

Wintersemester 2019/2020

How to „Hypothesen-Test“ mit SPSS

Ein Supplement zu den Teilschritten 4 und 5 in Leisens Leitfaden für statistische Auswertung



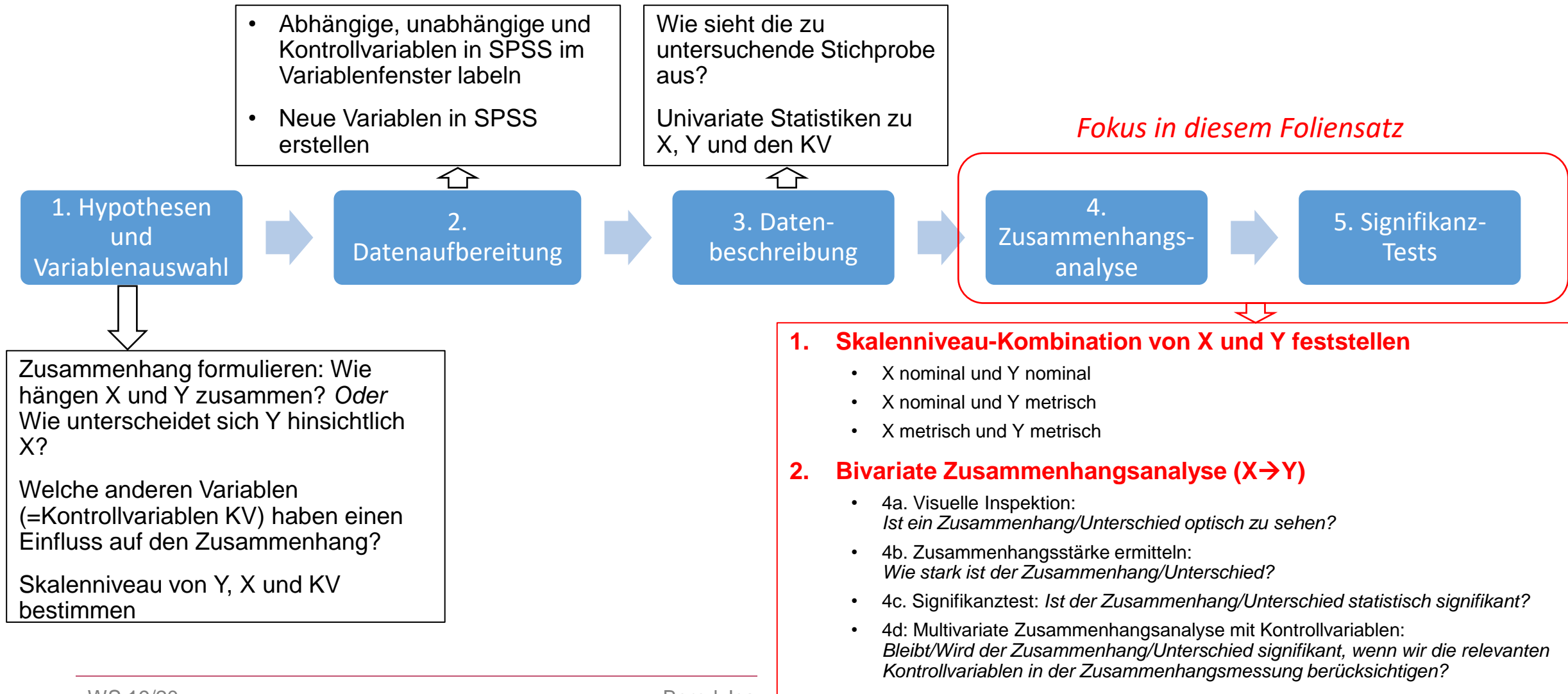
Universität Vechta
University of Vechta

Bernd Josef Leisen

Wissenschaftlicher Mitarbeiter

Management Sozialer Dienstleistungen/Dienstleistungsmanagement

Übersicht Vorgehen bei statistischen Hypothesentests



Hypothesen-Tests in SPSS durchführen und interpretieren

Vor dem Hypothesentest:

Skalenniveau-Kombination von X und Y feststellen

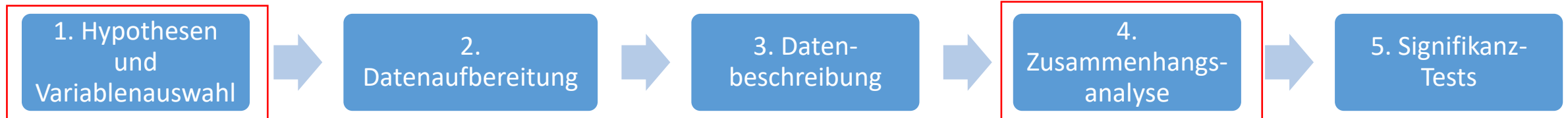
- X nominal und Y nominal
- X nominal und Y metrisch
- X metrisch und Y metrisch

Ggfls. Fälle selektieren

- Wenn nicht alle Personen relevant sind
 - z. B. beim Vergleich ausgewählter Personen, die Geld gewinnen können, gegen ihre Gruppenmitglieder
 - Oder: nur 2 Treatmentgruppen sollen miteinander verglichen werden
 - Oder: Nur Frauen sollen ausgewählt werden

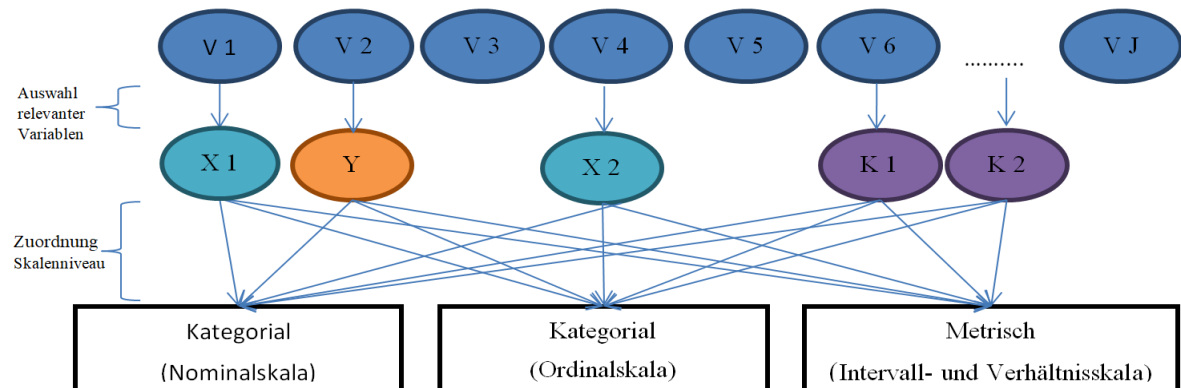
Skalenniveau-Kombination von X und Y feststellen

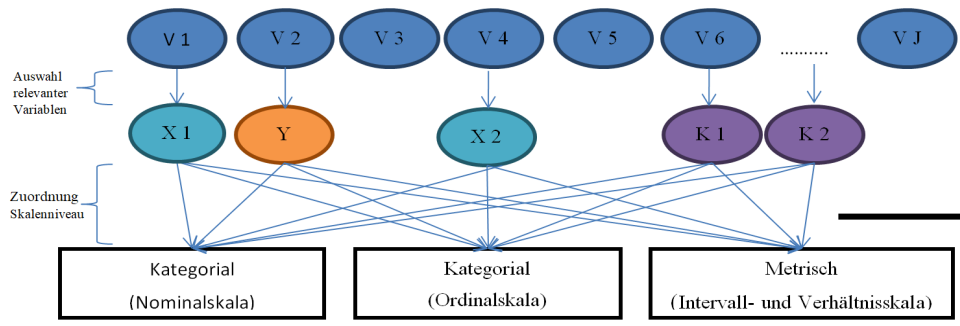
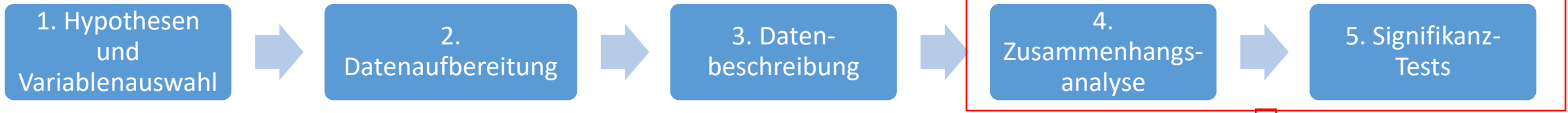
Hypothesen-Tests in SPSS durchführen und interpretieren



1. Hypothesen bilden und Variablenauswahl

- Auf Basis von **Vorwissen** Vermutung über den Zusammenhang zwischen zwei oder mehreren Variablen aufstellen (=Hypothese)
- Die hierfür **notwendigen Variablen** aus Ihrem Datensatz **auswählen** oder aus vorhandenen Variablen neu **erstellen**.
- Festlegung **abhängiger-** und **unabhängiger Variablen**: Welche Variable Y wird durch welche Variable(n) X_j beeinflusst? (für gerichtete Zusammenhangshypothesen)
- Festlegung von **Kontrollvariablen**: Welche Variablen könnten noch einen Einfluss auf Y oder X_j haben?
- Bestimmung des **Skalenniveaus**: nominal, ordinal oder metrisch (Intervall- oder Verhältnisskala)
 - Grund: Auswertungsverfahren unterscheiden sich nach Skalenniveaus der Variablen





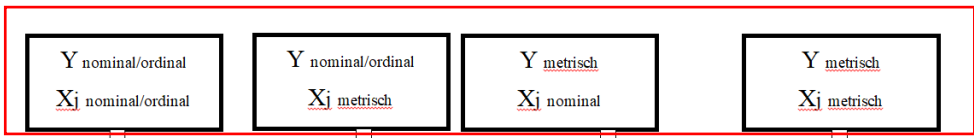
Skalenniveau-Kombination von X und Y feststellen

1. Welches Skalenniveau haben X und Y?

Nominal-skaliert: Daten können in keinerlei natürliche Reihenfolge gebracht werden
Bsp.: Geschlecht, Kontonummer, Treatments

Ordinal-skaliert: können in natürliche Reihenfolge gebracht werden, aber Abstände zwischen den einzelnen Werten sind nicht quantifizierbar
Bsp.: Schulnoten, Präferenzrangfolgen, vollverbalisierte Likertskalen

Metrisches Skalenniveau: natürliche Reihenfolge und quantifizierbare Abstände
Bsp.: Dauer in min, Geldbeträge, Alkoholwert,...



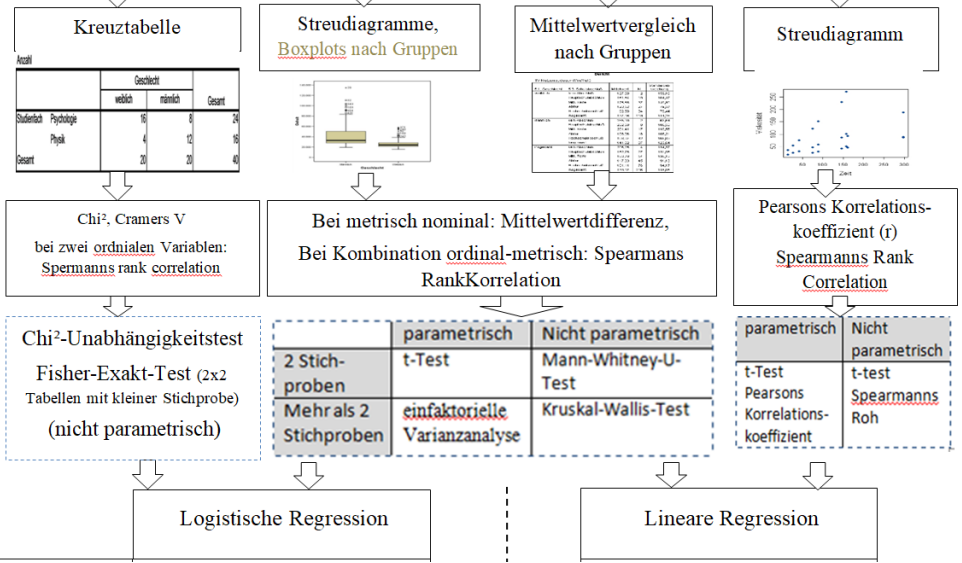
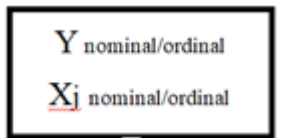
2. Welche Skalenniveau-Kombination haben X und Y?

Kritischer Arbeitsschritt: Y und X einem der 4 Kästen zuordnen

Nur wenn die Skalenniveaus müssen richtig bestimmt werden, können die richtigen Auswertungsverfahren aus der Liste abgelesen werden

Beispiel:

X = Treatment (0 = kein Feedback, 1 = Feedback) → nominal
 Y = Veränderung Alkoholwerte (-2,0 bis +2.5) → metrisch



Beispiel - Vorbereitung: nur bestimmte Fälle auswählen

Beispiel Datensatz Follow-Up-Survey

Wir haben 4 Treatmentgruppen → Ausprägungen der Variable „Treatment“ im Datensatz
0 = kein Feedback, 1 = Individualfeedback, 2 = Gruppenfeedback, 3 = Gesamtfeedback

Wir wollen aber nur 2 Treatmentgruppen miteinander vergleichen!

Personen, die **Feedback zum Alkoholkonsum in allen Gruppen auf der Party (T2)** erhalten

→ Variable im Datensatz „Treatment“ = 2

Personen, die kein Feedback erhalten haben

→ Variable im Datensatz „Treatment“ = 0

*Bedingung festlegen:
Ausgewählt werden nur Personen, bei denen
die Variable Treatment =0 oder =2 ist*

The screenshot shows the SPSS 'Fälle auswählen' dialog box. The 'Auswählen' section has 'Falls Bedingung zutrifft' selected, and the 'Falls...' button is circled in red. The condition 'Treatment=0 | Treatment=2' is entered in the text box. The 'Ausgabe' section has 'Nicht ausgewählte Fälle filtern' selected. The 'Funktionsgruppe' list is open, showing 'Alle' selected. The 'Funktionsgruppe' list is also circled in red. The 'Funktionsgruppe' list is open, showing 'Alle' selected.

- *In den Synthax einfügen und ausführen*
- *Alle Personen aus Treatment 1 und 0 werden dann gestrichen*



```
USE ALL.  
COMPUTE filter_$=(Treatment=0 | Treatment=2).  
VARIABLE LABELS filter_$ 'Treatment=0 | Treatment=2 (FILTER)'.  
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.  
FORMATS filter_$ (f1.0).  
FILTER BY filter_$.  
EXECUTE.
```

Hypothesen-Tests in SPSS durchführen und interpretieren

Es folgen drei Beispiele für die Durchführung von Hypothesentests in SPSS

Y nominal/ordinal
 X_j nominal/ordinal

Personen, die **Feedback zum Alkoholkonsum in allen Gruppen auf der Party (T2)** erhalten, fordern häufiger **Informationen zur Alkoholprävention** im Follow-Up-Survey an.

Y metrisch
 X_j nominal

Individuen, die ein **Live-Gruppen-Feedback(T1 oder T2 oder T3)** erhalten, verzeichnen einen signifikant geringeren **Anstieg des Atemalkoholwerts** von Messung1 zu Messung2 als Individuen, die kein Feedback (T0, treat = 0) erhalten.

Y metrisch
 X_j metrisch

Je mehr **Zeit**, Personen **außerhalb ihrer Peergruppe** verbringen, desto geringer ist Ihr **Anstieg des Atemalkoholwerts** von Messung1 zu Messung2 als Individuen, die kein Feedback (T0, treat = 0) erhalten.

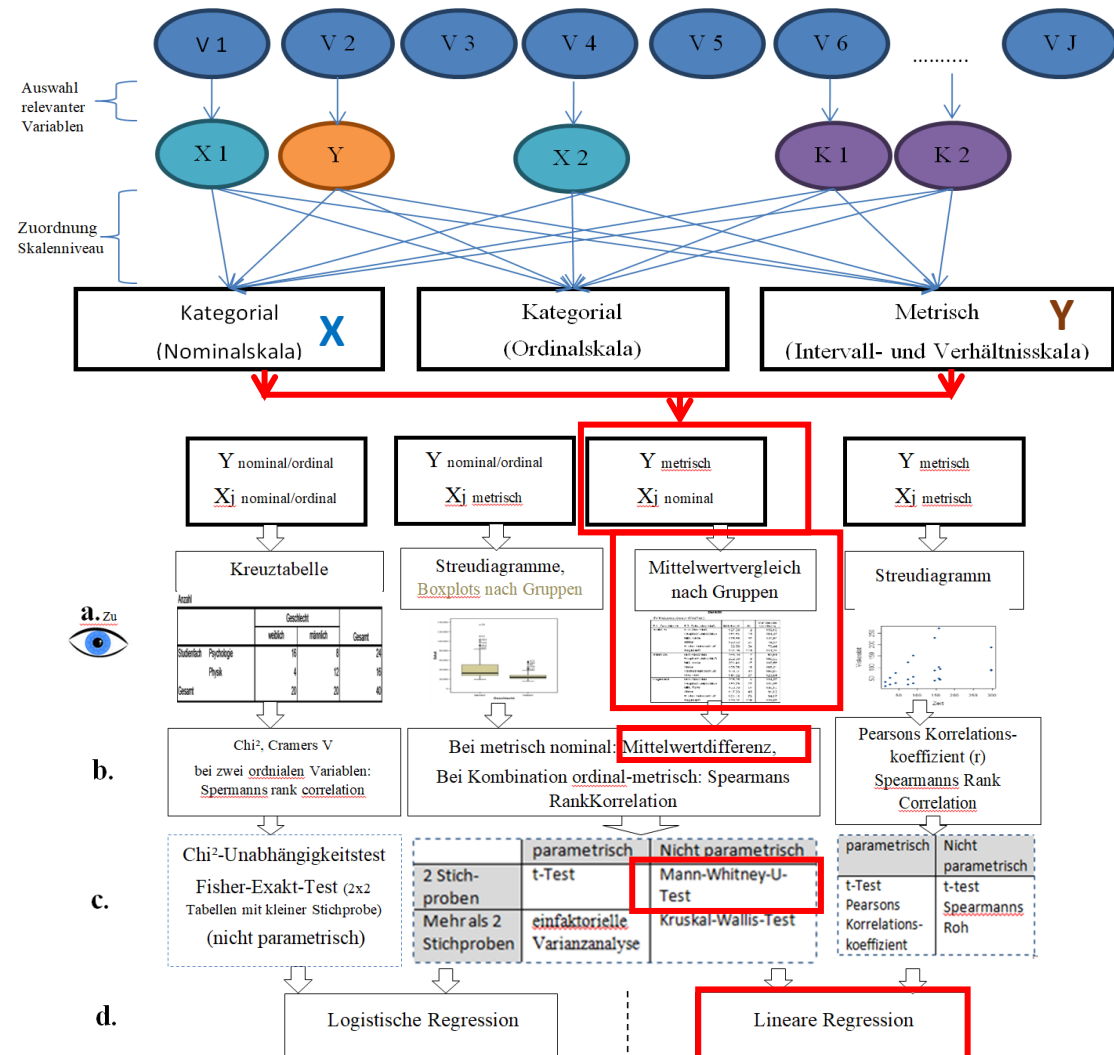
Beispiel – Hypothese: X nominal → Y metrisch

Hypothesen-Tests in SPSS durchführen und interpretieren

H1: Individuen, die ein **Live-Gruppen-Feedback (T1 oder T2 oder T3)** erhalten, verzeichnen einen **signifikant geringeren Anstieg des Atemalkoholwerts** von Messung 1 zu Messung 2 als Individuen, die kein Feedback (T0, treat = 0) erhalten.

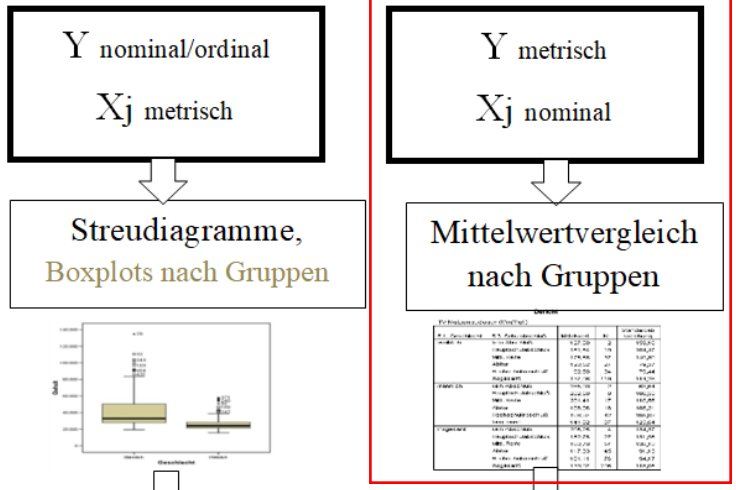
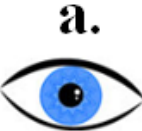
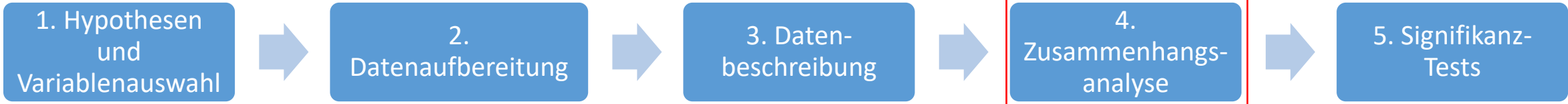
Abhängige Variable Y = Anstieg des Atemalkoholwerts (diffBAC) → metrisches Skalenniveau

Unabhängige X = Live-Gruppen-Feedback (/Ex-Post-Bonus/Ex-Ante-Bonus) → nominal



Beispiel Hypothesentest: X nominal \rightarrow Y metrisch

4a visuelle Inspektion und 4b Stärke des Zusammenhangs



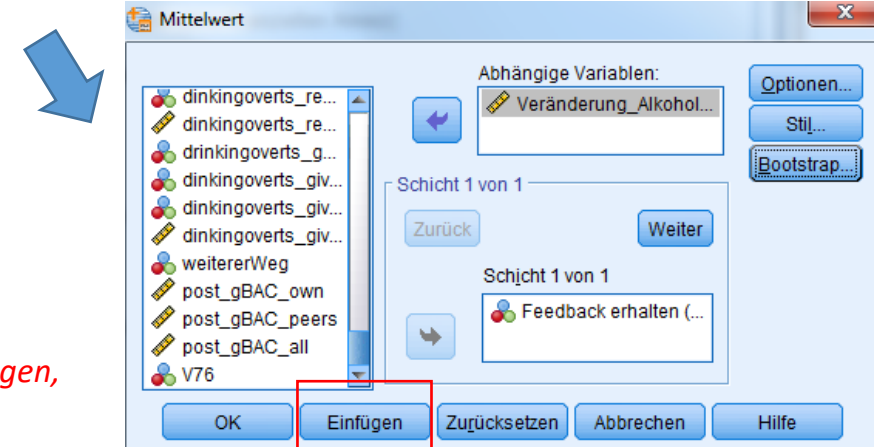
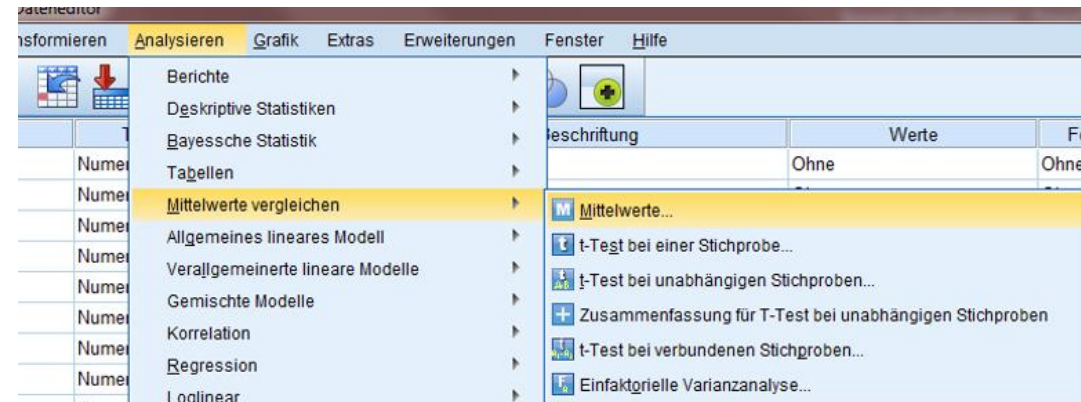
b.

Bei metrisch nominal: Mittelwertdifferenz,
Bei Kombination ordinal-metrisch: Spearmans
RankKorrelation

c.

	parametrisch	Nicht parametrisch
2 Stichproben	t-Test	Mann-Whitney-U-Test
Mehr als 2 Stichproben	einfaktorielle Varianzanalyse	Kruskal-Wallis-Test

Technische Umsetzung: Mittelwertvergleich in SPSS



In den Synthax einfügen,
Code und ausführen

Beispiel Hypothesentest: X nominal → Y metrisch

4a visuelle Inspektion und 4b Stärke des Zusammenhangs

4. bivariate Zusammenhangsanalyse

4a. Ist ein Zusammenhang/Unterschied optisch zu erkennen?

4b. Wie stark ist der Unterschied/Zusammenhang in der Stichprobe?

```
49  
50 ***4. bivariate Zusammenhangsanalyse  
51 4a. Ist ein Zusammenhang/Unterschied optisch zu erkennen?  
52 4b. Wie stark ist der Unterschied/Zusammenhang in der Stichprobe?  
53 4c. Ist der Unterschied statistisch signifikant?  
54  
55  
56 MEANS TABLES=Veränderung_Alkoholwert BY Feedback_mit_und_ohne_Bonus  
57 /CELLS=MEAN COUNT STDDEV.
```

In den Syntax einfügen,
Code und ausführen



Bericht			
Veränderung_Alkoholwert	Mittelwert	N	Std.- Abweichung
Feedback erhalten (mit oder ohne Bonus)			
kein Feedback erhalten	,2265	83	,43881
Feedback erhalten (mit oder ohne Bonus)	,2114	351	,42467
Insgesamt	,2143	434	,42694

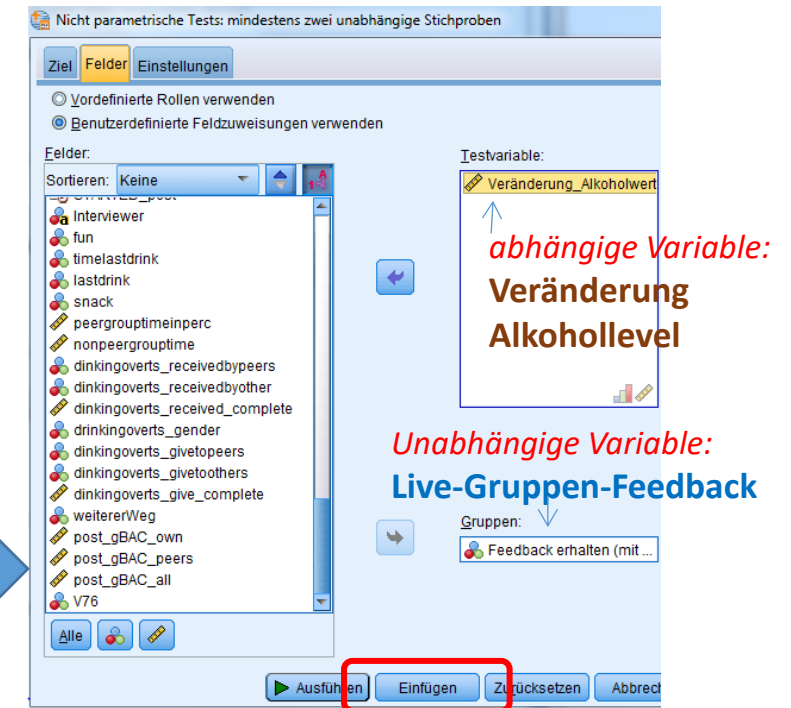
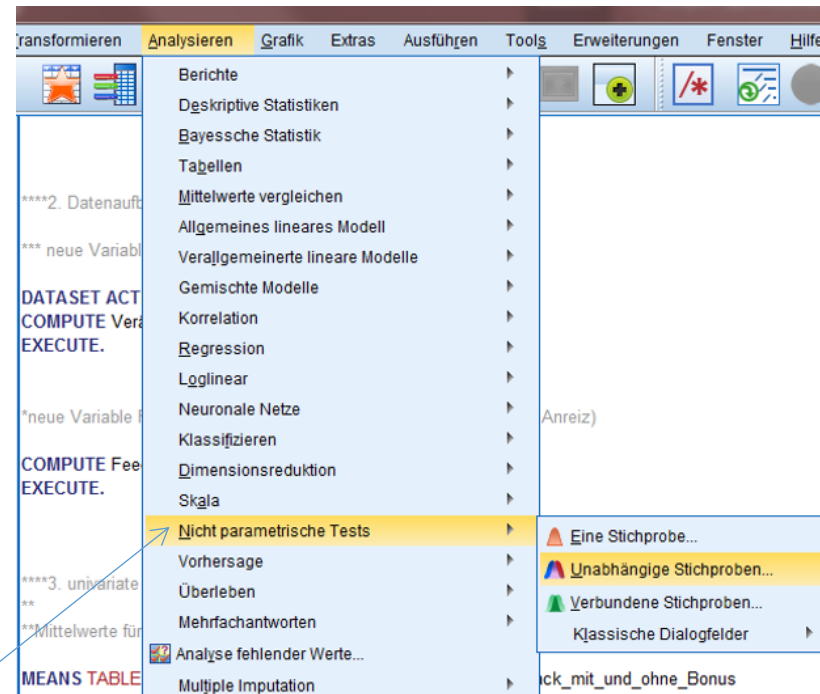
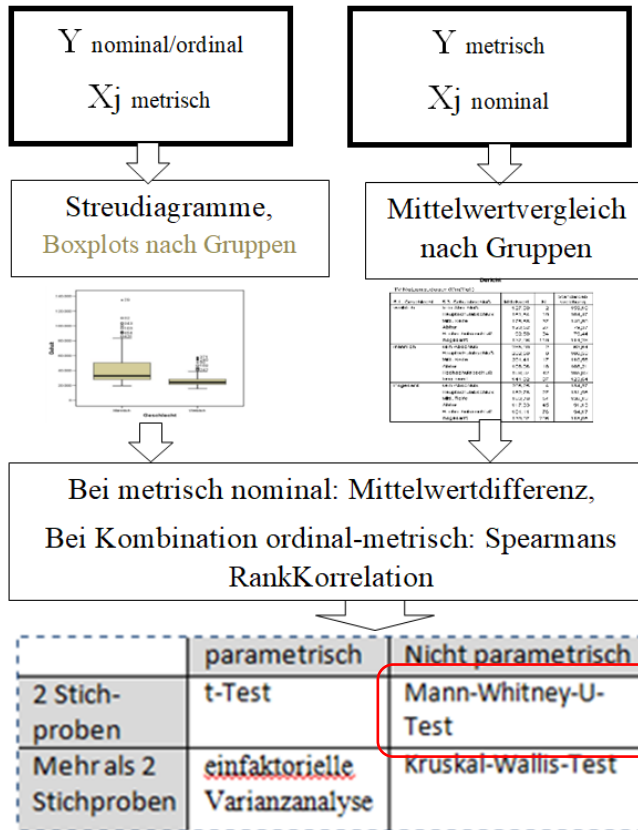
Durchschnittlicher Alkoholanstieg aller Partygäste 0.214 Promille

Mittelwertdifferenz: Ein kleiner und schwacher Unterschied zwischen Personen mit und ohne Feedback in Höhe von 0.0151 Promille ist zu sehen

Beispiel Hypothesentest: X nominal → Y metrisch

4c. Signifikanz des Zusammenhangs: Mann-Whitney-U-Test

- 4c. Ist der Unterschied statistisch signifikant?



```

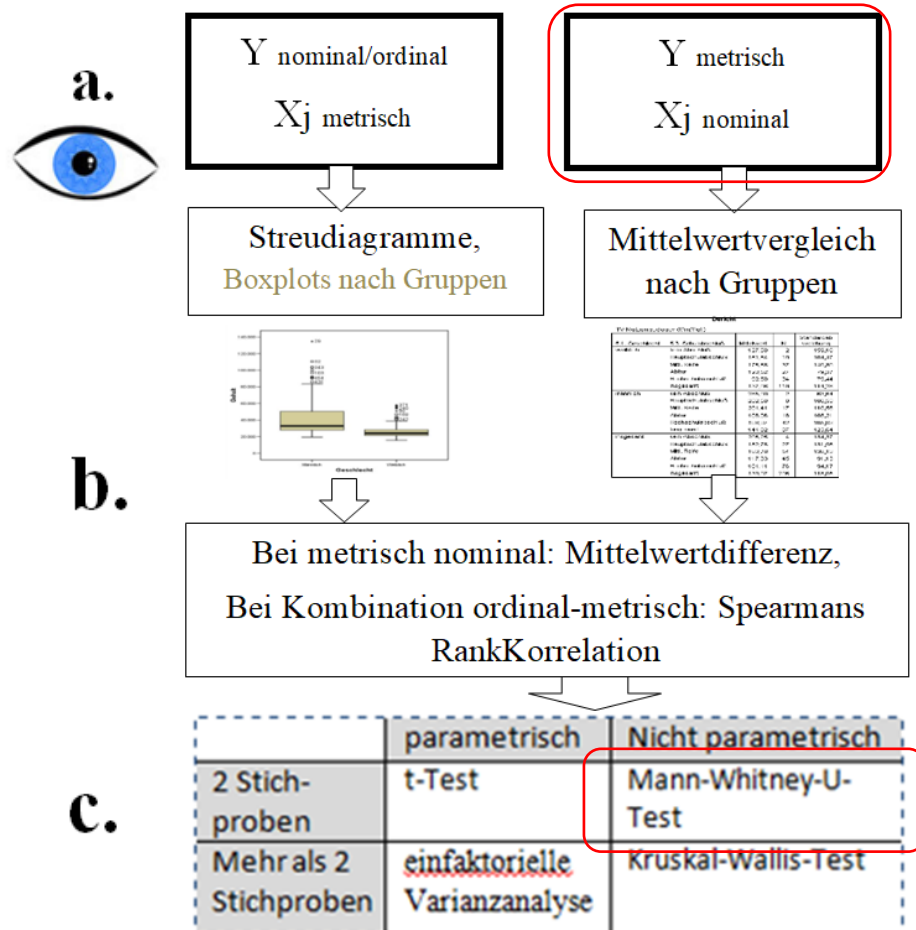
30 *Nonparametric Tests: Independent Samples.
31 NPTTESTS
32 /INDEPENDENT TEST (Veränderung_Alkoholwert) GROUP (Feedback_mit_und_ohne_Bonus)
33 /MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
34 /CRITERIA ALPHA=0.05 CILEVEL=95.
    
```

In den Syntax einfügen, Code und ausführen

Beispiel Hypothesentest: X nominal \rightarrow Y metrisch

4c. Signifikanztest: Mann-Whitney-U-Test

4c. Ist der Unterschied statistisch signifikant?



Interpretation: p-Wert = 0.819

Unter der Annahme der Nullhypothese (=Es gibt keinen Unterschied im Alkohollevelanstieg zwischen Personen mit und ohne Feedback) ist die Wahrscheinlichkeit eine Mittelwertdifferenz von 0,0151 zu finden 81,9%. \rightarrow Das ist über 10% ($p > 0.1$), d. h. Nullhypothese muss beibehalten werden

Signifikanzniveaus:

$p \leq 0.1 \rightarrow$ statistisch schwach signifikant *

$p \leq 0.05 \rightarrow$ statistisch signifikant **

$p \leq 0.01 \rightarrow$ statistisch stark signifikant ***

Nicht parametrische Tests

Hypothesentestübersicht				
	Nullhypothese	Test	Sig.	Entscheidung
1	Die Verteilung von Veränderung_Alkoholwert ist über die Kategorien von Feedback erhalten (mit oder ohne Bonus) identisch.	Mann-Whitney-U-Test bei unabhängigen Stichproben	,819	Nullhypothese beibehalten

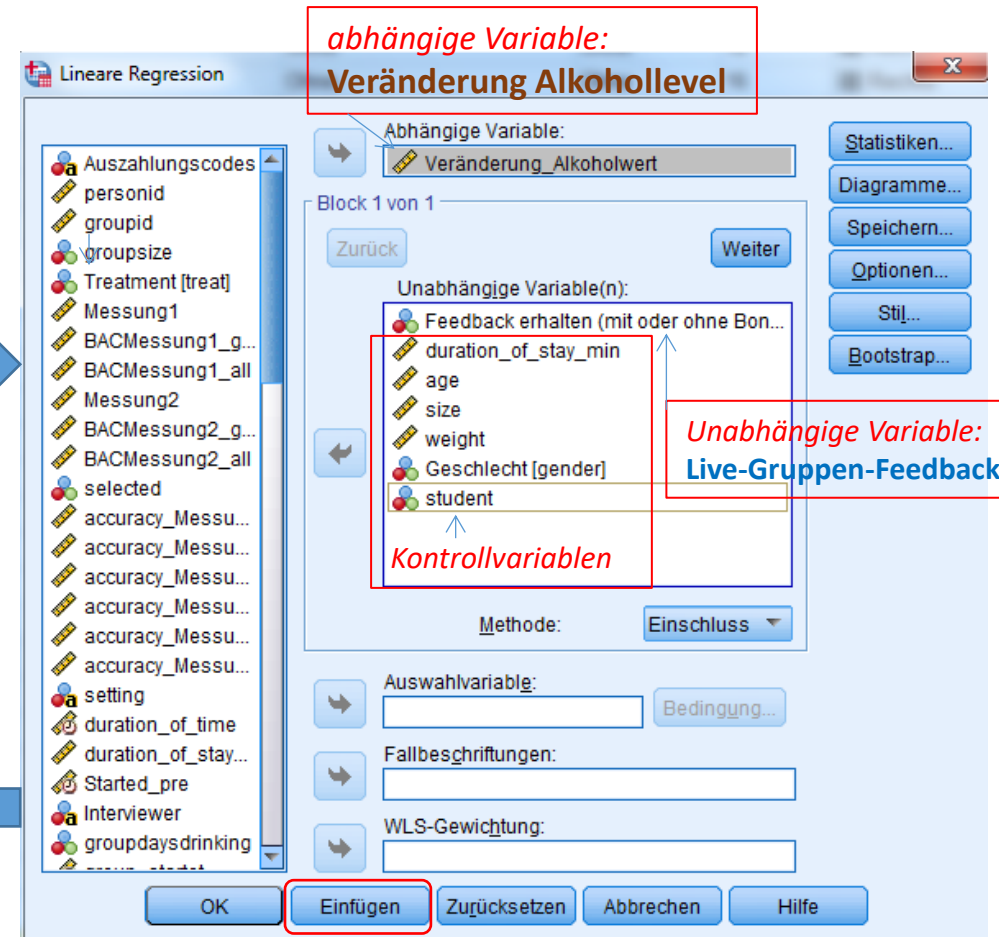
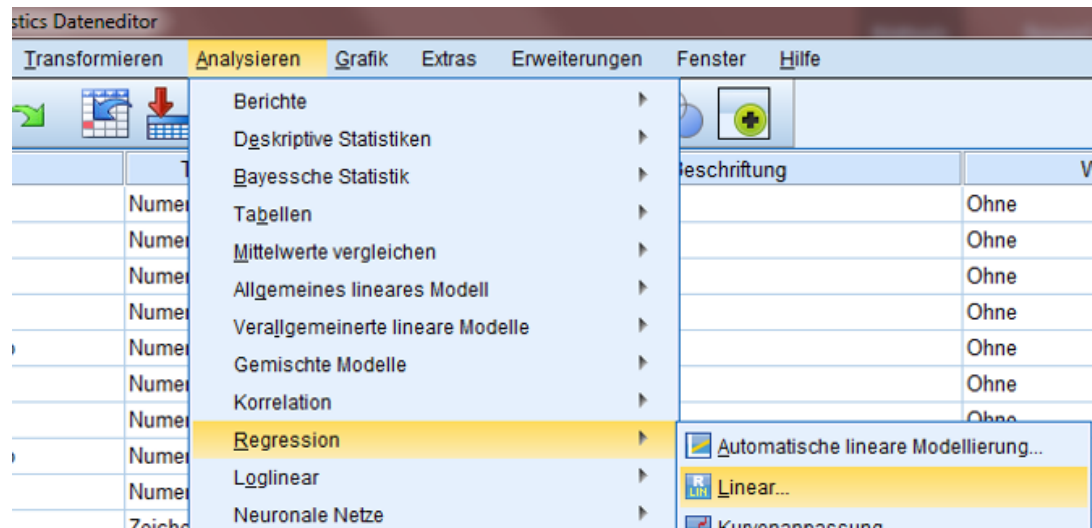
Asymptotische Signifikanzen werden angezeigt. Das Signifikanzniveau ist ,050.

Beispiel Hypothesentest: X nominal \rightarrow Y metrisch

4d. Multivariate Analyse in der linearen Regression

4d. Bleibt/Wird der Unterschied/Zusammenhang statistisch signifikant, wenn wir auf andere relevante Einflussfaktoren kontrollieren?

Technische Umsetzung in SPSS: Lineare Regression



abhängige Variable:
Veränderung Alkohollevel

Unabhängige Variable:
Live-Gruppen-Feedback

Kontrollvariablen

In den Synthax einfügen,
Code und ausführen

```
67 *4d. bleibt der Unterschied signifikant, wenn wir auf andere Einflussgrößen kontrollieren?
68
69 REGRESSION
70 /MISSING LISTWISE
71 /STATISTICS COEFF OUTS R ANOVA
72 /CRITERIA=PIN(.05) POUT(.10)
73 /NOORIGIN
74 /DEPENDENT Veränderung_Alkoholwert
75 /METHOD=ENTER Feedback_mit_und_ohne_Bonus duration_of_stay_min age size weight gender student.
```

Beispiel Hypothesentest: X nominal \rightarrow Y metrisch

4d. *Multivariate Analyse in der linearen Regression*

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,382 ^a	,146	,129	,40300

a. Einflußvariablen : (Konstante), student, Feedback erhalten (mit oder ohne Bonus), size, duration_of_stay_min, age, weight, Geschlecht

Modellgüte interpretieren:

Das Regressionsmodell mit allen unabhängigen Variablen erklärt 12,9% der Streuung der abhängigen Variable „Veränderung des Alkoholwertes“

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Sig.
		Regressionskoeffizient B	Std.-Fehler	Beta		
1	(Konstante)	-1,240	,552		-2,245	,025
	Feedback erhalten (mit oder ohne Bonus)	-,055	,054	-,050	-1,009	,313
	duration_of_stay_min	,002	,000	,316	6,140	,000
	age	-,001	,003	-,015	-,286	,775
	size	,007	,003	,162	2,165	,031
	weight	,001	,002	,025	,378	,706
	Geschlecht	,029	,061	,034	,478	,633
	student	-,070	,048	-,081	-1,462	,144

a. Abhängige Variable: Veränderung_Alkoholwert

Einflussstärke und Signifikanz des Einfluss von X interpretieren

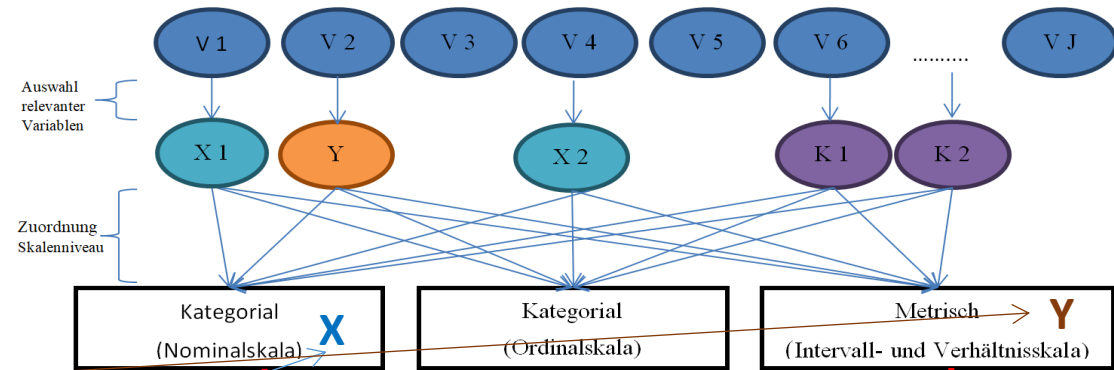
Personen mit Feedback haben im Durchschnitt einen um 0,055 geringeren Anstieg des Alkohollevels, wenn wir auf andere relevante Variablen kontrollieren.

Der Einfluss des Feedbacks ist mit einem p-Wert b von 0.313 nicht signifikant ($p \geq 0.1$)

Beispiel – Hypothese: X metrisch → Y metrisch

Skalenniveau-Kombination von X und Y feststellen

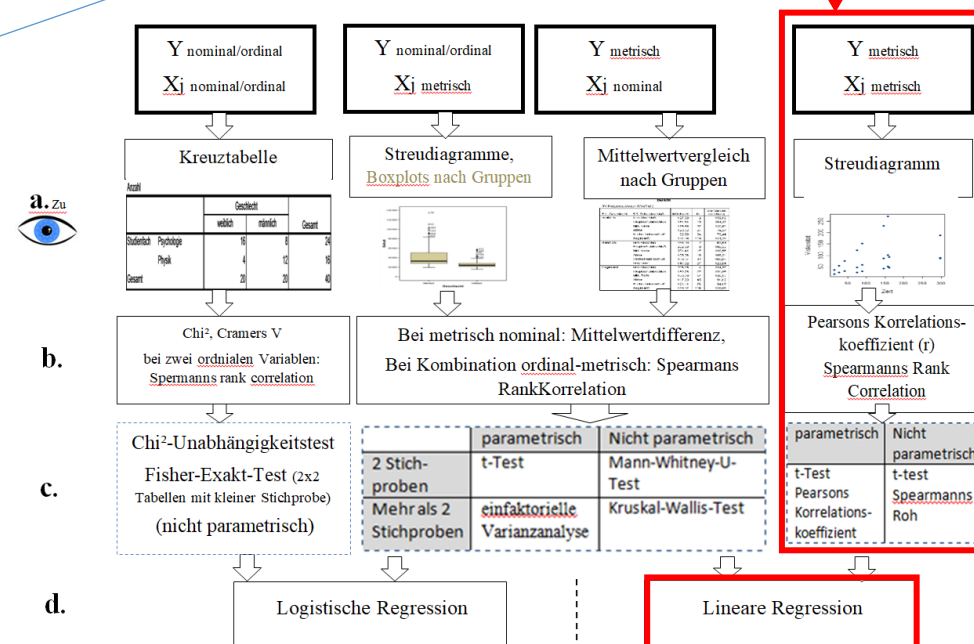
H1: Je mehr **Zeit**, Personen **außerhalb ihrer Peergruppe** verbringen, desto geringer ist Ihr **Anstieg des Atemalkoholwerts** von Messung1 zu Messung2 als Individuen, die kein Feedback (T0, treat = 0) erhalten.



Abhängige Variable Y = Anstieg des Atemalkoholwerts (diffBAC) → metrisches Skalenniveau

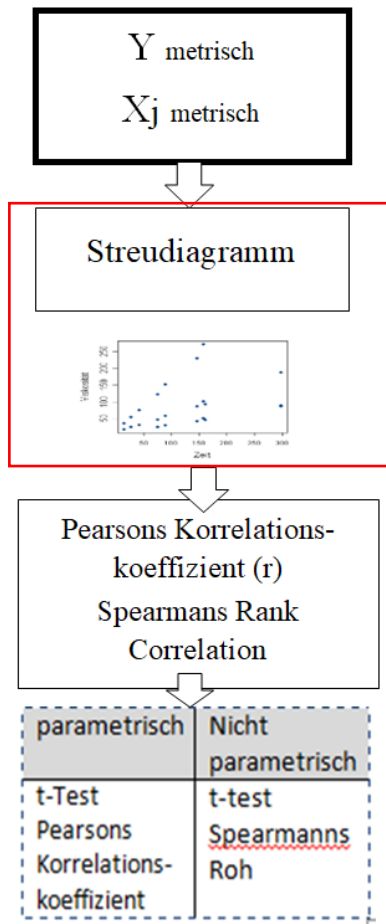
Unabhängige X = %Zeit außerhalb der Peergruppe – metrisches Skalenniveau

Kontrollvariablen K: Aufenthaltsdauer (duration_of_stay_min) → metrisch, Alter (age) → metrisch, Geschlecht (gender) → nominal, Gewicht (weight) → metrisch,



Beispiel – Hypothese: X metrisch → Y metrisch

4a. Optische Inspektion mittels Streudiagramm



Dateneditor

Formatieren Analysieren **Grafik** Extras Erweiterungen Fenster Hilfe

Diagrammerstellung...
Auswahl der Diagrammtafelvorgabe...

- + Weibull-Diagramm...
- + Untergruppen vergleichen
- + Regressionsvariablen-Diagramme

Klassische Dialogfelder

Typ	Breite	...
Zeichenfolge	8	
Numerisch	3	
Numerisch	3	
Numerisch	1	0 Treatment
Numerisch	32	2
Numerisch	21	15

Auswahl der Diagrammtafelvorgabe

Basis Detailliert Titel Optionen

Natürlich Name Typ

- fastturnik
- snack
- peergrouptimeperc
- nonpeergrouptime
- dinkingverts_receivedbypeers
- dinkingverts_receivedbyother
- dinkingverts_received_complete
- dinkingverts_gender
- dinkingverts_givetothers
- dinkingverts_givetothers
- dinkingverts_give_complete
- weitererWeg
- post_gBAC_own
- post_gBAC_peers
- post_gBAC_all
- V76
- Veränderung_Alkoholwert
- Feedback erhalten (mit oder ohne Bonus)

Visualisierung von:
Veränderung_Alkoholwert
nonpeergrouptime

3D-Dichte 3D-Histogramm Band

Klassiertes Streudiagramm Koordinaten auf einer Bezugskarte Linie

Streudiagramm Streudiagramm-Matrix (SPLOM)

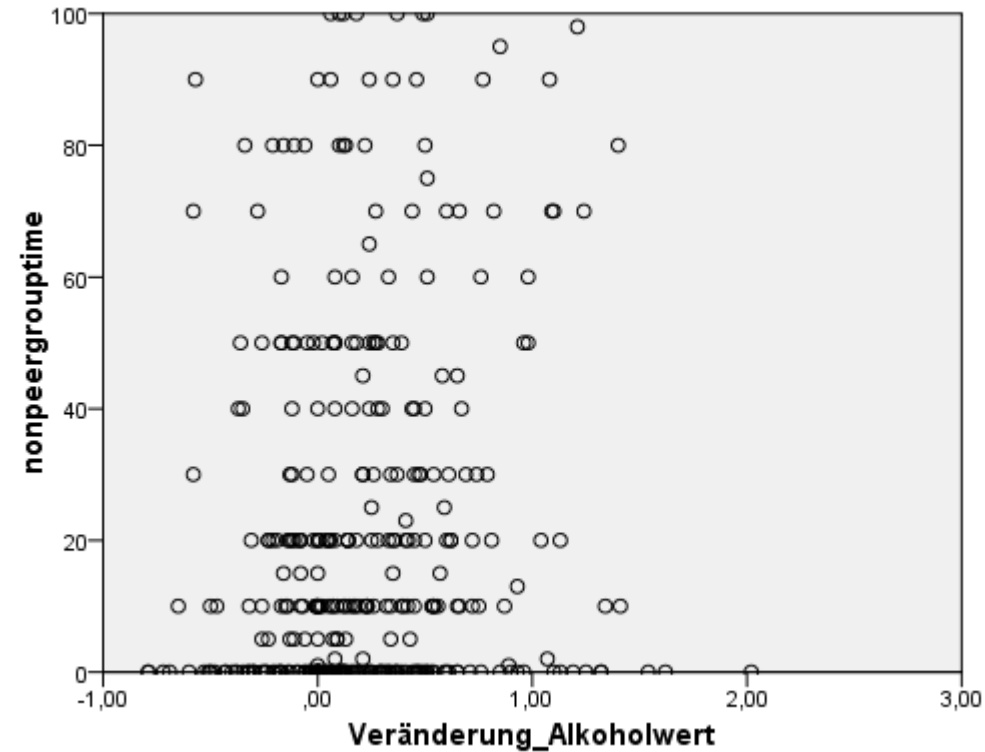
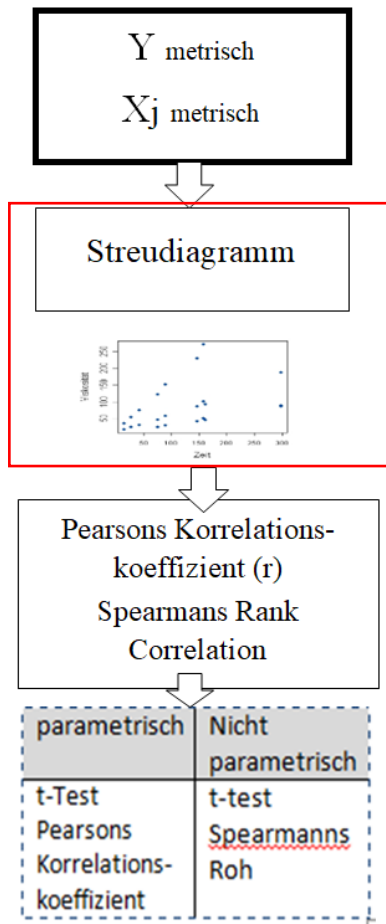
Auswertung:

Verwalten...

OK Einfügen Zurücksetzen Abbrechen Hilfe

Beispiel – Hypothese: X metrisch → Y metrisch

4a. Optische Inspektion mittels Streudiagramm



→ Ein schwacher Zusammenhang ist zu sehen.

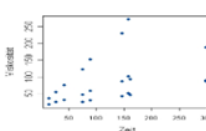
Beispiel – Hypothese: X metrisch → Y metrisch

4b Stärke des Zusammenhangs und 4c Signifikanztest

Y metrisch
X_j metrisch

↓

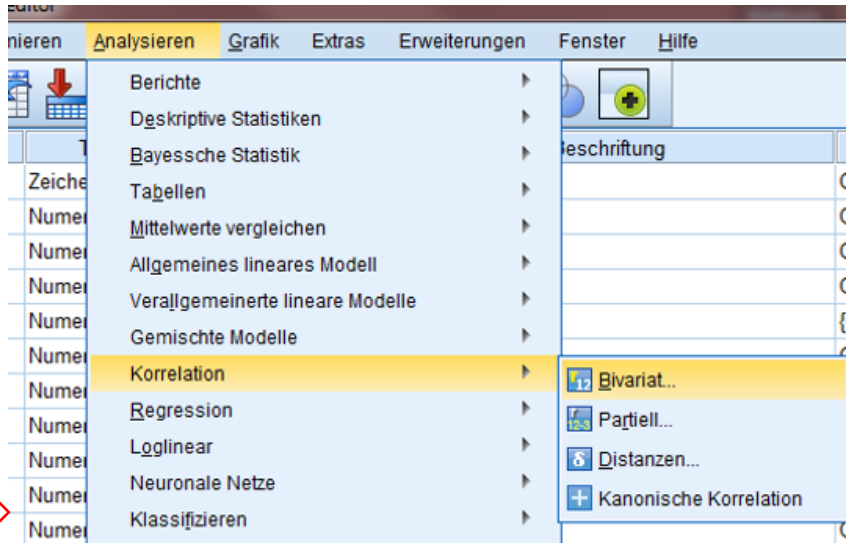
Streudiagramm

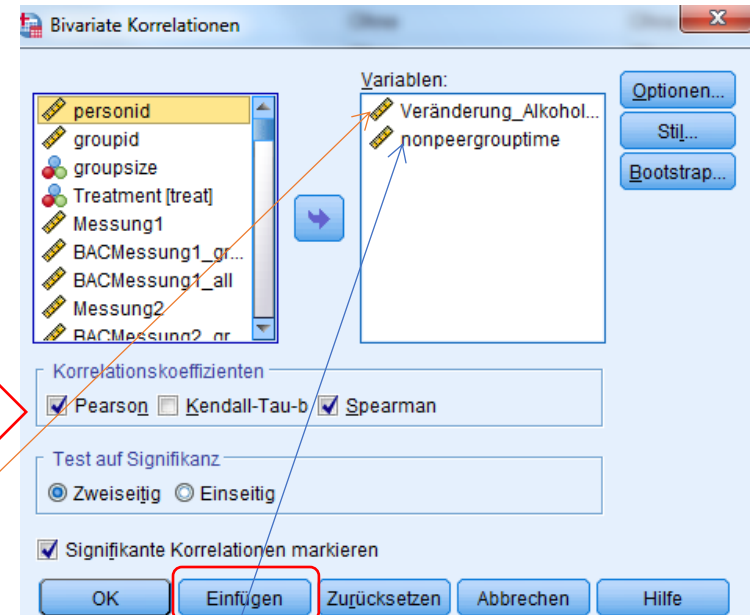


↓

Pearsons Korrelationskoeffizient (r)
Spearman's Rank Correlation

parametrisch	Nicht parametrisch
t-Test Pearsons Korrelationskoeffizient	t-test Spearman's Roh





In den Syntax einfügen,
Code und ausführen

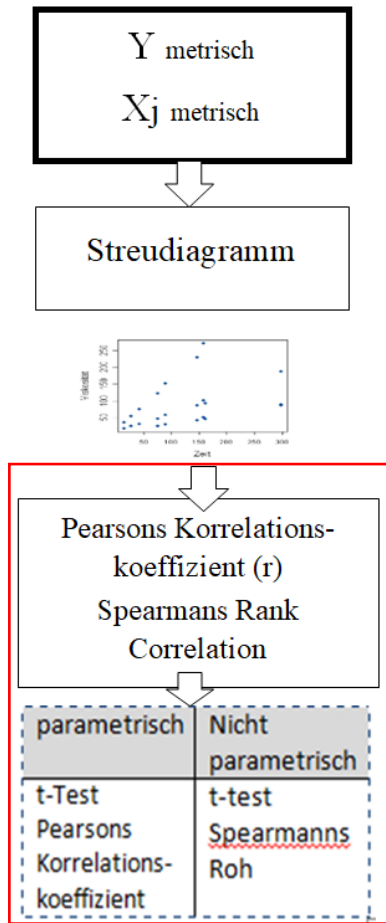
Abhängige Variable Y = Anstieg des Atemalkoholwerts (diffBAC) → metrisches Skalenniveau

Unabhängige X = %Zeit außerhalb der Peergruppe – metrisches Skalenniveau

Beispiel – Hypothese: X metrisch → Y metrisch

4b Zusammenhangsstärke und 4c Signifikanz

→ Nichtparametrische Korrelationen



	Veränderung_Alkoholwert	nonpeergrouptime
Spearman-Rho	Korrelationskoeffizient	.145**
	Sig. (2-seitig)	.005
	N	434
nonpeergrouptime	Korrelationskoeffizient	.145**
	Sig. (2-seitig)	.005
	N	371

** Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

4b Wie stark ist der Zusammenhang/Unterschied?
Korrelationskoeffizient $r=0.145$ zeigt einen schwach positiven Zusammenhang

4c Ist der Zusammenhang/Unterschied statistisch signifikant?

P-Wert = 0.005 → Der Zusammenhang ist hoch signifikant mit einem p-Wert von 0.005 ($p \leq 0.01$)

Interpretation von r nach Cohen (1988)

kleiner Effekt	$ r = .10$
mittlerer Effekt	$ r = .30$
großer Effekt	$ r = .50$

Signifikanzniveaus:

- $p \leq 0.1$ → statistisch schwach signifikant *
- $p \leq 0.05$ → statistisch signifikant **
- $p \leq 0.01$ → statistisch stark signifikant ***

Beispiel – Hypothese: X metrisch → Y metrisch

4d. Multivariate Analyse in der linearen Regression

Lineare Regression

Abhängige Variable: Veränderung_Alkoholwert

Block 1 von 1

Unabhängige Variable(n): nonpeergrouptime, duration_of_stay_min, age, size, weight, Geschlecht [gender], student

Methode: Einschluss

```

142
143 REGRESSION
144 /MISSING LISTWISE
145 /STATISTICS COEFF OUTS R ANOVA
146 /CRITERIA=PIN(.05) POUT(.10)
147 /NOORIGIN
148 /DEPENDENT Veränderung_Alkoholwert
149 /METHOD=ENTER nonpeergrouptime duration_of_stay_min age size weight gender student.
150

```

Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,377 ^a	,142	,124	,39884

a. Einflussvariablen : (Konstante), student, nonpeergrouptime, duration_of_stay_min, Geschlecht, age, weight, size

Das Regressionsmodell mit allen unabhängigen Variablen erklärt 12,4% der Streuung der abhängigen Variable „Veränderung des Alkoholwertes“

Beispiel – Hypothese: X metrisch → Y metrisch

4d) Ist der Einfluss der Non-Peer-Group-Time Y auf X die Alkoholveränderung signifikant, wenn wir gleichzeitig auf andere relevante Einflussgrößen kontrollieren?

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Sig.
		Regressionskoeffizient B	Std.-Fehler	Beta		
1	(Konstante)	-,981	,567		-1,731	,084
	nonpeergrouptime	,001	,001	,040	,753	,452
	duration_of_stay_min	,002	,000	,322	6,042	,000
	age	-,002	,003	-,034	-,609	,543
	size	,005	,003	,119	1,496	,136
	weight	,001	,002	,051	,740	,460
	Geschlecht	,009	,063	,011	,150	,881
	student	-,051	,049	-,060	-1,047	,296

a. Abhängige Variable: Veränderung_Alkoholwert

Interpretation

Stärke des Einflusses: $b = 0.001$

Mit jedem % der Zeit, das eine Person zusätzlich außerhalb seiner Peergruppe verbringt steigt der Alkoholzuwachs um 0.001 Promille, wenn wir auf andere relevante Variablen kontrollieren.

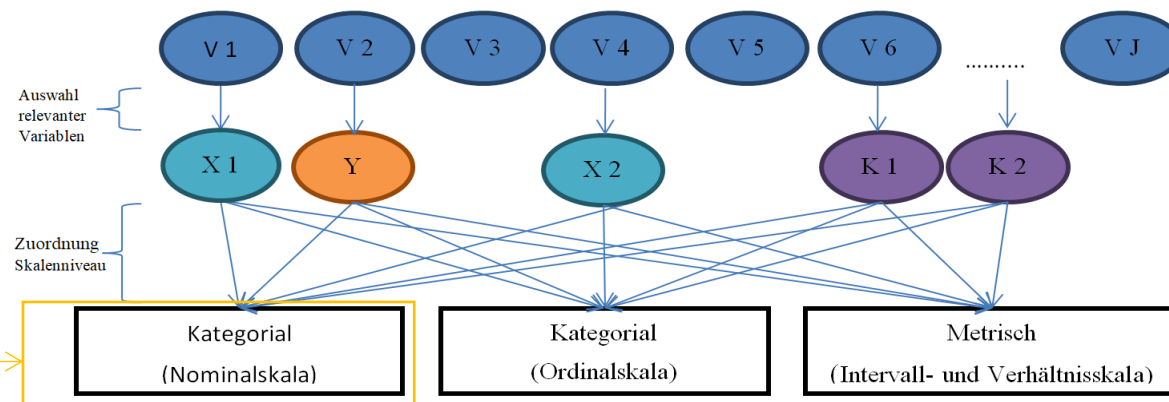
Signifikanz des Einflusses: $p = 0.0452$

Der Einfluss Non-Peergroup-Time ist somit mit einem p-Wert b von 0.452 nicht statistisch signifikant ($p \geq 0.1$)

Beispiel – Hypothese: X nominal → Y nominal

Beispiel Datensatz Teilgruppe3 – Follow-Up-Survey

H1: Personen, die **Feedback zum Alkoholkonsum in allen Gruppen auf der Party (T2)** erhalten, fordern häufiger **Informationen zur Alkoholprävention** im Follow-Up-Survey an.



Abhängige Variable Y = Informationen zur Alkoholprävention → nominales Skalenniveau
Unabhängige X = Feedback zum Alkoholkonsum in allen Gruppen auf der Party (T2) → nominales Skalenniveau



a.

Y nominal/ordinal
X_j nominal/ordinal

Kreuztabelle

Anzahl		Geschlecht		Gesamt
		weiblich	männlich	
Studienfach	Psychologie	16	8	24
	Physik	4	12	16
Gesamt		20	20	40

Chi², Cramers V

bei zwei ordinalen Variablen:
Spermans rank correlation

c.

Chi²-Unabhängigkeitstest

Fisher-Exakt-Test (2x2
Tabellen mit kleiner Stichprobe)
(nicht parametrisch)

Beispiel – Hypothese: X nominal → Y nominal

Vorbereitung: Fälle selektieren

Wir haben 4 Treatmentgruppen → Ausprägungen der Variable „Treatment“ im Datensatz
0 = kein Feedback, 1 = Individualfeedback, 2 = Gruppenfeedback, 3 = Gesamtfeedback

Wir wollen aber nur 2 Treatmentgruppen miteinander vergleichen!

Personen, die **Feedback zum Alkoholkonsum in allen Gruppen auf der Party (T2)** erhalten

→ Variable im Datensatz „Treatment“ = 2

Personen, die kein Feedback erhalten haben

→ Variable im Datensatz „Treatment“ = 0

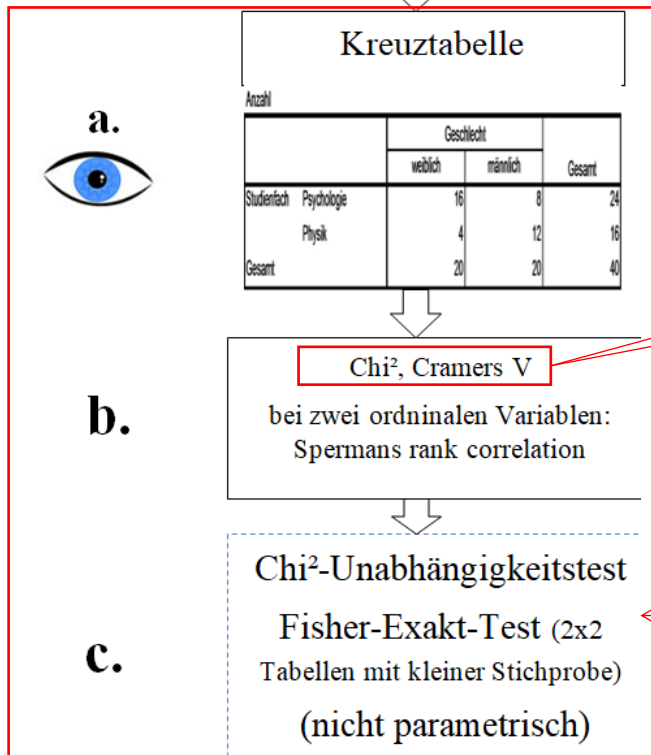
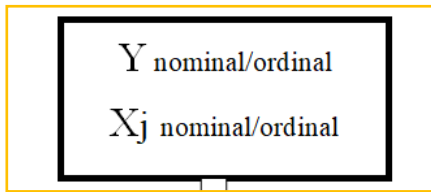
The screenshot shows the SPSS 'Fälle auswählen' (Select Cases) dialog box. The 'Auswählen' (Select) section has 'Falls Bedingung zutrifft' (If condition is satisfied) selected. The condition 'Treatment=0 | Treatment=2' is entered in the text box. The 'Ausgabe' (Output) section has 'Ausgewählte Fälle in neues Dataset kopieren' (Copy selected cases to new dataset) selected. The 'Erweiterungen' (Extensions) menu is circled in red in the top menu bar.

*In den Synthax einfügen und ausführen
Alle Personen aus Treatment 1 und 0 werden
dann gestrichen*

```
USE ALL.  
COMPUTE filter_$(Treatment=0 | Treatment=2).  
VARIABLE LABELS filter_$ 'Treatment=0 | Treatment=2 (FILTER)'.  
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.  
FORMATS filter_$ (f1.0).  
FILTER BY filter_$.  
EXECUTE.
```

Beispiel – Hypothese: X nominal → Y nominal

Kreuztabelle erstellen



Unabhängige Variable X **Feedback zum Alkoholkonsum in allen Gruppen auf der Party (T2)** in die Zeile der Kreuztabelle

Abhängige Variable **Informationen zur Alkoholprävention** Y in die Spalten der Kreuztabelle

Kreuztabellen

Zeile(n): Treatment [Treatment]

Spalten: Alkoholbezogene Informationen a...

Kreuztabellen: Statistik

Chi-Quadrat

Phi und Cramer-V

Signifikanzwerte werden zum Chi² automatisch ausgegeben

Technische Umsetzung in SPSS:

- Wählen Sie die Befehlsfolge „Analysieren“, „Deskriptive Statistiken“ → „Kreuztabellen...“. Es öffnet sich die Dialogbox „Kreuztabellen“
- Wählen Sie aus der Variablenliste die Zeilenvariable aus, und übertragen Sie diese in das Feld „Zeilen“, Übertragen Sie die Spaltenvariable in das Feld „Spalten“

Beispiel – Hypothese: X nominal → Y nominal

Kreuztabelle erstellen

Y nominal/ordinal
X_j nominal/ordinal

a.

Kreuztabelle

Anzahl		Geschlecht		Gesamt
		weiblich	männlich	
	Studienfach Psychologie	16	8	24
	Physik	4	12	16
	Gesamt	20	20	40

b.

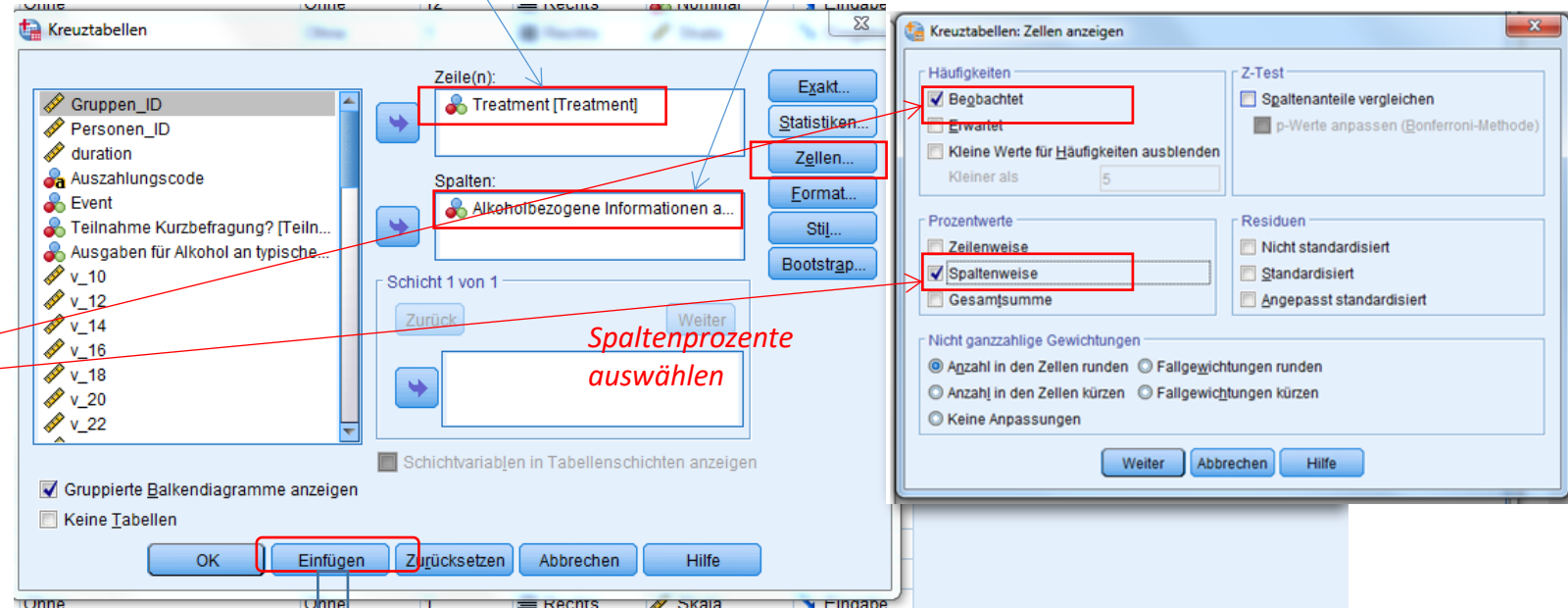
Chi², Cramers V
bei zwei ordinalen Variablen:
Sperman's rank correlation

c.

Chi²-Unabhängigkeitstest
Fisher-Exakt-Test (2x2
Tabellen mit kleiner Stichprobe)
(nicht parametrisch)

Unabhängige Variable X **Feedback zum Alkoholkonsum in allen Gruppen auf der Party (T2)** in die Zeile der Kreuztabelle

Abhängige Variable **Informationen zur Alkoholprävention** Y in die Spalten der Kreuztabelle



Spaltenprozent auswählen

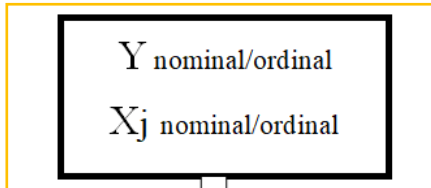
```

44 CROSSTABS
45 /TABLES=Treatment BY Wunsch_Informationen
46 /FORMAT=AVALUE TABLES
47 /STATISTICS=CHISQ PHI
48 /CELLS=COUNT COLUMN
49 /COUNT ROUND CELL
50 /BARCHART.
51
    
```


In den Synthax einfügen, Code für die Kreuztabelle (Crosstabs) markieren und hier ausführen

Beispiel – Hypothese: X nominal → Y nominal

4a) visuelle Inspektion in der Kreuztabelle



H1: Personen, die **Feedback zum Alkoholkonsum in allen Gruppen auf der Party (T2) erhalten**, fordern häufiger **Informationen zur Alkoholprävention im Follow-Up-Survey an**.

a. 

Kreuztabelle

Anzahl		Geschlecht		Gesamt
		weiblich	männlich	
	Studienfach Psychologie	16	8	24
	Physik	4	12	16
	Gesamt	20	20	40

Ergebnisse in SPSS interpretieren

4a. Ist ein Unterschied zwischen Personen mit und ohne Feedback zu sehen?

			Feedback erhalten?		
			Kein Feedback	Feedback zum Gruppenkonsum	Gesamt
Infos abgerufen?	Präventionsinfos abgerufen	Anzahl	2	2	4
		% innerhalb Feedback	11,76%	10,53%	11,10%
	Infos nicht abgerufen	Anzahl	15	17	32
		% innerhalb Feedback	88,34%	89,47%	88,90%
Gesamt	Anzahl				
	% innerhalb Feedback	17 100%	19 100%	36 100%	

b.

Chi², Cramers V
bei zwei ordinalen Variablen:
Spermans rank correlation

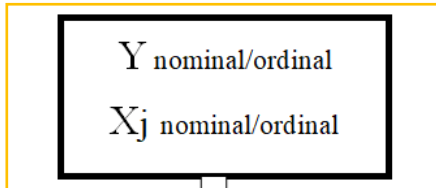
c.

Chi²-Unabhängigkeitstest
Fisher-Exakt-Test (2x2
Tabellen mit kleiner Stichprobe)
(nicht parametrisch)

Von allen Personen, die kein Feedback erhalten haben, haben 2 bzw. 11,76% die Alkoholpräventionsinformationen abgerufen. Bei den Personen mit Feedback zum Gruppenkonsum waren es mit 10,53% leicht weniger. → Ein schwacher Unterschied ist also zu sehen

Beispiel – Hypothese: X nominal → Y nominal

4b) Zusammenhangsstärke und 4c) Signifikanztest



H1: Personen, die **Feedback zum Alkoholkonsum in allen Gruppen auf der Party (T2) erhalten**, fordern häufiger **Informationen zur Alkoholprävention im Follow-Up-Survey an**.

Kreuztabelle

Anzahl		Geschlecht		Gesamt
		weiblich	männlich	
Studienfach	Psychologie	16	8	24
	Physik	4	12	16
Gesamt		20	20	40

Ergebnisse in SPSS interpretieren



b. Wie stark ist der Zusammenhang?

CramersV=0.02: Der Zusammenhang ist sehr schwach mit einem Cramers V Wert nahe 0

Chi², Cramers V
bei zwei ordinalen Variablen:
Sperman's rank correlation

Symmetrische Maße

	Wert	Näherungsw eise Signifikanz
Nominal- bzgl. Nominalmaß	Phi Cramer-V	-,020 ,020
Anzahl der gültigen Fälle	36	

Interpretation von Cramér's V nach Cohen (1988), zitiert nach Ellis (2010)

kleiner Effekt	V = 0.1
mittlerer Effekt	V = 0.3
großer Effekt	V = 0.5

Unter der Annahme der Nullhypothese (kein Zusammenhang zwischen Informationsanforderung und Gruppenfeedback) ist die Wahrscheinlichkeit eine solche Stichprobe mit CramersV-Wert von 0.014 + & Chi²-Wert von 0.14 zufällig zu ziehen 90,6%. → das über 10% (p>0.1) und somit zu hoch, um die Nullhypothese zu verwerfen

Chi²-Unabhängigkeitstest
Fisher-Exakt-Test (2x2
Tabellen mit kleiner Stichprobe)
(nicht parametrisch)

Chi-Quadrat-Tests

	Wert	df	Asymptotisch e Signifikanz (zweiseitig)	Exakte Signifikanz (2- seitig)	Exakte Signifikanz (1- seitig)
Chi-Quadrat nach Pearson	,014 ^a	1	,906		
Kontinuitätskorrektur ^b	,000	1	1,000		
Likelihood-Quotient	,014	1	,906		
Exakter Test nach Fisher				1,000	,655
Zusammenhang linear- mit-linear	,014	1	,907		
Anzahl der gültigen Fälle	36				

P-Wert für den Chi²-Unabhängigkeitstest
p > 0.1 → nicht statistisch signifikant → Nullhypothese muss beibehalten werden

p ≤ 0.1 → statistisch schwach signifikant *
p ≤ 0.05 → statistisch signifikant **
p ≤ 0.01 → statistisch stark signifikant ***

a. 2 Zellen (50,0%) haben eine erwartete Häufigkeit kleiner 5. Die minimale erwartete Häufigkeit ist 1,89.
b. Wird nur für eine 2x2-Tabelle berechnet

Beispiel – Hypothese: X nominal → Y nominal

4d. Multivariate Analyse in der logistischen Regression

4d) Bleibt/Wird der Zusammenhang zwischen **Feedback zum Alkoholkonsum in allen Gruppen auf der Party (T2)** und **Informationen zur Alkoholprävention** signifikant, wenn auf Kontrollvariablen kontrolliert wird?

Umsetzung einer logistischen Regression in SPSS

Kontrollvariablen hinzufügen

```
50  
51 LOGISTIC REGRESSION VARIABLES Wunsch_Informationen  
52 /METHOD=ENTER Gruppenfeedback Nettoeinkommen AusgabenfürAlkoholantypischemAbend  
53 /CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).  
54
```

In den Syntax einfügen, Code und ausführen

Beispiel – Hypothese: X nominal → Y nominal

4d. Multivariate Analyse in der logistischen Regression

4d) Bleibt/Wird der Zusammenhang zwischen **Feedback zum Alkoholkonsum in allen Gruppen auf der Party (T2)** und **Informationen zur Alkoholprävention** signifikant, wenn auf Kontrollvariablen kontrolliert wird?

Interpretation einer binär-logistischen Regression in SPSS

Schritt	-2 Log-Likelihood	Cox & Snell R-Quadrat	Nagelkerkes R-Quadrat
1	20,272 ^a	,104	,191

a. Schätzung beendet bei Iteration Nummer 6, weil die Parameterschätzer sich um weniger als ,001 änderten.

Interpretation Pseudo r^2 → Werte zwischen 0 und 1
 1 = perfekte Anpassung des Modell/hohe Erklärungskraft
 0 = gar keine Erklärungskraft des Regressionsmodell für die Vorhersage von Y → keine Verbesserung zur Vorhersage ohne Variablen

Interpretation: Pseudo- r^2 von 0.1 und 0.191

→ Das Modell zur Vorhersage von **Informationen zur Alkoholprävention** mit den drei Variablen hat eine höhere Erklärungskraft als das Null-Modell ohne Variablen

→ Mittelmäßige Erklärungskraft/Anpassungsgüte

	Regressionskoeffizient B	Standardfehler	Wald	df	Sig.	Exp(B)
Schritt 1 ^a Gruppenfeedback	,137	1,142	,014	1	,904	1,147
Nettoeinkommen	,000	,002	,024	1	,877	1,000
Ausgaben für Alkohol an typischem Abend	-,094	,073	1,630	1	,202	,910
Konstante	-,333	1,365	,060	1	,807	,717

a. In Schritt 1 eingegebene Variablen: Gruppenfeedback, Nettoeinkommen, Ausgaben für Alkohol an typischem Abend.

Regressionskoeffizient hat ein positives Vorzeichen → Gruppenfeedback erhöht die Wahrscheinlichkeit, Präventionsinformationen im Survey aufzurufen.

Unter der Annahme der Nullhypothese (=kein Zusammenhang zwischen Feedback und Informationsaufrufung) zufällig eine Verteilung mit Koeffizienten von 0.137 und Standardfehler von 1.142 zu finden beträgt 90,4% → Das ist über 10% und somit ist der Einfluss des **Gruppenfeedbacks** nicht statistische signifikant.