Hochschule Harz
Hochschule für angewandte Wissenschaften

Probeklausur Statistik I
Sommersemester 2026
Bbgl. Bachelor BWL

Christian Reinboth

Aufgabe I: Sind diese Aussagen richtig oder falsch?

1) Der Modus lässt sich nur bestimmen, wenn eine biomodale Verteilung vorliegt.

2) Zwei Teilgesamtheiten einer Grundgesamtheit können sich niemals überschneiden.

3) Qualitative und quantitative Methoden der Datenerhebung und Datenanalyse können grundsätzlich nicht miteinander kombiniert werden.

4) Der Median ist derjenige Wert, der genau in der Mitte der geordneten Verteilung liegt.

5) Der Interquartilsabstand (IQR) ist der Abstand zwischen oberem und unterem Quartil.

6) Die kumulierte empirische Verteilungsfunktion lässt sich grafisch als Treppenfunktion darstellen.

7) Die Größe der Box im Box-Plot wird durch das obere sowie das untere Quartil definiert.

8) Die Berechnung der Varianz setzt mindestens ordinalskalierte Daten voraus.

9) Objektivität, Reliabilität und Validität sind die einzigen relevanten Kriterien für die Bewertung der Güte erhobener Daten.

10) Rangplatzbindungen erhöhen die Aussagekraft des Spearman-Rangkorrelationskoeffizienten.

Aufgabe I: Sind diese Aussagen richtig oder falsch?

1) Der Modus lässt sich nur bestimmen, wenn eine biomodale Verteilung vorliegt.

2) Zwei Teilgesamtheiten einer Grundgesamtheit können sich niemals überschneiden.

3) Qualitative und quantitative Methoden der Datenerhebung und Datenanalyse können grundsätzlich nicht miteinander kombiniert werden.

4) Der Median ist derjenige Wert, der genau in der Mitte der geordneten Verteilung liegt.

5) Der Interquartilsabstand (IQR) ist der Abstand zwischen oberem und unterem Quartil.

6) Die kumulierte empirische Verteilungsfunktion lässt sich grafisch als Treppenfunktion darstellen.

7) Die Größe der Box im Box-Plot wird durch das obere sowie das untere Quartil definiert.

8) Die Berechnung der Varianz setzt mindestens ordinalskalierte Daten voraus.

9) Objektivität, Reliabilität und Validität sind die einzigen relevanten Kriterien für die Bewertung der Güte erhobener Daten.

10) Rangplatzbindungen erhöhen die Aussagekraft des Spearman-Rangkorrelationskoeffizienten.

Aufgabe II: Grafische Darstellungsformen

Eine Umfrage unter Professor*innen am Fachbereich Biowissenschaften einer kleineren Universität erbrachte folgende Angaben zur Anzahl der Jahre an persönlicher Forschungserfahrung.

Anonymisierte Teilnehmerkennung	Bisherige Forschungs-erfahrung (in Jahren)
OV37KS	17
OV28ER	8
FG13WS	3
DV66AZ	21
TF45WR	12
GH73TG	20
UF84FB	17
OV33ZU	16
ED85FT	7
FV75DD	5

Zeichnen Sie einen erweiterten Box-Plot für die Verteilung der Jahresangaben und benennen Sie die hierfür erforderlichen Größen.

Wie viel Forschungserfahrung müsste ein/e zusätzliche/r (!) Professor*in aufweisen, um im Box-Plot als Ausreißer im oberen Bereich sichtbar zu werden?

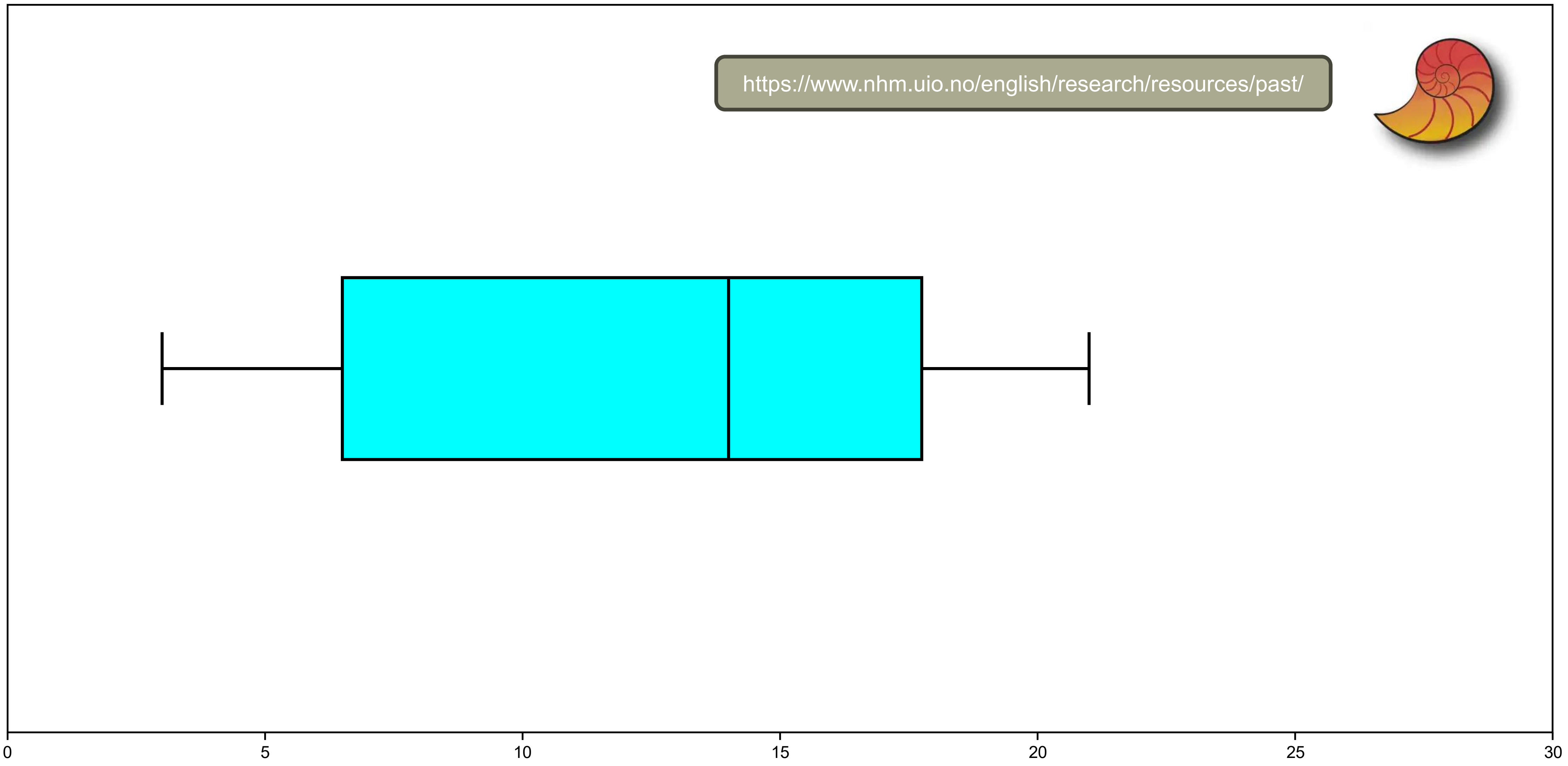
Aufgabe II: Konstruktion des Box-Plots

Bildung der geordneten Wertereihe: 3, 5, 7, 8, 12, 16, 17, 17, 20, 21

$x_{0,25}$ → 0,25 * 10 = 2,5 → Wert an der 3. Stelle der geordneten Verteilung → 7
$x_{0,50}$ → 0,50 * 10 = 5,0 → Mittel der Werte an der 5. und 6. Stelle → (12 + 16) / 2 = 14
$x_{0,75}$ → 0,75 * 10 = 7,5 → Wert an der 8. Stelle der geordneten Verteilung → 17

IQR = 17 – 7 = 10
1,5-facher IQR = 15

Die Box läuft damit von 7 bis 17, der Median wird bei 14 eingezeichnet. Der untere Zaun könnte minimal bis 7-15 = -8 laufen, der kleinste Wert in diesem Bereich ist die 3. Nach unten gibt es somit keine Ausreißer oder Extremwerte. Der obere Zaun könnte maximal bis 17 + 15 = 32 laufen, der größte Wert in diesem Bereich ist die 21. Auch nach oben gibt es somit keine Ausreißer oder Extremwerte.


https://www.nhm.uio.no/english/research/resources/past/
0
5
10
15
20
25
30

Aufgabe II: Zusatzfrage (für die 1,0-Kandidat*innen)

Wie viel Forschungserfahrung müsste ein/e zusätzliche/r (!) Professor*in aufweisen, um im Box-Plot als Ausreißer im oberen Bereich sichtbar zu werden?

<u>Einfache Lösung ohne Berücksichtigung der Veränderungen in der Größe des Datensatzes:</u>

Ein/e solche/r Professor*in müsste mindestens 33 (oder zumindest mehr als 32) Jahre an Forschungserfahrung aufweisen, um als Ausreißer im oberen Bereich sichtbar zu werden (also außerhalb des aktuellen Zauns liegen).

<u>Komplexe Lösung unter Berücksichtigung der Veränderungen in der Größe des Datensatzes (da ja eine Person zusätzlich aufgenommen wird):</u>

Bei einem zusätzlichen Professor hätte der Datensatz 11 Werte, der zusätzliche Wert soll am oberen Ende liegen. Die neue geordnete Reihe sieht somit wie folgt aus:

3, 5, 7, 8, 12, 16, 17, 17, 20, 21, a

Unabhängig von der Größe von a lassen sich nun neue Werte für die Identifikation von Ausreißern bestimmen.

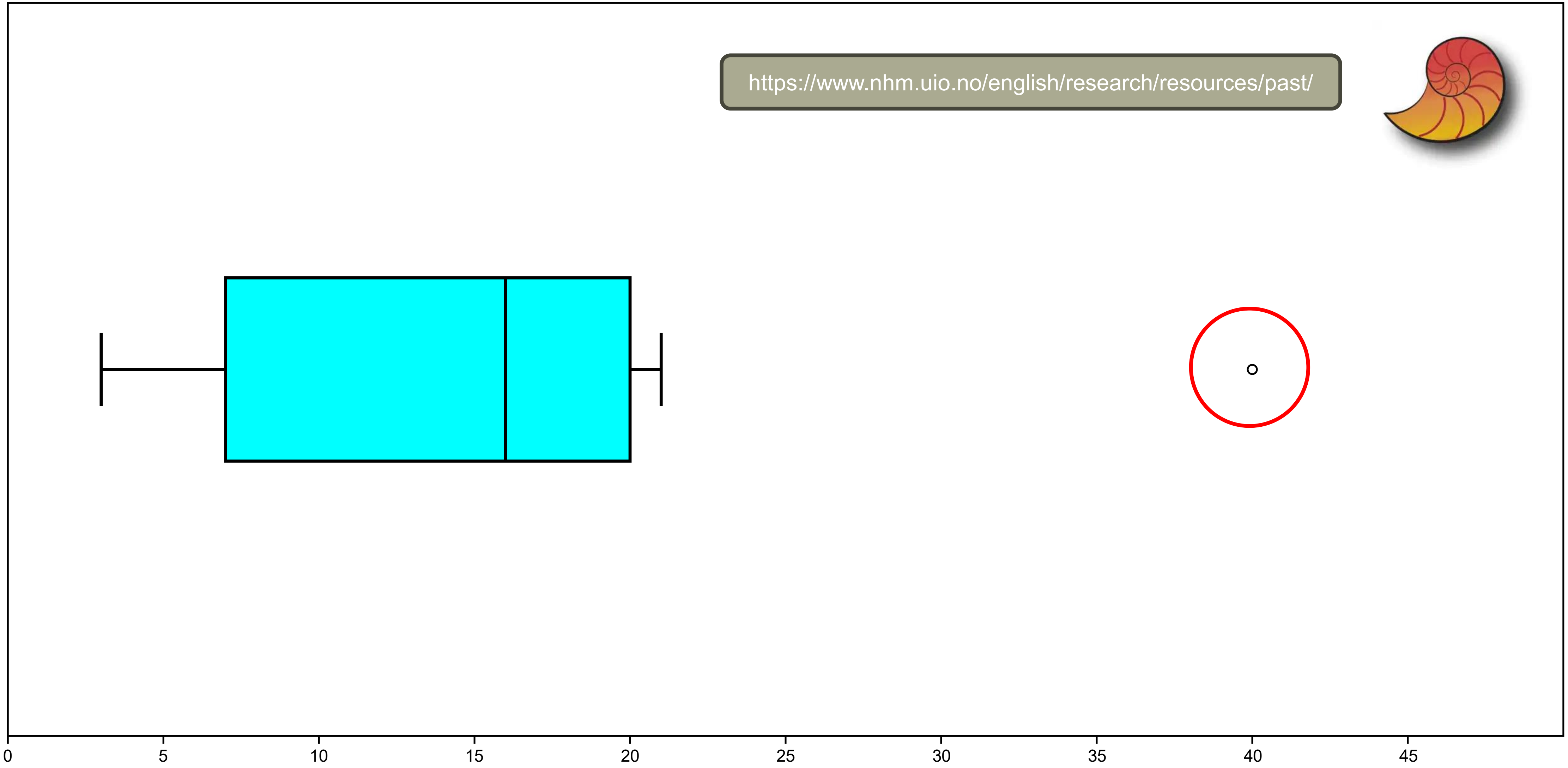
Aufgabe II: Zusatzfrage (für die 1,0-Kandidat*innen)

$X_{0,25} \rightarrow 0{,}25 * 11 = 2{,}75 \rightarrow$ Wert an der 3. Stelle der geordneten Verteilung $\rightarrow 7$
$X_{0,75} \rightarrow 0{,}75 * 11 = 8{,}25 \rightarrow$ Wert an der 9. Stelle der geordneten Verteilung $\rightarrow 20$

IQR = 20 – 7 = 13
1,5-facher IQR = 19,5

Der obere Zaun würde also maximal bis (20 + 19,5) = 39,5 laufen. Ein zusätzlicher (!) Professor müsste damit mindestens 40 (oder zumindest mehr als 39,5) Jahre an Forschungserfahrung aufweisen, um als Ausreißer im oberen Bereich sichtbar zu werden.

Die Korrektheit dieser Überlegung lässt sich durch die Erweiterung des Datensatzes um einen 11. Wert von 40 testen. Weist der Box-Plot, wenn er mit PAST generiert wird, nun wirklich einen Ausreißer auf?



https://www.nhm.uio.no/english/research/resources/past/

Aufgabe III: Verteilungsparameter

Die schon aus der vorherigen Aufgabe bekannten Daten sollen für das Forschungsberichtswesen der Hochschule weiter aufbereitet werden. Berechnen Sie für die Angaben zur Forschungserfahrung die folgenden Lage- und Streuungsmaße:

a) Arithmetisches Mittel
b) Um 5% getrimmtes arithmetisches Mittel
c) Modus
d) Spannweite
e) Varianz
f) Standardabweichung

Anonymisierte Teilnehmerkennung	Bisherige Forschungs-erfahrung (in Jahren)
OV37KS	17
OV28ER	8
FG13WS	3
DV66AZ	21
TF45WR	12
GH73TG	20
UF84FB	17
OV33ZU	16
ED85FT	7
FV75DD	5

Aufgabe III: Verteilungsparameter

Arithmetisches Mittel

17 + 8 + 3 + 21 + 12 + 20 + 17 + 16 + 7 + 5 = 126

126 / 10 = **12,6**

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Um 5% getrimmtes arithmetisches Mittel

immer aufrunden!

0,05 * 10 = 0,5 → Kürzung um jeweils den größten und den kleinsten Wert im Datensatz (3 und 21)

17 + 8 + 12 + 20 + 17 + 16 + 7 + 5 = 102

102 / 8 = **12,75**

Modus

Die Verteilung ist nicht unimodal, weshalb sie auch **keinen Modus** aufweist.

Spannweite

21 – 3 = **18**

Aufgabe III: Verteilungsparameter

Varianz

Arithmetisches Mittel: 12,6

$(17-12{,}6)^2 = 19{,}36$ $(20-12{,}6)^2 = 54{,}76$
$(8-12{,}6)^2 = 21{,}16$ $(17-12{,}6)^2 = 19{,}36$
$(3-12{,}6)^2 = 92{,}16$ $(16-12{,}6)^2 = 11{,}56$
$(21-12{,}6)^2 = 70{,}56$ $(7-12{,}6)^2 = 31{,}36$
$(12-12{,}6)^2 = 0{,}36$ $(5-12{,}6)^2 = 57{,}76$

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Summe der quadrierten Differenzen vom arithmetischen Mittel: 378,4

378,4 / 10 = **37,84**

Aufgabe III: Verteilungsparameter

Standardabweichung

Positive Wurzel aus der Varianz → **6,15**

$$s = +\sqrt{s^2}$$

Was, wenn alternativ mit der Stichprobenvarianz gerechnet wurde?

Die Stichprobenvarianz berechnet sich mit (n-1) und beläuft sich auf 378,4 / 9 = **42,04**

Die Standardabweichung liegt in diesem Fall bei **6,48**

$$s^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Aufgabe III: Kontrolle mit PAST

Univariate statistics

	All
N	10
Min	3
Max	21
Sum	126
Mean	12,6
Std. error	2,050474
Variance	42,04444
Stand. dev	6,484169
Median	14
25 prcntil	6,5
75 prcntil	17,75
Skewness	-0,1983201
Kurtosis	-1,608032
Geom. mean	10,6948
Coeff. var	51,46166

Bootstrap

Bootstrap type: Simple

Bootstrap N: 9999

Recompute

Close Copy Print

https://www.nhm.uio.no/english/research/resources/past/

Aufgabe IV: Zusammenhangsmaße

Neben der bisherigen Forschungserfahrung wird unter anderem auch erhoben, an wie vielen begutachteten (sog. peer-reviewten) Publikationen die Professor*innen im Vorjahr beteiligt waren.

Berechnen und interpretieren Sie für die vorliegenden Daten a) den Rangkorrelationskoeffizienten nach Spearman sowie b) den Bravais-Pearson-Korrelationskoeffizienten.

Anonymisierte Teilnehmer-kennung	Bisherige Forschungs-erfahrung (in Jahren)	Peer-reviewte Publikationen im Vorjahr
OV37KS	17	11
OV28ER	8	5
FG13WS	3	2
DV66AZ	21	15
TF45WR	12	7
GH73TG	20	17
UF84FB	17	8
OV33ZU	16	9
ED85FT	7	3
FV75DD	5	3

Aufgabe IV: Berechnung des Spearman-Koeffizienten

x	rg(x)	y	rg(y)	d	d²
17	7,5	11	8	-0,5	0,25
8	4	5	4	0	0
3	1	2	1	0	0
21	10	15	9	1	1
12	5	7	5	0	0
20	9	17	10	-1	1
17	7,5	8	6	1,5	2,25
16	6	9	7	-1	1
7	3	3	2,5	0,5	0,25
5	2	3	2,5	-0,5	0,25

Die Summe aller quadrierten Rangplatzdifferenzen liegt bei 6 → Einsetzen in die Formel

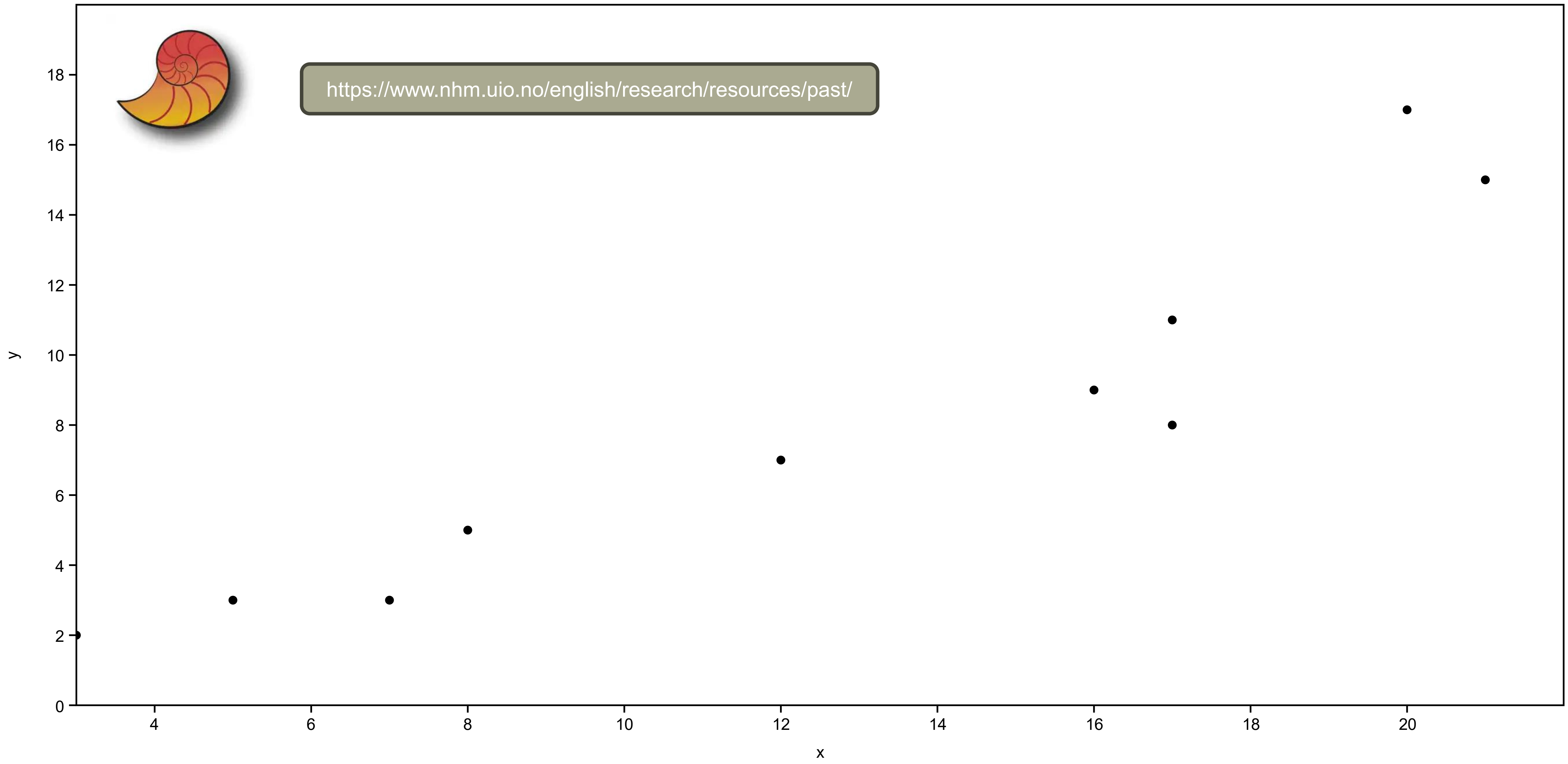
Aufgabe IV: Berechnung des Spearman-Koeffizienten

$\rho = 1 - (6 * 6) / [(10^2 - 1) * 10] = 1 - 36 / 990 =$ **<u>0,9636</u>**

$$\rho = 1 - \frac{6 \sum d_i^2}{(n^2 - 1)n}$$

Es liegt eine **starke, positive, monotone** Korrelation vor.
Zwei **verbundene Ränge** verringern die Aussagekraft des Spearman-Rangkorrelationskoeffizienten.

https://www.nhm.uio.no/english/research/resources/past/



Aufgabe IV: Berechnung des Bravais-Pearson-Koeffizienten

n	x	y	x^2	y^2	(x*y)
1	17	11	289	121	187
2	8	5	64	25	40
3	3	2	9	4	6
4	21	15	441	225	315
5	12	7	144	49	84
6	20	17	400	289	340
7	17	8	289	64	136
8	16	9	256	81	144
9	7	3	49	9	21
10	5	3	25	9	15
Summe	126	80	**1966**	**876**	**1288**
Mittel	**12,6**	**8**	//	//	//

Aufgabe IV: Berechnung des Bravais-Pearson-Koeffizienten

1288 – 10 * 12,6 * 8 = 280

$$\sqrt{(1966 - 10 * 12{,}6^2)} = 19{,}45$$

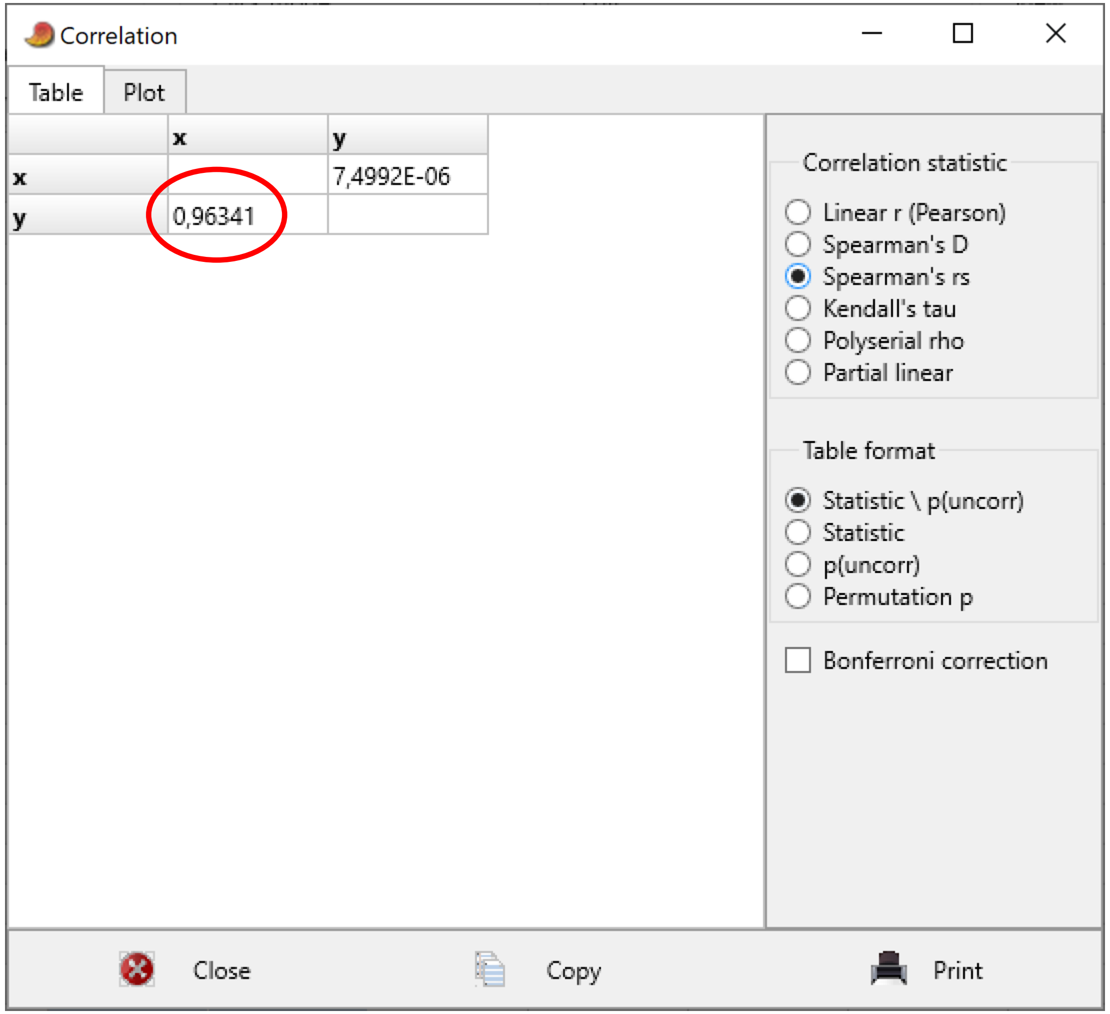
$$\sqrt{(876 - 10 * 8^2)} = 15{,}36$$

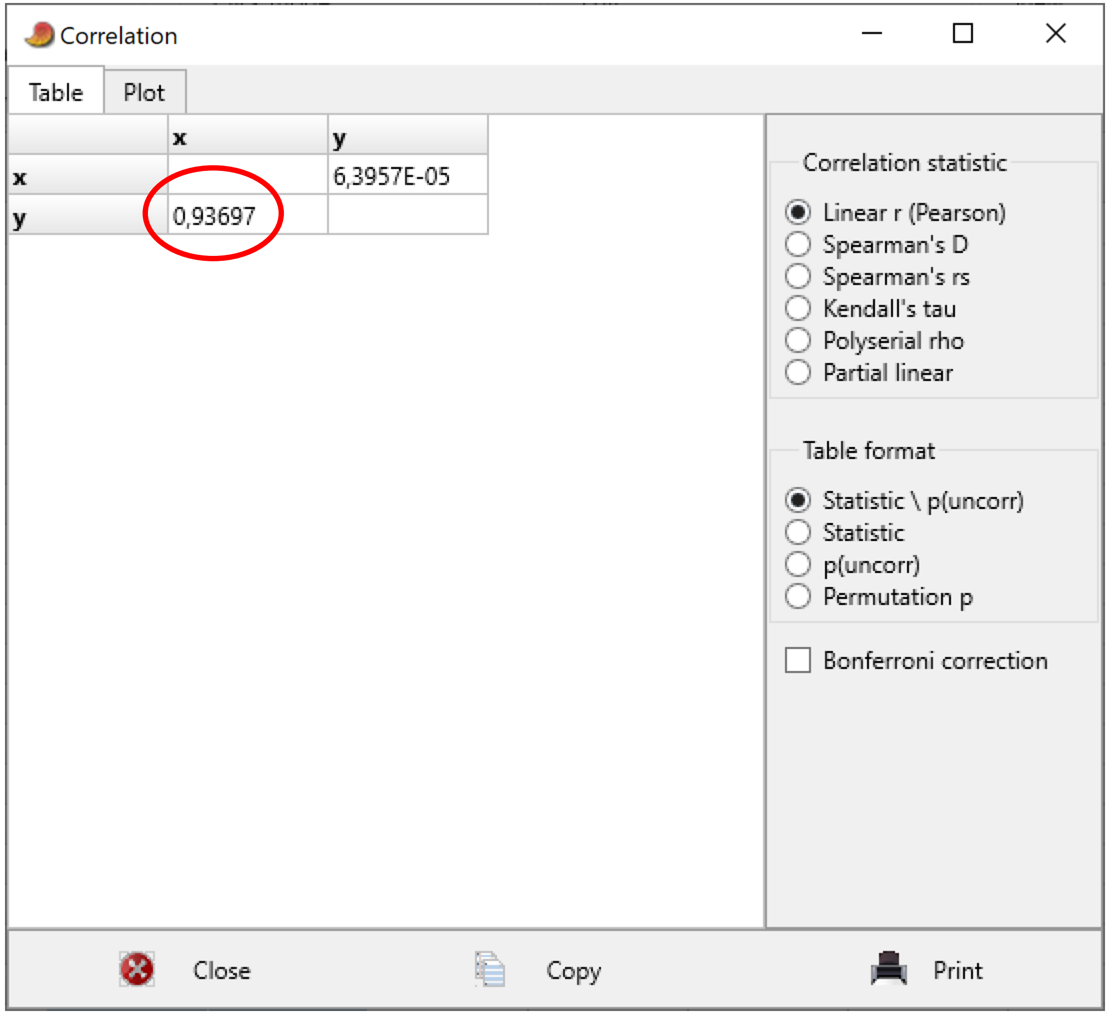
1288 / (19,45 * 15,36) = **0,9369**

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

Es liegt eine **starke, positive, lineare** Korrelation vor.

Aufgabe IV: Kontrolle mit PAST





Fragen im Nachgang gerne auch per E-Mail!

Viel Erfolg bei der Klausur!

Hochschule Harz
Hochschule für angewandte Wissenschaften

Christian Reinboth
Telefon +49 3943 – 896
E-Mail creinboth@hs-harz.de
Friedrichstraße 57 – 59
38855 Wernigerode