

Übungsaufgaben – Sitzung 6

Aufgabe 1

Immer mehr Menschen ist es wichtig, ethisch verantwortlich zu konsumieren. Dafür wählen einige Menschen das Mittel des Produktboykotts, d.h. sie verzichten auf den Konsum bestimmter Produkte, da sie mit den Herstellungsbedingungen, Klimafolgen etc. nicht einverstanden sind. Sie interessieren sich nun dafür, wer genau dazu tendiert, Produkte zu boykottieren. Nach theoretischen Überlegungen stellen Sie die Vermutung auf, dass Menschen, die parteipolitisch den Grünen nahestehen auch eher dazu tendieren, bestimmte Produkte zu boykottieren, da ein bedeutender Teil der Wählerschaft sich für Umwelt- und Arbeitsschutzziele engagiert und über ein hohes Einkommen verfügt, sodass diese freier in ihren Konsumentscheidungen sind. Im European Social Survey finden Sie dazu folgende Daten:

- 184 Befragte haben in den letzten 12 Monaten ein Produkt boykottiert und stehen den Grünen nahe. 744 Befragte, die ein Produkt boykottierten, weisen keine Nähe zu den Grünen auf.
- 88 Befragte geben von sich an, den Grünen nahezustehen, und in den letzten 12 Monaten kein Produkt boykottiert zu haben, während 1336 Befragte, die kein Produkt boykottiert haben, den Grünen nicht nahestehen.

a. Legen Sie eine entsprechende Kreuztabelle an. Worauf müssen Sie achten?

In einer Kreuztabelle stehen die Ausprägungen der abhängigen Variablen (Y) in den Zeilen, die der unabhängigen Variablen (X) in den Spalten. Entsprechend der theoretischen Annahme, ist die unabhängige Variable die Nähe zu den Grünen, die abhängige Variable der Boykott eines Produktes. Auch die Randsummen, d.h. die unbedingten Verteilungen sollten nicht vergessen werden. Entsprechende Kreuztabelle:

	Nähe zu Grünen (X ₁)	Keine Nähe zu Grünen (X ₂)	Σ Reihensumme
Produkt boykottiert (Y ₁)	184	744	928
Nicht boykottiert (Y ₂)	88	1336	1424
Σ Spaltensumme	272	2080	2352

Da es sich um nominale Variablen handelt, spielt die Anordnung der Ausprägungen keine Rolle. Dementsprechend sind auch folgende Lösungen richtig:

	keine Nähe zu Grünen	Nähe zu Grünen	Σ
boykottiert	744	184	928
Nicht boykottiert	1336	88	1424
Σ	2080	272	2352

Variante 2

	Nähe zu Grünen	Keine Nähe zu Grünen	Σ
Nicht boykottiert	88	1336	1424
boykottiert	184	744	928
Σ	272	2080	2352

Variante 3

	Keine Nähe zu Grünen	Nähe zu Grünen	Σ
Nicht boykottiert	1336	88	1424
boykottiert	744	184	928
Σ	2080	272	2352

Variante 4

- b. Führen Sie die Spaltenprozentuierung durch und berechnen Sie die Prozentsatzdifferenz. Hängen die Nähe zu den Grünen und die Neigung zum Produktboykott zusammen?

	Nähe zu Grünen (X ₁)	Keine Nähe zu Grünen (X ₂)	Σ Reihensumme
Produkt boykottiert (Y ₁)	$\frac{184}{272} * 100 = 67,65 \text{ [\%]}$	$\frac{744}{2080} * 100 = 35,77 \text{ [\%]}$	$\frac{928}{2352} * 100 = 39,46 \text{ [\%]}$
Nicht boykottiert (Y ₂)	$\frac{88}{272} * 100 = 32,35 \text{ [\%]}$	$\frac{1336}{2080} * 100 = 64,23 \text{ [\%]}$	$\frac{1424}{2352} * 100 = 60,54 \text{ [\%]}$
Σ Spaltensumme	100%	100%	100%

Prozentsatzdifferenz:

$$d\% = 67,65\% - 35,77\% = 31,88 \text{ [Prozentpunkte]}$$

alternativ

$$d\% = 32,35\% - 64,23\% = -31,88 \text{ [Prozentpunkte]}$$

Die Prozentsatzdifferenz zeigt hier einen Zusammenhang zwischen der Nähe zu den Grünen und dem Boykott von Produkten an, da $d\% = 31,88 \text{ PP} \neq 0$.

- c. Welches Problem bringt die Prozentsatzdifferenz mit sich? Berechnen Sie ein angemessenes, alternatives Zusammenhangsmaß und interpretieren Sie es.

Das Problem ist, dass die Prozentsatzdifferenz zwar anzeigt, dass es einen Zusammenhang gibt, nicht aber wie stark dieser ist. Eine mögliche Lösung wäre die Berechnung von Odds Ratio und dessen Normierung über Yule's Q.

$$Odds_1 = \frac{Y_{1|X1}}{Y_{2|X1}} = \frac{184}{88} = 2,09$$

$$Odds_2 = \frac{Y_{1|X2}}{Y_{2|X2}} = \frac{744}{1336} = 0,56$$

$$\text{Odds Ratio} = \frac{\text{Odds}_1}{\text{Odds}_2} = \frac{2,09}{0,56} = 3,73$$

Dass eine Person Produkte boykottiert, im Vergleich dazu, dass die Person keine Produkte boykottiert, kommt unter der Bedingung, dass diese Person den Grünen nahesteht, 3,73-mal häufiger vor, als für eine Person, die den Grünen nicht nahesteht

$$\text{Yule's } Q = \frac{OR - 1}{OR + 1} = \frac{3,73 - 1}{3,73 + 1} = 0,58$$

Es handelt sich um einen starken Zusammenhang.

Wenn die Kreuztabellen anders angelegt worden sind (siehe a.), führt dies zu anderen, ebenfalls richtigen Ergebnissen. In der Interpretation muss dann jedoch entsprechend auf die Anordnung der Ausprägung in der Kreuztabelle geachtet werden.

Variante 2:

$$\text{Odds}_1 = \frac{Y_{1|x2}}{Y_{2|x2}} = \frac{744}{1336} = 0,56$$

$$\text{Odds}_2 = \frac{Y_{1|x1}}{Y_{2|x1}} = \frac{184}{88} = 2,09$$

$$\text{Odds Ratio} = \frac{\text{Odds}_1}{\text{Odds}_2} = \frac{0,56}{2,09} = 0,27$$

$$\text{Yule's } Q = \frac{OR - 1}{OR + 1} = \frac{0,27 - 1}{0,27 + 1} = -0,57$$

Interpretation: Dass eine Person Produkte boykottiert, im Vergleich dazu, dass die Person keine Produkte boykottiert, kommt unter der Bedingung, dass sie den Grünen nicht nahesteht, nur 0,27-mal so häufig vor, als für eine Person, die den Grünen nahesteht.

Es handelt sich um einen starken Zusammenhang.

Variante 3:

$$\text{Odds}_1 = \frac{Y_{1|x2}}{Y_{2|x2}} = \frac{88}{184} = 0,48$$

$$\text{Odds}_2 = \frac{Y_{1|x1}}{Y_{2|x1}} = \frac{1336}{744} = 1,8$$

$$\text{Odds Ratio} = \frac{\text{Odds}_1}{\text{Odds}_2} = \frac{0,48}{1,8} = 0,27$$

$$\text{Yule's } Q = \frac{OR - 1}{OR + 1} = \frac{0,27 - 1}{0,27 + 1} = -0,57$$

Interpretation: Dass eine Person keine Produkte boykottiert, im Vergleich dazu, dass die Person Produkte boykottiert, kommt unter der Bedingung, dass sie den Grünen nahesteht, nur 0,27-mal so häufig vor, wie für eine Person, die den Grünen nicht nahesteht.

Es handelt sich um einen starken Zusammenhang.

Variante 4:

$$Odds_1 = \frac{Y_{1|X1}}{Y_{2|X1}} = \frac{1336}{744} = 1,8$$

$$Odds_2 = \frac{Y_{1|X2}}{Y_{2|X2}} = \frac{88}{184} = 0,48$$

$$Odds\ Ratio = \frac{Odds_1}{Odds_2} = \frac{1,8}{0,48} = 3,75$$

$$Yule's\ Q = \frac{OR - 1}{OR + 1} = \frac{3,75 - 1}{3,75 + 1} = 0,58$$

Interpretation: Dass eine Person keine Produkte boykottiert, im Vergleich dazu, dass sie Produkte boykottiert, kommt unter der Bedingung, dass sie den Grünen nicht nahesteht, 3,75-mal häufiger vor, als für eine Person, die den Grünen nahesteht.

Es handelt sich um einen starken Zusammenhang.

Wichtig: Unabhängig davon nach welcher Variante Sie Ihre Kreuztabelle angelegt haben und welche der vier Rechenwege und Ergebnisse folgten, die inhaltliche Aussage ist in allen vier Varianten die gleiche: In dieser Stichprobe, *gibt es einen Zusammenhang zwischen der ideologischen Nähe zu den Grünen und der Neigung, das Mittel des Boykotts bestimmter Produkte zu nutzen. Dabei tendieren Menschen, die den Grünen nahestehen, häufiger dazu, Produkte zu boykottieren, als Menschen, die den Grünen nicht nahestehen.*

Aufgabe 2

Sie interessieren sich für die Frage, wann Menschen eine Gesellschaft als gerecht erachten. 2018 beantworteten im European Social Survey 1395 Teilnehmer*innen die Frage, ob sie die Gesellschaft als fair empfinden, wenn Einkommen und Wohlstand gleichmäßig verteilt seien. Sie vermuten, dass die Zustimmung zu diesem Gerechtigkeitsprinzip mit der Positionierung der Befragten auf der Links-Rechts-Achse zusammenhängt. Folgende Zahlen finden Sie dazu im ESS:

- 170 Befragte stimmen der Aussage zu und sind eher links verortet. 290 Befragte stimmen der Aussage zu und sehen sich selbst in der politischen Mitte. 189 Befragte stimmen der Aussage zu und sind eher rechts auf der Achse zu finden.
- 64 der eher linken Befragten sind hinsichtlich dieser Frage unentschieden. In der politischen Mitte sind es 230 Befragte und in der politisch eher rechten Gruppe 165.
- 49 Befragte stimmen der Aussage nicht zu und stufen sich politisch eher links ein, 157 Befragte der politischen Mitte lehnen die Aussage ab und 81 Befragte, die politisch eher rechts stehen, lehnen die Aussage ab.

a. Legen Sie eine entsprechende Kreuztabelle an.

	links	Pol. Mitte	rechts	Σ
Zustimmung	170	290	189	649
unentschieden	64	230	165	459
Ablehnung	49	157	81	287
Σ	283	677	435	1395

- b. Hängen die Zustimmung zu der Aussage, eine Gesellschaft sei gerecht, wenn Einkommen und Wohlstand gleichmäßig verteilt seien und die Selbsteinstufung auf der Links-Rechts-Achse zusammen? Berechnen Sie ein angemessenes Zusammenhangsmaß und normieren Sie es.

Beide Variablen sind nominal und haben mehr als 3 Ausprägungen, dementsprechend muss Chi-Square berechnet und mit Cramers V normiert werden.

Hinweis: Die Zustimmung zu der Aussage kann auch als ordinale Variable interpretiert werden. Wenn Sie ein Zusammenhangsmaß zwischen zwei Variablen unterschiedlicher Skalenniveaus berechnen, richten Sie sich in der Wahl des Zusammenhangsmaßes nach dem niedrigeren Skalenniveau. Entsprechend bliebe es hier bei Chi-Square, sodass die Frage, ob die Zustimmung als nominale oder ordinale Variable skaliert ist, für die weitere Berechnung in dieser Aufgabe nicht wichtig ist.

Berechnung von Chi-Square:

1. Berechnung erwarteter Häufigkeiten: $e_{ij} = \frac{f_{i.} \cdot f_{.j}}{n}$

	links	Pol. Mitte	rechts	Σ
Zustimmung	$\frac{649 \cdot 283}{1395}$ = 131,66	$\frac{649 \cdot 677}{1395}$ = 314,96	$\frac{649 \cdot 435}{1395}$ = 202,38	649
unentschieden	$\frac{459 \cdot 283}{1395}$ = 93,12	$\frac{459 \cdot 677}{1395}$ = 222,75	$\frac{459 \cdot 435}{1395}$ = 143,13	459
Ablehnung	$\frac{287 \cdot 283}{1395}$ = 58,22	$\frac{287 \cdot 677}{1395}$ = 139,28	$\frac{287 \cdot 435}{1395}$ = 89,49	287
Σ	283	677	435	1395

Würden schon hier alle erwarteten Häufigkeiten den beobachteten Häufigkeiten aus der Tabelle oben entsprechen, könnte die Berechnung abgebrochen und der Schluss gezogen werden, dass kein Zusammenhang vorläge. Da dies nicht der Fall ist, wird die Berechnung fortgesetzt.

2. Für die folgenden Schritte (Berechnung der Differenz zwischen beobachteten und erwarteten Häufigkeiten, Quadrierung, Teilen durch die erwartete Häufigkeit) wird die Hilfstabelle angelegt

i Wert von Y Zustimmung	j Wert von pol. Einstufung	f_{ij}	e_{ij}	$f_{ij} - e_{ij}$	$(f_{ij} - e_{ij})^2$	$\frac{(f_{ij} - e_{ij})^2}{e_{ij}}$
Zustimmung	Links	170	131,66	38,34	1496,96	11,37
Zustimmung	Mitte	290	314,96	-24,96	623,00	1,98
Zustimmung	Rechts	189	202,38	-13,38	179,02	0,88
Unentschieden	Links	64	93,12	-29,12	847,97	9,11
Unentschieden	Mitte	230	222,75	7,25	52,49	0,24
Unentschieden	Rechts	165	143,13	21,87	478,30	3,34
Ablehnung	Links	49	58,22	-9,22	85,00	1,46
Ablehnung	Mitte	157	139,28	17,72	314,00	2,25
Ablehnung	Rechts	81	89,49	-8,49	72,08	0,81
						$\Sigma = 31,44 = \chi^2$

Hinweis: Wenn Sie einige der Zwischenschritte zusammengefasst haben, kann es durch die hier vorgenommene Rundung zu Abweichungen kommen. Sofern Sie korrekt gerundet haben, sind diese Ergebnisse dann auch korrekt.

$$\text{Cramers } V = \sqrt{\frac{X^2}{N * (q - 1)}} = \sqrt{\frac{31,44}{1395 * (3 - 1)}} = 0,11$$

Interpretation: Die Zustimmung zu der Aussage, die Gesellschaft sei gerecht, wenn Einkommen und Wohlstand gleichmäßig verteilt seien sowie die Selbsteinstufung auf der Links-Rechts-Achse scheinen zusammenzuhängen, da $X^2 = 31,44 \neq 0$. Anhang der Normierung erkennt man jedoch, dass es sich nur um einen schwachen Zusammenhang handelt.

Aufgabe 3 – mit SPSS

Die CDU wirbt in ihren Wahlkampagnen oft damit, Kriminalität konsequent bekämpfen und die Polizei stärken zu wollen. Sie vermuten daher, dass ein hohes Vertrauen gegenüber der Polizei auch mit einer Wahlpräferenz für die CDU einhergeht. Dies wollen Sie anhand des Datensatzes der Prä-Evaluation überprüfen. Öffnen Sie diesen mit SPSS.

- a. Suchen Sie die Variablen „Wahlpräferenz“ und „Vertrauen in politische Institutionen - Polizei“. Machen Sie sich mit der Skalierung, d.h. dem Skalenniveau und den möglichen Ausprägungen vertraut.

Wahlpräferenz

- Variablenname: A1
- Skalenniveau: nominal
- Mögliche Ausprägungen: 1=CDU; 2=SPD; 3=Grüne; 4=AfD; 5=Die Linke; 6=FDP

Vertrauen in politische Institutionen – Polizei

- Variablenname: A4c
- Skalenniveau: ordinal
- Mögliche Ausprägungen: 0=“vertraue gar nicht“ bis 10=“vertraue voll und ganz“

- b. Legen Sie mit Hilfe von SPSS eine entsprechende Kreuztabelle zwischen den beiden Variablen an und lassen Sie SPSS die Spaltenprozentuierung durchführen. Können Sie bereits erste Erkenntnisse gewinnen? Welche Probleme kommen dabei auf? Diskutieren Sie die Probleme und überlegen Sie sich eine angemessene Lösung.

Analysieren > Deskriptive Statistiken > Kreuztabellen.

Sie stellen die Hypothese auf, dass ein hohes Vertrauen in die Polizei mit einer stärkeren Wahlpräferenz für die CDU einhergeht. Entsprechend ist das Vertrauen in die Polizei die unabhängige Variable. Ziehen Sie diese Variable von der Liste links in das Feld für die Spalten rechts. Die Wahlpräferenz ist die abhängige Variable. Ziehen Sie diese in das Feld für die Zeilen. Klicken Sie auf „Zellen“ und stellen Sie sicher, dass das Häkchen unter Prozentwerte > Spaltenweise gesetzt ist. Bestätigen Sie Ihre Auswahl.

Oder wählen Sie: Datei > Neu > Syntax

Geben Sie im erscheinenden Editor folgende Syntax ein, markieren diese und führen Sie mit dem grünen Pfeilknopf oben aus:

```
CROSSTABS
  /TABLES=A1 BY A4c
  /FORMAT=AVALUE TABLES
  /CELLS=COUNT COLUMN
  /COUNT ROUND CELL.
```

Folgende Kreuztabelle erscheint:

Wahlpräferenz * Vertrauen in politische Institutionen - Polizei Kreuztabelle														
			vertraue gar nicht	1,00	2,00	3,00	4,00	5,00	6,00	7,00	8,00	9,00	vertraue voll und ganz	Gesamt
Wahlpräferenz	CDU	Anzahl	0	0	2	1	1	3	3	7	7	0	1	25
		% von Vertrauen in politische Institutionen - Polizei	0,0%	0,0%	18,2%	9,1%	8,3%	13,6%	15,0%	36,8%	24,1%	0,0%	11,1%	15,9%
SPD		Anzahl	1	1	2	1	1	5	5	3	9	5	1	34
		% von Vertrauen in politische Institutionen - Polizei	16,7%	16,7%	18,2%	9,1%	8,3%	22,7%	25,0%	15,8%	31,0%	41,7%	11,1%	21,7%
Grüne		Anzahl	1	3	2	4	7	8	9	3	11	4	5	57
		% von Vertrauen in politische Institutionen - Polizei	16,7%	50,0%	18,2%	36,4%	58,3%	36,4%	45,0%	15,8%	37,9%	33,3%	55,6%	36,3%
AfD		Anzahl	0	0	0	0	0	0	0	0	1	0	0	1
		% von Vertrauen in politische Institutionen - Polizei	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	3,4%	0,0%	0,0%	0,6%
Die Linke		Anzahl	4	2	4	5	3	5	2	3	1	1	1	31
		% von Vertrauen in politische Institutionen - Polizei	66,7%	33,3%	36,4%	45,5%	25,0%	22,7%	10,0%	15,8%	3,4%	8,3%	11,1%	19,7%
FDP		Anzahl	0	0	1	0	0	1	1	3	0	1	1	8
		% von Vertrauen in politische Institutionen - Polizei	0,0%	0,0%	9,1%	0,0%	0,0%	4,5%	5,0%	15,8%	0,0%	8,3%	11,1%	5,1%
9,00		Anzahl	0	0	0	0	0	0	0	0	0	1	0	1
		% von Vertrauen in politische Institutionen - Polizei	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	8,3%	0,0%	0,6%
Gesamt		Anzahl	6	6	11	11	12	22	20	19	29	12	9	157
		% von Vertrauen in politische Institutionen - Polizei	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Die ausgegebene Kreuztabelle ist aufgrund der vielen Ausprägungen der beiden Variablen sehr komplex. So kann nur schwierig ein Überblick gewonnen werden. Eine mögliche Lösung wäre jeweils eine Gruppierung der beiden Variablen, sodass eine übersichtlichere Kreuztabelle entsteht.

Da in dieser Aufgabe nur theoretisch interessant ist, inwieweit das Vertrauen in die Polizei die Wahlpräferenz zur CDU beeinflusst, kann die Variable in einen Dummy umcodiert werden, mit der Ausprägung 1 für die Wahl der CDU und 0 für die Wahl einer anderen Partei. Wählen Sie hierzu in SPSS:

Transformieren > Umcodieren in andere Variable.

Im erscheinenden Fenster ziehen Sie die Variable „Wahlpräferenz“ in die Mitte. Legen Sie in dem Fenster, das erscheint, wenn Sie „Alte und neue Werte“ klicken fest, dass Fälle, die die alte Ausprägung 1 aufweisen, in der neuen Dummy-Variable ebenfalls eine 1 zugewiesen bekommen. Legen Sie außerdem fest, dass alle anderen Werte in der neuen Variablen zu einer 0 werden. Bestätigen Sie mit „Weiter“. Geben Sie Ihrer Variablen einen Namen, z.B. „CDUWahl“, und beschriften Sie sie entsprechend. Bestätigen Sie dann mit „Ok“.

Arbeiten Sie alternativ mit folgender Syntax. Geben Sie sie in den Editor ein, markieren Sie sie und führen Sie sie mit einem Klick auf den Pfeilknopf oben aus.

```
RECODE A1 (1=1) (ELSE=0) INTO CDUWahl.
VARIABLE LABELS CDUWahl 'Wahlpräferenz für die CDU'.
EXECUTE.
```

Achtung: Falls Sie den Variablen andere Namen geben, müssen Sie die Syntax entsprechend anpassen!

Für die Variable „Vertrauen in politische Institutionen – Polizei“ wären verschiedene Gruppierungen möglich. Unter anderem folgende:

Alte Ausprägung	Neue Ausprägung	
	Beschreibung	Quantitative Zuordnung
0 - 3	Wenig Vertrauen	1
4 - 7	Mäßig Vertrauen	2
8 - 10	Starkes Vertrauen	3

Gehen Sie dazu ähnlich vor wie zuvor. Wählen sie wieder:

Transformieren > Umcodieren in andere Variable.

In dem erscheinenden Fenster wählen Sie die Variable „Vertrauen in politische Institution – Polizei“ aus und ziehen Sie sie in die Mitte. Dann wählen Sie unter „Alte und neue Werte“ jeweils die Zuweisung der Werte vor und bestätigen am Ende mit weiter. Legen Sie einen Variablennamen, z.B. „VerPoGru“ und eine Beschreibung („Vertrauen in die Polizei – Gruppiert) fest und bestätigen Sie mit „Weiter“.

Alternativ steht Ihnen auch hier der Weg per Syntax offen:

```
RECODE A4c (Lowest thru 3=0) (4 thru 7=1) (8 thru 10=2) INTO VerPoGrup.
VARIABLE LABELS VerPoGrup 'Vertrauen in die Polizei - Gruppiert'.
EXECUTE.
```

Bestenfalls nehmen Sie noch eine Beschriftung der Ausprägungen der neuen Variablen vor. Gehen Sie dazu in die Variablenansicht. Klicken Sie in das Feld in der Spalte „Werte“ bei den neuen Variablen und dann auf das erscheinende Kästchen mit den drei Punkten. Nehmen Sie im erscheinenden Fenster die Beschriftung vor. Legen Sie in der Spalte „Fehlend“ noch die 99 als fehlenden Wert fest.

Legen Sie nun eine Kreuztabelle zwischen den neuen beiden Variablen an. Wählen Sie dazu entweder:

Analysieren > Deskriptive Statistiken > Kreuztabellen.

Ziehen Sie die Präferenz für die Wahl der CDU in die Zeilen, das gruppierte Vertrauen in die Polizei in die Spalten. Klicken Sie auf „Zellen“ und stellen Sie sicher, dass das Häkchen unter Prozentwerte > Spaltenweise gesetzt ist. Bestätigen Sie Ihre Auswahl. Oder arbeiten Sie mit dieser Syntax:

```
CROSSTABS
  /TABLES=CDUWahl BY VerPoGrup
  /FORMAT=AVALUE TABLES
  /CELLS=COUNT COLUMN
  /COUNT ROUND CELL.
```

Folgende Kreuztabelle erscheint:

Wahlpräferenz für die CDU * Vertrauen in die Polizei - Gruppiert Kreuztabelle

			Vertrauen in die Polizei - Gruppiert			Gesamt
			Wenig Vertrauen	Mäßiges Vertrauen	Hohes Vertrauen	
Wahlpräferenz für die CDU	keine Präferenz für die CDU	Anzahl	35	62	45	142
		% von Vertrauen in die Polizei - Gruppiert	92,1%	81,6%	84,9%	85,0%
	Präferenz für die CDU	Anzahl	3	14	8	25
		% von Vertrauen in die Polizei - Gruppiert	7,9%	18,4%	15,1%	15,0%
Gesamt		Anzahl	38	76	53	167
		% von Vertrauen in die Polizei - Gruppiert	100,0%	100,0%	100,0%	100,0%

Mit dieser Tabelle können Sie einen besseren Überblick erhalten. Allgemein haben nur wenige der Studierenden damals eine Präferenz für die CDU gehabt (15%). Es scheint sich ein Zusammenhang abzuzeichnen: Die allermeisten derjenigen, die wenig Vertrauen in die Polizei haben, haben auch keine Präferenz für die CDU. Die Unterschiede zwischen den bedingten Verteilungen, sind gering, aber sichtbar: 81,6% mitmäßigem Vertrauen haben keine Präferenz für die CDU. Unter den Studierenden mit wenig Vertrauen sind es 92,1%. Wahrscheinlich wird es sich um einen geringen Zusammenhang handeln.

- c. Berechnen Sie nun mit Hilfe von SPSS ein angemessenes Zusammenhangsmaß und dessen Normierung. Interpretieren Sie Ihre Ergebnisse.

Da es sich um eine nominale und eine ordinale Variable handelt sowie eine 2x3 Kreuztabelle, muss Chi-Square berechnet werden.

Wählen Sie dazu:

Analysieren > Deskriptive Statistiken > Kreuztabellen.

Ziehen Sie die Variable die neue gruppierte Variable, die das Vertrauen in die Polizei misst, in die Spalten und die Präferenz für die CDU in die Zeilen. Klicken Sie auf „Statistiken“ und setzen Sie das Häkchen bei Chi-Square sowie bei „Chi und Cramer V“. Bestätigen Sie mit „Weiter“, dann mit „Ok“.

Alternativ geben Sie folgende Syntax ein:

```
CROSSTABS
  /TABLES=CDUWahl BY VerPoGrup
  /FORMAT=AVALUE TABLES
  /STATISTICS=CHISQ PHI
  /COUNT ROUND CELL.
```

Der für Chi-Square ausgegebene Wert liegt bei rund 2,21, d.h. die Präferenz für die CDU wird in dieser Stichprobe tatsächlich durch das Vertrauen in die Polizei bedingt. Cramer's V liegt ungefähr bei 0,12. Es handelt sich um einen schwachen Zusammenhang.

Ergänzender Hinweis. Durch jede Gruppierung gehen Informationen verloren. Dadurch ergeben sich auch bei der Berechnung von Zusammenhangsmaßen Verzerrungen. Wenn Sie interessehalber Chi-Square und Cramer's V für den Zusammenhang der ursprünglichen Variablen A4c „Vertrauen in politische Institutionen – Polizei“ mit ihren 11 Ausprägungen und dem Dummy für die CDU-Präferenz, erhalten Sie: Chi-Square = 12,47 und Cramer's V = 0,27. Das würde auf einen mittleren Zusammenhang hindeuten.

Aufgabe 4 – mit SPSS

Gemeinhin wird gesagt, dass Personen, die sich als männlich identifizieren, selbstbewusster seien als Personen, die sich nicht als männlich identifizieren. Falls dies stimmt, wäre es folgerichtig, wenn sich

die männlichen Studierenden häufiger zu den Leistungsstärkeren ihres Studiengangs zählen würden als die nicht männlichen Studierende.

- a. Überprüfen Sie diese Vermutung. Suchen Sie dazu zunächst die Variablen „Geschlecht“ und "Ich zähle mich zu den TOP XY Prozent meines Studiengangs“ heraus.

Geschlecht

- Variablenname: A8
- Skalenniveau: nominal
- Mögliche Ausprägungen: 1=weiblich; 0=männlich

„Erwartung 4 - Ich würde mich unter die Top ?? meines Jahrgangs einordnen“

- Variablenname: E4
- Skalenniveau: metrisch
- Mögliche Ausprägungen: 0 bis 100 [Prozent].

- b. Warum ist es hier *nicht* sinnvoll, direkt eine Kreuztabelle anzulegen? Wie können Sie diesem Problem begegnen. Schlagen Sie eine Lösungsstrategie vor und wenden Sie diese in SPSS an, sodass Sie ein entsprechendes Zusammenhangsmaß berechnen können.

Variable E4 ist metrisch skaliert. Entsprechend hat Sie sehr viele Ausprägungen. Eine Kreuztabelle würde sehr breit und unübersichtlich werden. Auch hier würde eine Gruppierung der Variablen E4 eine mögliche Lösung bieten.

Folgende Gruppierung wäre möglich. Wie immer wären auch andere Gruppierungen möglich, abhängig von Ihrer theoretischen Überlegung.

Alte Ausprägung	Neue Ausprägung	
	Beschreibung: Ich zähle mich zu	Quantitative Zuordnung
Bis 25	... den Leistungstärksten	1
26 - 50	... den Leistungsstarken	2
51 - 75	... dem Mittelfeld	3
76 - 100	... den Leistungsschwächeren	4

Wählen Sie dazu:

Transformieren > Umcodieren in andere Variable.

Variable E4 hinzufügen. „Alte und neue Werte“ klicken. Im erscheinenden Pop-Up nun die Neucodierung vornehmen. Mit „Weiter“ bestätigen. Geben Sie der Variablen einen Namen, zum Beispiel „Leistungsgruppen“ und fügen Sie eine Beschriftung ein, wie etwa: „Ich zähle mich unter die TOP ?? (E4) – Gruppiert“.

Alternativ führen Sie die Gruppierung mit folgender Syntax durch:

```
RECODE E4 (Lowest thru 25=0) (26 thru 50=1) (51 thru 75=2) (76 thru 100=3) INTO Leistungsgruppen.
VARIABLE LABELS Leistungsgruppen 'Ich zähle mich unter die Top ?? Prozent (E4) - Gruppiert'.
EXECUTE.
```

Legen Sie auch für diese Variable die Codierung fehlender Werte fest.

Erstellen Sie nun die Kreuztabelle inklusive Spaltenprozentuierung und lassen Sie sich Chi-Square und Cramer's V ausgeben. Wählen Sie dazu:

Analysieren > Deskriptive Statistiken > Kreuztabellen.

Ziehen Sie die neue, gruppierte Variable „Leistungsgruppen“ in die Zeilen, die Variable Geschlecht in die Spalten. Stellen Sie unter „Statistiken“ sicher, dass Chi-Square und Cramer's V ausgegeben werden und unter „Zellen“, dass die Spaltenprozentuierung durchgeführt wird. Bestätigen Sie mit okay.

Oder verwenden Sie folgende Syntax:

```
CROSSTABS   /TABLES=Leistungsgruppen BY A8  
/FORMAT=AVALUE TABLES  
/STATISTICS=CHISQ PHI  
/CELLS=COUNT COLUMN  
/COUNT ROUND CELL.
```

c. Hält die These Ihrer Überprüfung stand? Interpretieren Sie die Ergebnisse.

Zunächst fällt unter Betrachtung der Spaltenprozentuierung auf, dass sich die Einordnung zu den Leistungsgruppen unterscheidet. Am deutlichsten ist dies in der obersten Leistungsgruppe zu sehen, zu der sich 16,2% der Männer, aber nur 3,8% der Frauen zählen. Entsprechend ist ein Zusammenhang nicht unwahrscheinlich. Mit dem Chi-Square-Wert von 7,16 kann sicher von einem Zusammenhang ausgegangen werden. Entsprechend des Wertes von 0,22 für Cramer's V kann von einem mittlerem Zusammenhang ausgegangen werden.

Wenn die Selbsteinstufung zu einer Leistungsgruppe als Zeichen von Selbstvertrauen interpretiert werden kann, dann scheint die Hypothese, dass es einen Zusammenhang zwischen der Geschlechtsidentität und dem Selbstvertrauen gibt, zu bestätigen.