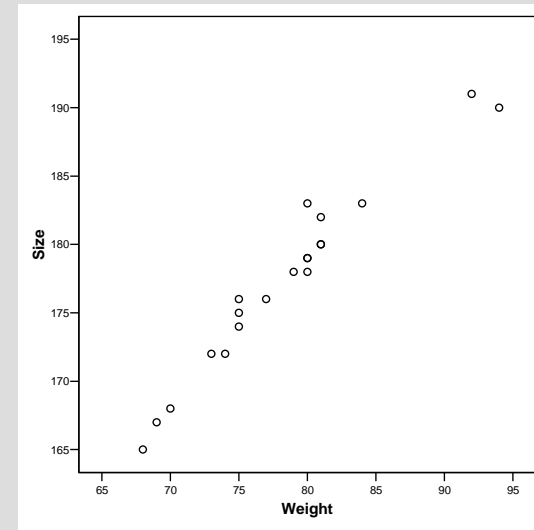
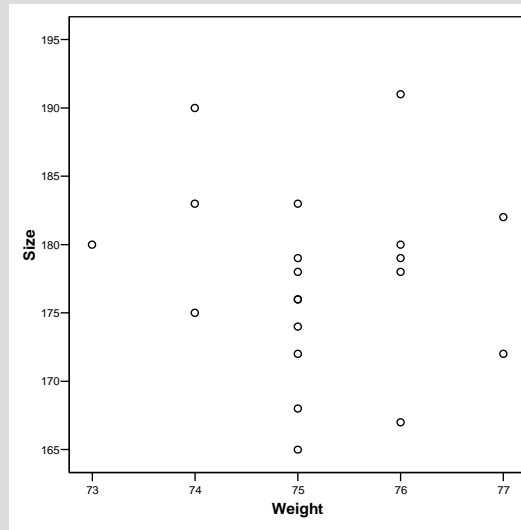
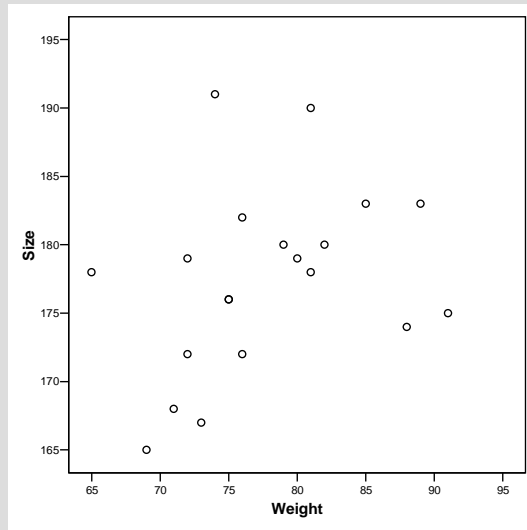


Correlational analysis

Students' performance on an exam are influenced by multiple factors. Two possible factors are (i) anxiety and (ii) study time. In order to test the effect of these two factors on students' performance the researcher conducts a correlational analysis with 15 students. He uses a standard test to determine their anxiety and measures the amount of time they spend studying for the exam.

Cases (= children)	Anxiety	Study time	Score on exam
1,00	33,00	45,00	61,00
2,00	64,00	68,00	72,00
3,00	33,00	100,00	84,00
4,00	22,00	44,00	39,00
5,00	70,00	62,00	50,00
6,00	66,00	61,00	55,00
7,00	59,00	52,00	71,00
8,00	84,00	66,00	71,00
9,00	56,00	79,00	66,00
10,00	44,00	44,00	51,00
11,00	22,00	16,00	29,00
12,00	44,00	61,00	45,00
13,00	80,00	60,00	70,00
14,00	66,00	61,00	58,00
15,00	79,00	60,00	65,00

Scatterplot



Correlational analyses

Interval	1. Pearson's r
----------	----------------

Correlational analyses

Interval	1. Pearson's r
Ordinal	1. Spearman's Rho 2. Kendall's Tau

Correlational analyses

Interval	1. Pearson's r
Ordinal	1. Spearman's Rho 2. Kendall's Tau
Nominal	1. Phi coefficient 2. Cramer's V 3. Pearson χ^2 test 4. Loglinear analysis

Correlational coefficient and effect size

Correlation coefficient	Shared variance (effect size)	
$r = 0.0$	0.00	Kein Zusammenhang
$r = 0.1$	0.01 (1%)	Geringe Korrelation
$r = 0.2$	0.04 (4%)	
$r = 0.3$	0.09 (9%)	Mittlere Korrelation
$r = 0.4$	0.16 (16%)	
$r = 0.5$	0.25 (25%)	
$r = 0.6$	0.36 (36%)	Hohe Korrelation
$r = 0.7$	0.49 (49%)	
$r = 0.8$	0.64 (64%)	Sehr hohe Korrelation
$r = 0.9$	0.81 (81%)	
$r = 1.0$	1.00 (100%)	

Example

Example: Age and MLU

Ein wichtiges Maß für den Entwicklungsstand eines Kindes ist die Mittlere Länge von Äußerungen (MLU = Mean length of Utterances). Danach nimmt die Anzahl der Wörter und/oder Silben in einer Äußerung mit dem Alter zu. Überprüfen Sie diese Hypothese an den folgenden Daten.

Cases	Children	Age in months	MLU
1	John	24	2.10
2	Bill	23	2.16
3	Sue	32	2.25
4	Jane	20	1.93
5	Ann	43	2.64
6	Susan	58	5.63
7	Jack	28	1.96
8	William	34	2.23
9	Mary	53	5.19
10	Peter	46	3.45
11	Pete	49	3.21
12	Allan	36	2.84

Example: Age and MLU

Long Report: There is an association between age and MLU. The correlation coefficient ($r = .887$) shows that 78.6% (r^2) of the variation in MLU is accounted for by the variation in age. The associated probability level of 0.001 shows that such a result is unlikely to have arisen from sampling error.

Short Report: As can be seen in the table above, there is a strong correlation between age and MLU ($r = .887$, $p = .001$).

Kendell's Tau

Example: Typicality and frequency

In einer Studie zur lexikalischen Semantik mussten 15 Probanden eine List von 10 Wörtern danach beurteilen, wie typisch das jeweilige Wort für die Kategorie ‚Fahrzeug‘ ist. Dazu mussten Sie die 10 Wörter in eine geordnete Reihenfolge von typisch bis untypisch bringen. Parallel dazu wurde die Häufigkeit der 10 Wörter in einem großen Korpus ermittelt und es wurde eine zweite Ranglist erstellt, in der die Wörter aufgrund ihrer Häufigkeit einem Rang auf einer Liste von 1 bis 10 zugewiesen wurden. Frage: Gibt es einen Zusammenhang zwischen der Häufigkeit der Wörter und dem Urteil der Probanden?

Example: Typicality and frequency

Words	Typicality rank	Frequency rank
car	1	1
truck	2	2
sports car	3	6
motor bike	4	5
train	5	3
bicycle	6	4
ship	7	8
boat	8	7
scat board	9	9
space shuttle	10	10

Kendall's tau ($\tau = .733$, $p = .003$)

Spearman's rho ($r_s = .879$, $p = .001$)

Partial correlation

Example: Typicality and frequency

Zipfs Gesetz besagt, dass die Länge eines Wortes mit seiner Häufigkeit korreliert: Häufig gebrauchte Wörter sind in der Regel kürzer als selten gebrauchte Wörter. Ein Forscher möchte die Gültigkeit des Zipfschen Gesetzes empirisch überprüfen. Zu diesem Zweck untersucht er die Länge und die Verwendungshäufigkeit von 10 willkürlich ausgewählten Wörtern. Als Indikator für die Verwendungshäufigkeit der Wörter ermittelt der Forscher deren Anzahl in einem großen Korpus. Für die Länge der Wörter verwendet er zwei Maße: (i) Anzahl der Phoneme, (ii) Anzahl der Silben.

Example: Frequency and length

Wort	Anzahl der Phoneme	Anzahl der Silben	Häufigkeit des Wortes
1	8	3	10,989
2	5	1	8,549
3	4	1	4,209
4	9	3	9,897
5	10	3	8,897
6	4	1	5,345
7	5	2	5,289
8	6	2	6,001
9	8	3	7,987
10	8	2	8,988

Partial correlation

Phoneme & Silben	$r = .898, p = .001$
Phoneme & Häufigkeit	$r = .795, p = .006$
Silben & Häufigkeit	$r = .677, p = .031$

Partial correlation:	Phoneme & Häufigkeit (Silben Konstant)	$r = .578, p = .103$
----------------------	---	----------------------

Nominal data

Nominanal data

50 boys and 50 girls are asked to select toys from a cupboard. The available toys were previously categorized as mechanical and non-mechanical. The hypothesis is that boys prefer mechanical toys and girls prefer non-mechanical toys.

	Mechanical	Non-mechanical	Total
Boys	30	20	50
Girls	15	35	50
Total	45	55	100

Nominal data

Es gibt eine signifikante Korrelation zwischen Geschlecht (Boys vs. Girls) und der Präferenz für ein bestimmtes Spielzeug (mechanisch vs. nicht-mechanisch):

Symmetrische Maße		
	Wert	Näherungsweise Signifikanz
Nominal- bzgl. Nominalmaß	Phi Cramer-V	,302 ,302
Anzahl der gültigen Fälle		100