sum_{i=1}^n (rg(x_i) - \overline{rg_x})^2} \cdot \sqrt{\sum_{i=1}^n (rg(y_i) - \overline{rg_y})^2}}$$

Oder lieber ein wenig einfacher mit SPSS...

Rangkorrelations-Koeffizient

Beispiel

Korrelationen

			Beurt. Arbeitsmotiv ation A	Beurt. Arbeitsmotiv ation B
Spearman-Rho	Beurt. Arbeitsmotivation A	Korrelationskoeffizient	1,000	,800
		Sig. (2-seitig)	.	,200
		N	4	4
	Beurt. Arbeitsmotivation B	Korrelationskoeffizient	,800	1,000
		Sig. (2-seitig)	,200	.
		N	4	4



Interpretation:

Wie sag' ich's Tante Erna?



Interpretation *Rangkorrelation*

„Es gibt einen starken Zusammenhang zwischen den Beurteilungen der beiden Teamleiter.“

Oder:

„Es gibt eine starke Übereinstimmung zwischen den Beurteilungen der beiden Teamleiter.“

Kovarianz

Was? Wofür?

Kovarianz

Was?

Nicht normiertes
Zusammenhangsmaß
für metrische Variablen

➤ „Mutter“ des
Korrelationskoeffizienten r

Kovarianz

Wofür?

- Um Zusammenhänge zwischen metrischen Variablen zu quantifizieren

Kovarianz

Nachteil

- Maßstabsabhängig
 - daher kaum zur Interpretation geeignet!
- Kann je nach Skala sehr groß oder klein sein
- Lösung:
 - Korrelationskoeffizient r nach Bravais-Pearson

Kovarianz

Beispiel Maßstab

Zusammenhang zwischen Kopfumfang und Intelligenz

Versuchs- person Nr.	Kopfumfang in cm	Intelligenz
1	55	102
2	60	117
3	57	120

Kovarianz

Beispiel Maßstab

Versuchs- person Nr.	Kopfumfang in cm	Intelligenz
1	55	102
2	60	117
3	57	120

Kovarianz = 17 (in cm gemessen)

Kovarianz

Beispiel Maßstab

Versuchs- person Nr.	Kopfumfang in m	Intelligenz
1	0,55	102
2	0,60	117
3	0,57	120

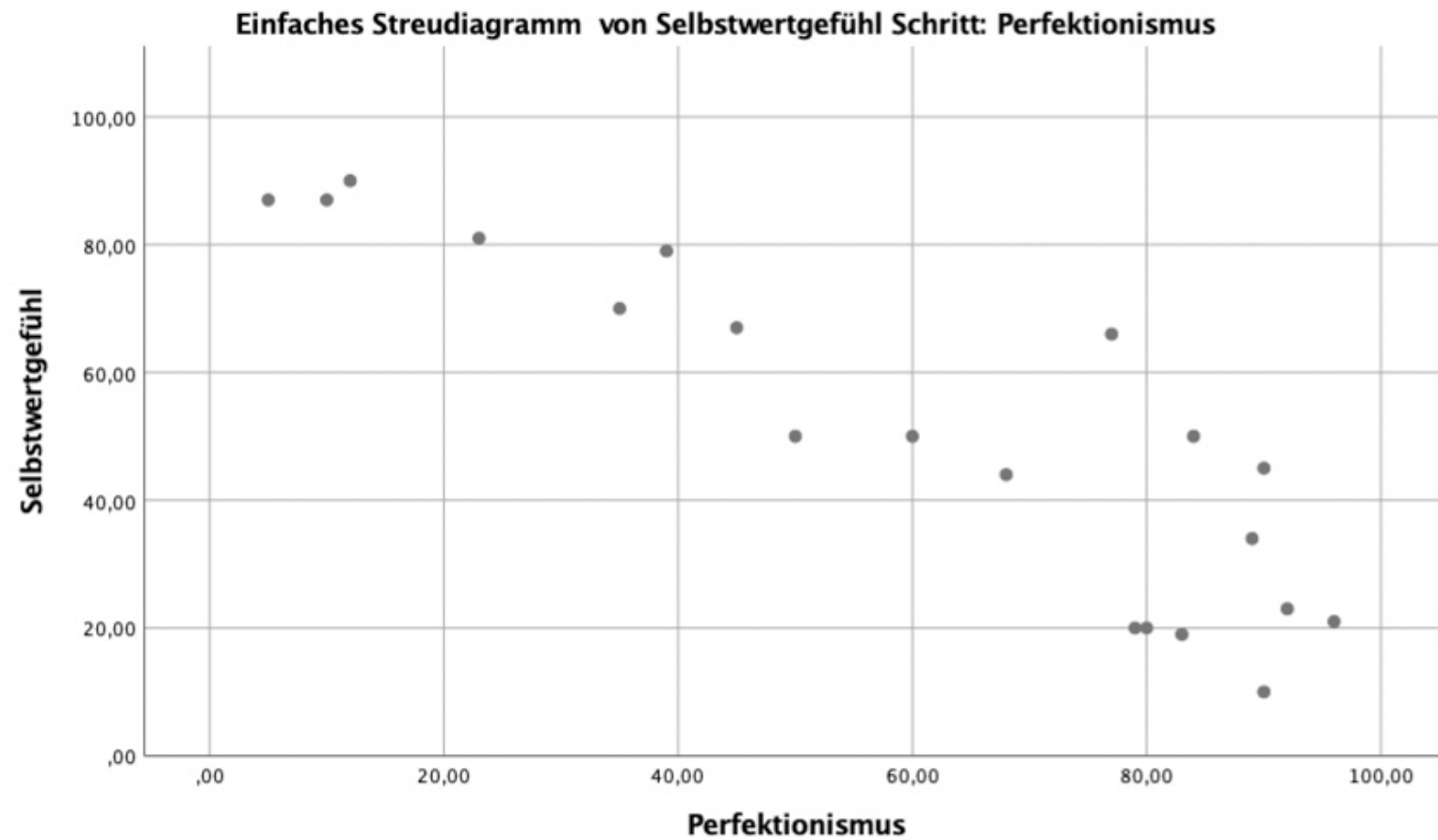
Kovarianz = 0,17 (in m gemessen)

Kovarianz

Richtung des Zusammenhangs

Kovarianz

grafisch



Kovarianz

Richtung

Positive Kovarianz:

- Gleichgerichtete Tendenz
 - Hohe Werte in dem einen Merkmal gehen mit hohen Werten in dem anderen einher (bzw. niedrige mit niedrigen)

Negative Kovarianz:

- Entgegengesetzte Tendenz
 - Hohe Werte in dem einen Merkmal gehen mit niedrigen Werten in dem anderen einher

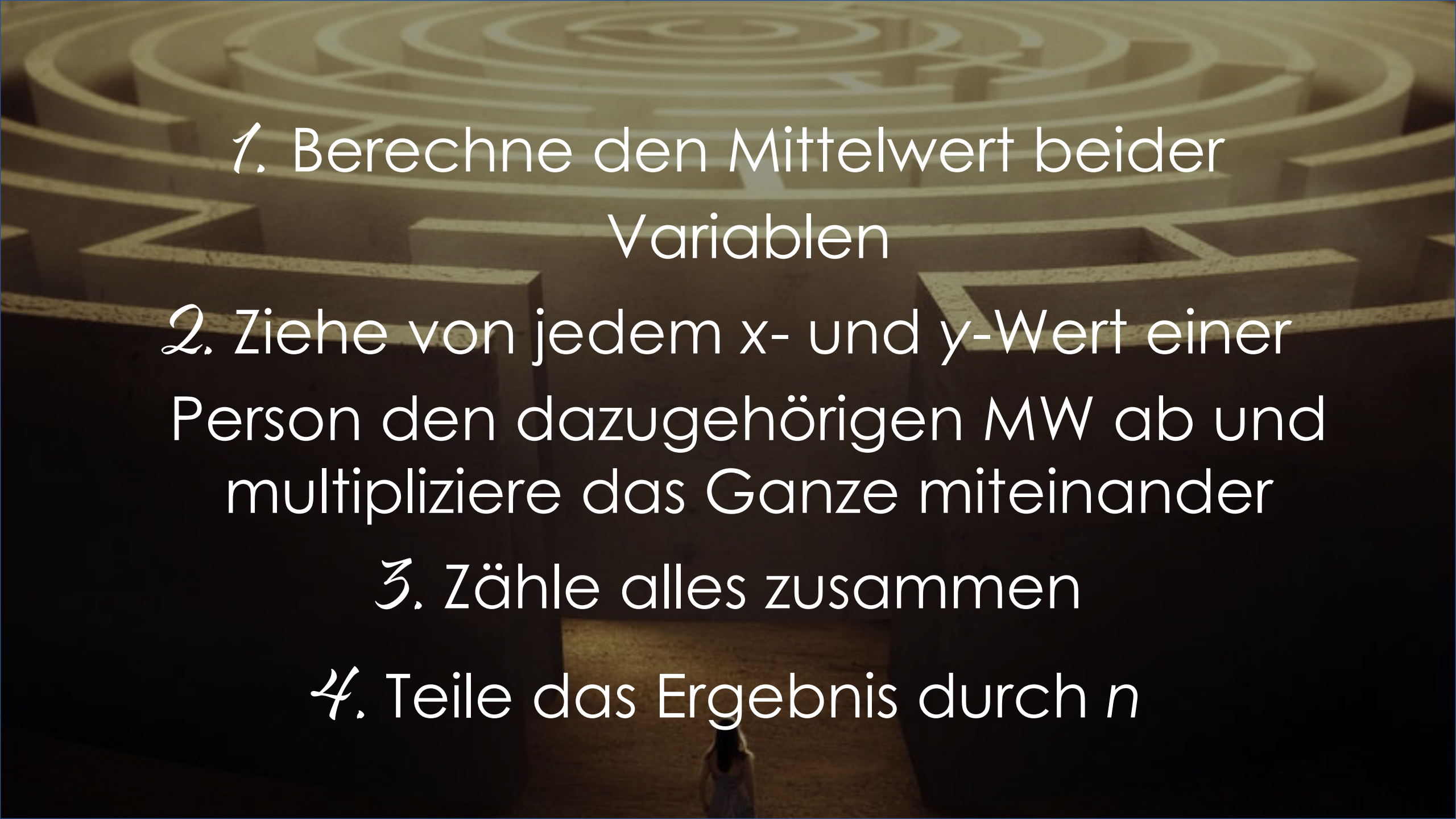
Kovarianz

Berechnung

Kovarianz

Formel

$$s_{xy} := \frac{1}{n} \cdot [(x_1 - \bar{x})(y_1 - \bar{y}) + \dots + (x_n - \bar{x})(y_n - \bar{y})] = \frac{1}{n} \cdot \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

- 
1. Berechne den Mittelwert beider Variablen
 2. Ziehe von jedem x - und y -Wert einer Person den dazugehörigen MW ab und multipliziere das Ganze miteinander
 3. Zähle alles zusammen
 4. Teile das Ergebnis durch n

Kovarianz

Mit SPSS

Kovarianz

Mit SPSS

Korrelationen

		Perfektionis- mus	Selbstwertgef ühl
Perfektionismus	Korrelation nach Pearson	1	-,891**
	Signifikanz (2-seitig)		,000
	Quadratsummen und Kreuzprodukte	17766,550	-13701,550
	Kovarianz	935,082	-721,134
	N	20	20
Selbstwertgefühl	Korrelation nach Pearson	-,891**	1
	Signifikanz (2-seitig)	,000	
	Quadratsummen und Kreuzprodukte	-13701,550	13324,550
	Kovarianz	-721,134	701,292
	N	20	20

** . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Korrelationskoeffizient r

Was? Wofür?

r
Was?

Maßstabsunabhängiges,
normiertes

Zusammenhangsmaß
für metrische Variablen

- Kann nur Werte zwischen
-1 und +1 annehmen

r

Wofür?

- Misst, wie stark zwei Merkmale miteinander kovariieren
 - Kann nur bei linearen Zusammenhängen angewendet werden!

Hinter die Löffelchen schreiben:

Wenn $r = 0$ ist,
kann dennoch ein
nicht-linearer
Zusammenhang vorliegen!

r

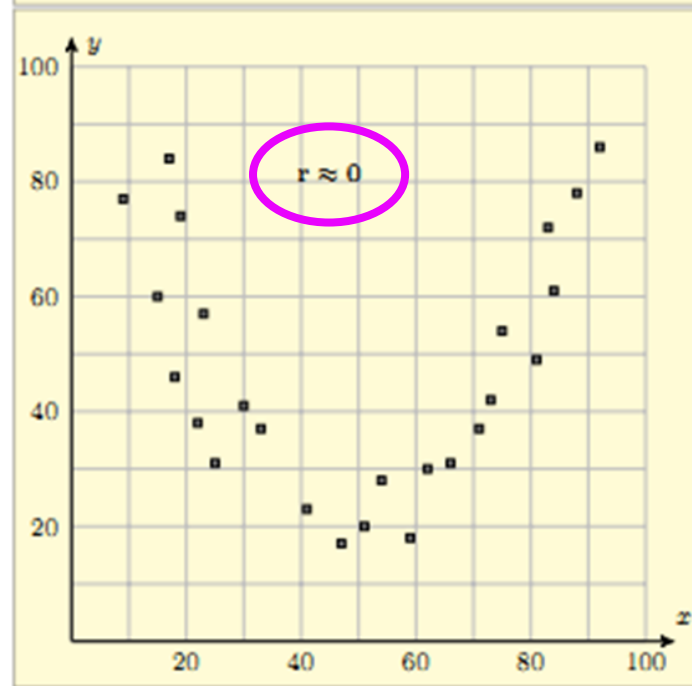
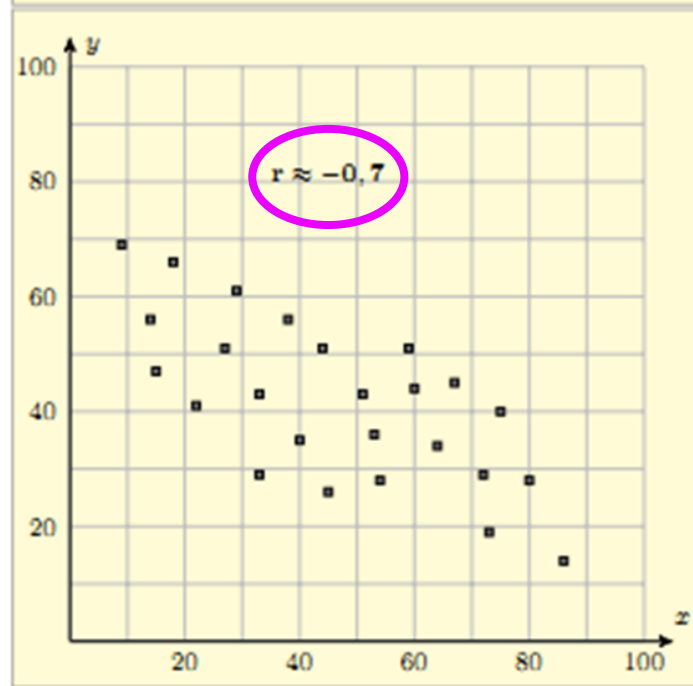
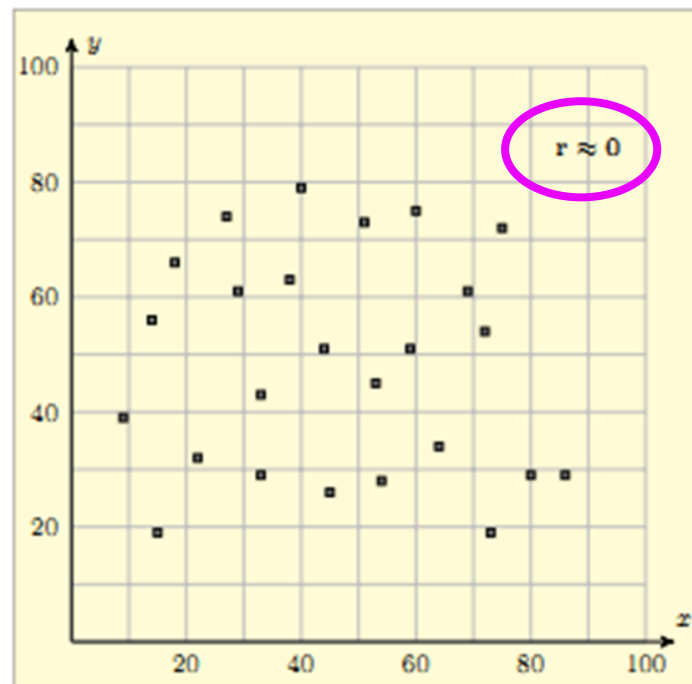
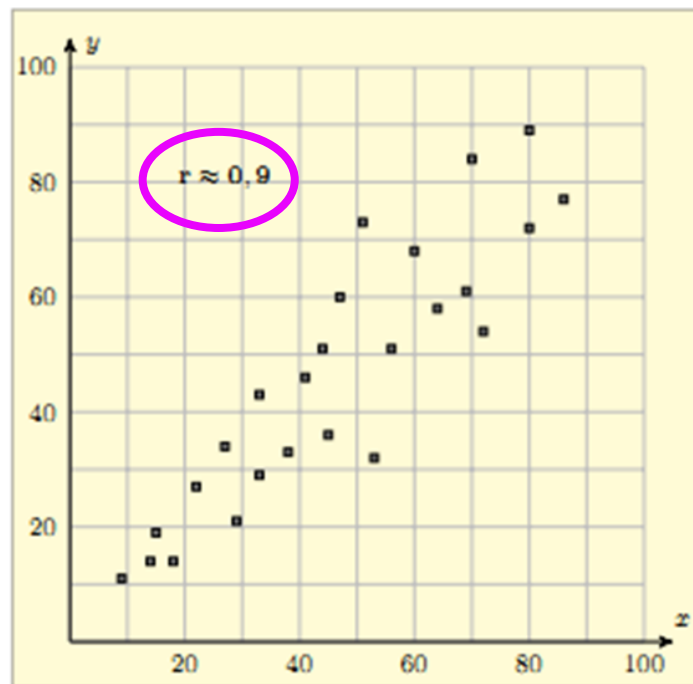
*Stärke des
Zusammenhangs*

0.1 schwach

0.3 mittel

0.5 stark

➤ Betrag: gilt für positive oder negative Werte



Korrelationskoeffizient r

Richtung

Positiver Zusammenhang: r größer 0

Negativer Zusammenhang: r kleiner 0

Unkorreliertheit (linear!): r gleich 0

Korrelationskoeffizient r

Berechnung

Korrelationskoeffizient r

Berechnung

$$r := \frac{s_{xy}}{s_x \cdot s_y}$$

- Kovarianz wird durch das Produkt der Standardabweichungen der beiden beteiligten Variablen geteilt
 - Immer gleiches Vorzeichen wie Kovarianz!

Korrelationskoeffizient r

Mit SPSS



Korrelationskoeffizient r

Mit SPSS

Korrelationen

		Perfektionismus	Selbstwertgefühl
Perfektionismus	Korrelation nach Pearson	1	-,891**
	Signifikanz (2-seitig)		,000
	N	20	20
Selbstwertgefühl	Korrelation nach Pearson	-,891**	1
	Signifikanz (2-seitig)	,000	
	N	20	20

** . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.



Interpretation:

Wie sag' ich's Tante Erna?



Interpretation

Korrelationskoeffizient

„Es gibt einen starken negativen Zusammenhang zwischen Perfektionismus und dem Selbstwertgefühl: je niedriger das Selbstwertgefühl eines Menschen, desto mehr neigt er zum Perfektionismus.“

„Und natürlich umgekehrt: je höher das Selbstwertgefühl eines Menschen, desto weniger neigt er zum Perfektionismus.“

SPSS

Praktische Anwendung



SPSS

Wo findet man Datensätze?

- **Windows:** C:/Programme/IBM/SPSS/Statistics/26 (oder eine andere Version)/Samples/German
- **Mac:** Im Finder: Applications oder Programme/IBM/SPSS/Statistics/26 (oder eine andere Version)/Samples/German
- **Datensätze von Andy Field:**
<https://edge.sagepub.com/field5e/student-resources/datasets>

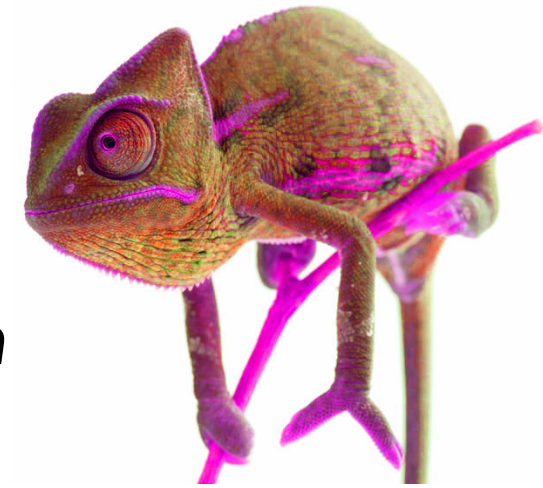
Geschafft!
Zeit für fette Belohnung...



Dankeschön fürs Mitmachen!
Und jetzt heißt's üben!



Verwendete Literatur



- Bortz, J., & Schuster, C. (2017). *Statistik für Human- und Sozialwissenschaftler*. Berlin: Springer.
- Bühner, M., & Ziegler, M., (2008): *Statistik für Psychologen und Sozialwissenschaftler*. München [u.a.]: Pearson.
- Eid, M., Gollwitzer, M., & Schmitt, M. (2010): *Statistik und Forschungsmethoden*. Lehrbuch ; mit Online-Materialien. 1. Aufl. Weinheim [u.a.]: Beltz.
- Field, A. (2018). *Discovering Statistics using IBM SPSS Statistics*. London: SAGE.
- Sedlmeier, P., & Renkewitz, F. (2018). *Forschungsmethoden und Statistik für Psychologen und Sozialwissenschaftler*. München: Pearson.

Nachweis *Bilder*

- Alle Bilder © 123RF.com
- Ausnahme: Rosa Chamäleon-Kopf – von Unsplash