

Übungsaufgaben – Sitzung 7

Aufgabe 1

Sie untersuchen, wie sich die Zufriedenheit mit der Regierung erklären lässt. Sie stellen fest, dass in Wahlkämpfen oft mit dem Thema Bildung geworben wird. Da Sie vermuten, dass Bildungsinstitutionen für die meisten Bürger*innen relevant sind und zugleich eine staatliche Institution sind, stellen Sie folgende Arbeitshypothese auf: Je zufriedener Menschen mit den Bildungsinstitutionen ihres Landes, desto zufriedener sind sie auch mit ihrer Regierung. Diese Hypothese wollen Sie für Deutschland anhand des European Social Survey überprüfen, in dem Sie für das Jahr 2018 folgende (gruppierte) Angaben finden:

- 199 Befragte sind sowohl mit der Regierung als auch mit dem Bildungssystem zufrieden. 151 Befragte sind zufrieden mit der Regierung und mittelmäßig zufrieden mit dem Bildungssystem. 43 Befragte sind zufrieden mit der Regierung, aber unzufrieden mit dem Bildungssystem.
- Ferner sind 310 Befragte mittelmäßig zufrieden mit der Regierung und zugleich zufrieden mit dem Bildungssystem. 481 Befragte sind sowohl mit der Regierung als auch mit dem Bildungssystem mittelmäßig zufrieden. 232 Befragte sind mittelmäßig zufrieden mit der Regierung und unzufrieden mit dem Bildungssystem.
- 148 Befragte sind mit der Regierung unzufrieden, aber mit dem Bildungssystem zufrieden. 330 mit der Regierung unzufriedene Befragte sind mit dem Bildungssystem mittelmäßig zufrieden. Zudem finden sich in der Stichprobe 364 Befragte, die sowohl mit der Bildung als auch mit der Regierung unzufrieden sind.

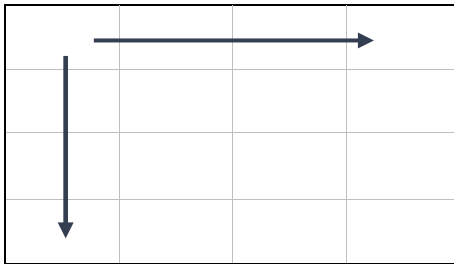
a. Legen Sie eine entsprechende Kreuztabelle an.

Zufriedenheit mit Bildungssystem Zufriedenheit mit der Regierung	zufrieden	Mittelmäßig zufrieden	unzufrieden	Σ Reihensumme
zufrieden	199	151	43	393
Mittelmäßig zufrieden	310	481	232	1023
unzufrieden	148	330	364	842
Σ Spaltensumme	657	962	639	2258

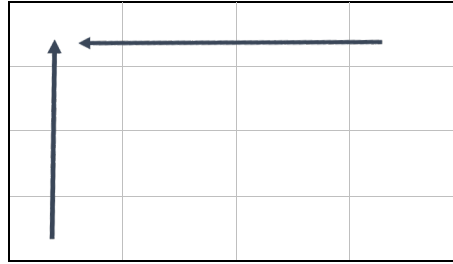
Zufriedenheit mit Bildungssystem Zufriedenheit mit der Regierung	unzufrieden	Mittelmäßig zufrieden	zufrieden	Σ Reihensumme
unzufrieden	364	330	148	842
Mittelmäßig zufrieden	232	481	310	1023
Zufrieden	43	151	199	393
Σ Spaltensumme	639	962	657	2258

Alternative Variante

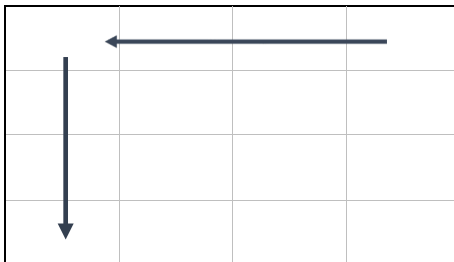
Beide Kreuztabellen sind korrekt. Achten Sie beim Anlegen der Kreuztabellen darauf, dass Sie in der Anordnung der Ausprägungen für beide Variablen die gleiche Logik, d.h. auf- oder absteigend, wählen. Eine gegenläufige Logik erschwert die Interpretation später.



Beide aufsteigend (korrekt)



Beide absteigend (korrekt)



Eine aufsteigend, eine absteigend (möglich, aber Interpretation erschwert)

- b. Berechnen Sie ein ordinales Zusammenhangsmaß. Hängen die Zufriedenheit mit dem Bildungssystem und die Zufriedenheit mit der Regierung zusammen?

Beide Variablen sind ordinal skaliert. Es kämen daher Goodman und Kruskal's γ , Kennedalls τ_B und Somer's d sowie Spearman's r infrage. Da eine Vermutung bezüglich der Richtung des Zusammenhangs vorliegt, eignet sich Somer's d jedoch am besten. Spearman's r bringt theoretische Probleme mit sich, die Verwendung sollte daher vermieden werden.

$$\text{Somers } d = \frac{N_c - N_D}{N_c + N_D + T_Y}$$

Konkordante Paare N_c

$$\begin{aligned} &= 199 * (481 + 232 + 330 + 364) + 151 * (232 + 364) + 310 \\ &\quad * (330 + 364) + 481 * 364 = 279\,993 + 89\,996 + 215\,140 + 175\,084 \\ &= \mathbf{760\,213} \end{aligned}$$

Diskordante Paare N_D

$$\begin{aligned} &= 43 * (310 + 481 + 148 + 330) + 151 * (310 + 148) + 232 \\ &\quad * (148 + 330) + 481 * 148 = 54\,567 + 69\,158 + 110\,896 + 71\,188 \\ &= \mathbf{305\,809} \end{aligned}$$

Auf Y verbundene Paare T_Y

$$\begin{aligned} &= 199 * (151 + 43) + 151 * 43 + 310 * (481 + 232) + 481 * 232 + 148 \\ &\quad * (330 + 364) + 330 * 364 \\ &= 38\,606 + 6\,493 + 221\,030 + 111\,592 + 102\,712 + 120\,120 \\ &= \mathbf{600\,553} \end{aligned}$$

$$\text{Somers } d = \frac{760\,213 - 305\,809}{760\,213 + 305\,809 + 600\,553} = \mathbf{0,27}$$

Die Zufriedenheit mit dem Bildungssystem und die Zufriedenheit mit der Regierung hängen zusammen. Es handelt sich um einen mittleren Zusammenhang.

Alternativ, wenn keine Vermutung über die Richtung des Zusammenhangs bestanden hätte, hätte man auch Goodman und Kruskal's γ oder Kendall's τ_B berechnen können. Für Letzteres benötigen Sie noch die auf X verbundenen Paare

Auf X verbundene Paare

$$= 199 * (310 + 148) + 310 * 148 + 151 * (481 + 330) + 481 * 330 + 43 * (232 + 364) + 232 * 364 \\ = 91\,142 + 45\,880 + 122\,461 + 158\,730 + 25\,628 + 84\,448 = \mathbf{528\,289}$$

$$\text{Kendalls } \tau_B = \frac{N_c - N_D}{\sqrt{(N_c + N_D + T_X) * ((N_c + N_D + T_Y))}} \\ = \frac{760\,213 - 305\,809}{\sqrt{(760\,213 + 305\,809 + 528\,289) * (760\,213 + 305\,809 + 600\,553)}} = \mathbf{0,28}$$

$$\text{Goodmans and Kruskals } \gamma = \frac{N_c - N_D}{N_c + N_D} = \frac{760\,213 - 305\,809}{760\,213 + 305\,809} = \mathbf{0,43}$$

Kendall's τ_B lässt ebenfalls auf einen mittelstarken Zusammenhang schließen, wohingegen Goodman und Kruskal's γ den Zusammenhang überschätzt. Grund ist die Vernachlässigung der verbundenen Paare.

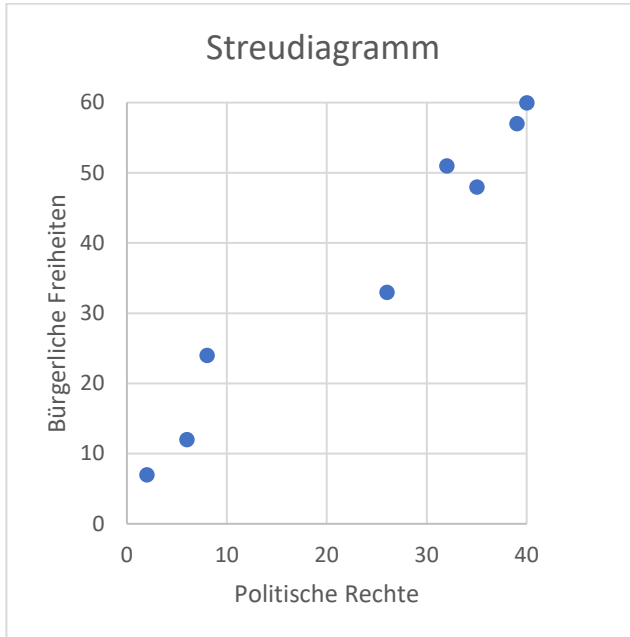
Aufgabe 2

Die Nichtregierungsorganisation Freedom House versucht jährlich zu quantifizieren, wie es um die Freiheit in verschiedenen Ländern der Welt steht. Dabei wird zum einen versucht zu messen, wie gut Bürger*innen in einem Staat politische Rechte, z.B. Wahl- oder Partizipationsrechte, wahrnehmen können. Zum anderen drückt Freedom House in einem weiteren Wert aus, wie es um die bürgerlichen Freiheiten, z.B. Meinungs-, Rede oder Versammlungsfreiheit, in einem Staat steht. Die Skala für die politischen Rechte reicht von 0 bis 40, die Skala für die bürgerlichen Freiheiten von 0 bis 60. In beiden Fällen zeichnen höhere Werte einen höheren Grad an Freiheit aus. Sie fragen sich, ob der Zugang zu politischen Rechten und der Grad an bürgerlichen Freiheiten in einem Staat zusammenhängen und schauen sich dafür die Werte für 8 Staaten an.

Staat	Politische Rechte	Bürgerliche Freiheiten
USA	32	51
Mali	8	24
Norwegen	40	60
El Salvador	26	33
Belgien	39	57
Ägypten	6	12
Rumänien	35	48
Aserbaidshan	2	7

- Berechnen Sie, ob es einen Zusammenhang gibt. Wählen Sie dafür die Kovarianz und den Korrelationskoeffizienten Pearson's r aus. Interpretieren Sie Ihr Ergebnis.

Beide Variablen sind grundsätzlich metrisch skaliert. Bei metrischen Variablen empfiehlt es sich, zunächst ein Streudiagramm anzulegen, um Übersicht zu gewinnen. Das ist von Ihnen hier nicht gefordert, soll aber für ein besseres Verständnis einmal gezeigt werden:



Die Werte scheinen fast perfekt einer steigenden Gerade zu folgen. Es ist zu vermuten, dass ein starker Zusammenhang besteht, da mit steigenden Werten für politische Rechte auch steigende Werte für bürgerliche Freiheiten einhergehen.

$$\text{Persons } r = \frac{\text{cov}(x; y)}{s_x * s_y}$$

$$\text{cov}(x; y) = \frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y})$$

Für die Berechnung der Kovarianz müssen zunächst die arithmetischen Mittelwerte für x und y berechnet werden:

$$\bar{x} = \frac{1}{n} \sum x_i$$

$$\bar{x} = \frac{1}{8} * (32 + 8 + 40 + 26 + 39 + 6 + 35 + 2) = \frac{1}{8} * 188 = 23,5$$

$$\bar{y} = \frac{1}{8} * (51 + 24 + 60 + 33 + 57 + 12 + 48 + 7) = \frac{1}{8} * 292 = 36,5$$

Für die Berechnung der Kovarianz kann am besten eine Hilfstabelle angelegt werden.

Staat	Politische Rechte	Bürgerliche Freiheiten	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x}) * (y_i - \bar{y})$
USA	32	51	8,5	14,5	123,25
Mali	8	24	-15,5	-12,5	193,75
Norwegen	40	60	16,5	23,5	387,75
El Salvador	26	33	2,5	-3,5	-8,75
Belgien	39	57	15,5	20,5	317,75
Ägypten	6	12	-17,5	-24,5	428,75
Rumänien	35	48	11,5	11,5	132,25
Aserbaidshjan	2	7	-21,5	-29,5	634,25

$$\begin{aligned} cov(x; y) &= \frac{1}{8}(123,25 + 193,75 + 387,75 + (-8,75) + 317,75 + 428,75 + 132,25 \\ &\quad + 634,25) = \frac{1}{8} * 2209 = \mathbf{276,13} \end{aligned}$$

Anhand der Kovarianz lässt sich schon erkennen, dass die beiden Variablen „Politische Rechte“ und „Bürgerliche Freiheiten“ zusammenhängen, da $cov(x; y) = 276,13 \neq 0$.

Für die Berechnung von Pearson's r bedarf es noch der beiden Standardabweichungen:

$$s_x = \sqrt{var(x)} = \sqrt{\frac{1}{n} \sum (x_i - \bar{x})^2}$$

s_x

$$\begin{aligned} &= \sqrt{\frac{1}{8}((32 - 23,5)^2 + (8 - 23,5)^2 + (40 - 23,5)^2 + (26 - 23,5)^2 + (39 - 23,5)^2 + (6 - 23,5)^2 + (35 - 23,5)^2 + (2 - 23,5)^2)} \\ &= \sqrt{\frac{1}{8}(72,25 + 240,25 + 272,25 + 6,25 + 240,25 + 306,25 + 132,25 + 462,25)} = \sqrt{\frac{1}{8} * 1732} = \mathbf{14,71} \end{aligned}$$

s_y

$$\begin{aligned} &= \sqrt{\frac{1}{8}((51 - 36,5)^2 + (24 - 36,5)^2 + (60 - 36,5)^2 + (33 - 36,5)^2 + (57 - 36,5)^2 + (12 - 36,5)^2 + (48 - 36,5)^2 + (7 - 36,5)^2)} \\ &= \sqrt{\frac{1}{8}(210,25 + 156,25 + 552,25 + 12,25 + 420,25 + 600,25 + 132,25 + 870,25)} = \sqrt{\frac{1}{8} * 2954} = \mathbf{19,22} \end{aligned}$$

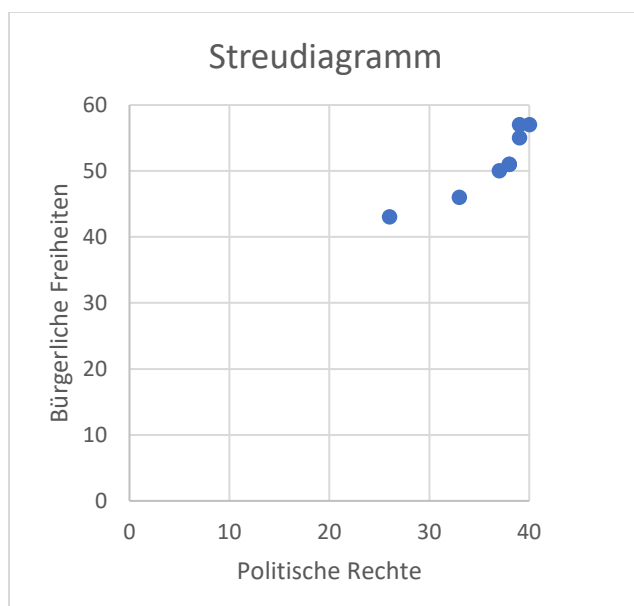
$$Person's r = \frac{cov(x; y)}{s_x * s_y} = \frac{276,13}{14,71 * 19,22} = \mathbf{0,98}$$

Aufgrund eines sehr hohen Pearson's r Wertes kann von einer sehr starken Korrelation ausgegangen werden.

- b. Die obige Auswahl an Staaten kommt Ihnen willkürlich und wenig vergleichbar vor. Sie entscheiden sich, den Zusammenhang noch einmal für 8 Staaten der Europäischen Union zu untersuchen, da diese Ihnen vergleichbarer erscheinen. Berechnen Sie erneut den Korrelationskoeffizienten r . Was fällt im Vergleich zur ersten Aufgabe auf? Diskutieren Sie Unterschiede, gerade auch im Hinblick auf jene Maße, die der Berechnung der Kovarianz zugrunde liegen.

Staat	Politische Rechte	Bürgerliche Freiheiten
Belgien	39	57
Bulgarien	33	46
Dänemark	40	57
Deutschland	39	55
Frankreich	38	51
Griechenland	37	50
Litauen	38	51
Ungarn	26	43

Auch hier kann zunächst ein Streudiagramm für einen ersten Überblick angelegt werden (nicht gefordert):



Die Fälle scheinen im Hinblick auf beide Merkmale sehr homogen zu sein und ballen sich auf beiden Variablen im Bereich hoher Werte. Weiterhin kann auch hier ein Trend vermutet werden: Je höher die Werte auf der Variablen „Politische Rechte“, desto höher auch die Werte auf der Variablen „Bürgerliche Freiheiten“.

Dies kann noch durch die Berechnung der Kovarianz und des Korrelationskoeffizienten weiter überprüft werden:

$$\text{cov}(x; y) = \frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y})$$

Berechnung der arithmetischen Mittelwerte

$$\bar{x} = \frac{1}{8} * (39 + 33 + 40 + 39 + 38 + 37 + 38 + 26) = \frac{1}{8} * 290 = 36,25$$

$$\bar{y} = \frac{1}{8} * (57 + 46 + 57 + 55 + 51 + 50 + 51 + 43) = \frac{1}{8} * 410 = 51,25$$

Anlegen der Hilfstabelle zur Berechnung der Kovarianz:

Staat	Politische Rechte	Bürgerliche Freiheiten	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x}) * (y_i - \bar{y})$
Belgien	39	57	2.75	5.75	15.81
Bulgarien	33	46	-3.25	-5.25	17.06
Dänemark	40	57	3.75	5.75	21.56
Deutschland	39	55	2.75	3.75	10.31
Frankreich	38	51	1.75	-0.25	-0.44
Griechenland	37	50	0.75	-1.25	-0.94
Litauen	38	51	1.75	-0.25	-0.44
Ungarn	26	43	-10.25	-8.25	84.56

$$\begin{aligned} cov(x; y) &= \frac{1}{8} * (15,81 + 17,06 + 21,56 + 10,31 + (-0,44) + (-0,94) + (-0,44) + 84,56) \\ &= \frac{1}{8} * 147,48 = \mathbf{18,44} \end{aligned}$$

Berechnung der Standardabweichungen für die spätere Berechnung von Pearson's r:

$$\begin{aligned} s_x &= \sqrt{\frac{1}{8}((39 - 36,25)^2 + (33 - 36,25)^2 + (40 - 36,25)^2 + (39 - 36,25)^2 + (38 - 36,25)^2 + (37 - 36,25)^2 + (38 - 36,25)^2 + (26 - 36,25)^2)} \\ &= \sqrt{\frac{1}{8}(7,56 + 10,56 + 14,06 + 7,56 + 3,06 + 0,56 + 3,06 + 105,06)} = \sqrt{\frac{1}{8} * 151,48} = \mathbf{4,35} \end{aligned}$$

$$\begin{aligned} s_y &= \sqrt{\frac{1}{8}((57 - 51,25)^2 + (46 - 51,25)^2 + (57 - 51,25)^2 + (55 - 51,25)^2 + (51 - 51,25)^2 + (50 - 51,25)^2 + (51 - 51,25)^2 + (43 - 51,25)^2)} \\ &= \sqrt{\frac{1}{8}(33,06 + 27,56 + 33,06 + 14,06 + 0,06 + 1,56 + 0,06 + 68,06)} = \sqrt{\frac{1}{8} * 177,48} = \mathbf{4,71} \end{aligned}$$

$$\text{Persons } r = \frac{cov(x; y)}{s_x * s_y} = \frac{18,44}{4,35 * 4,71} = \mathbf{0,90}$$

Allgemein kann ein Zusammenhang zwischen dem Zugang zu den politischen Rechten in einem Staat und dem Grad an bürgerlichen Freiheiten festgestellt werden, da $cov(x; y) = 18,44 \neq 0$. Obwohl Pearson's r in dieser Stichprobe etwas geringer ist, kann man noch immer eine sehr starke Korrelation erkennen.

Die Samples unterscheiden sich in einigen Merkmalen. In der ersten Stichprobe liegen die Mittelwerte deutlich unter denen der zweiten Stichprobe, d.h. Freedom House misst in diesen EU-Staaten einen höheren Grad an politischen Rechten und bürgerlichen Freiheiten als in der global gemischten Stichprobe.

Außerdem betragen die Standardabweichungen in der EU-Stichprobe nur etwa ein Drittel bzw. ein Viertel der Standardabweichungen in der gemischten Stichprobe, d.h. die Staaten in der EU-Stichprobe sind sich deutlich ähnlicher als diejenigen in der gemischten Stichprobe. Hier wird noch

[illegible]

Eine Gruppierung empfiehlt sich also zum Zwecke besserer Übersichtlichkeit. Für die Positionierung auf der Links-Rechts-Achse können Sie wieder die Einteilung in Links (0-3), Mitte (4-6) und Rechts wählen (7-10). Für die Positionierung zu TTIP können Sie beispielsweise folgende Einteilung wählen:

Alte Ausprägung	Neue Ausprägung	
	Beschreibung:	Quantitative Zuordnung
1-2	Keine Zustimmung	0
3-5	Mäßige Zustimmung	1
6-7	Hohe Zustimmung	2

Wie immer wären auch andere Einteilungen möglich. Für die Gruppierung wählen Sie Transformieren > Umcodieren in andere Variable und füllen Sie das Fenster mit den jeweiligen Angaben aus (für mehr Details zu Gruppierungen: siehe Aufgaben Sitzung 2-3 und Aufgaben Sitzung 6).

Nutzen Sie alternativ folgende Syntax:

```
RECODE A3 (1 thru 2=0) (3 thru 5=1) (6 thru 7=2) INTO PostTTIPGrup.
VARIABLE LABELS PostTTIPGrup 'Gruppierte Position TTIP'.
EXECUTE.
RECODE A5 (0 thru 3=0) (4 thru 6=1) (7 thru 10=2) INTO LRGr.
VARIABLE LABELS LRGr 'Links Rechts in drei Gruppen'.
EXECUTE.
```

Wenn Sie beide Variablen als ordinal betrachten, dürfen Sie Somer's d, Goodman und Kruskals γ , Kendall's τ_B ausrechnen. Auch die Berechnung des nominalen Zusammenhangsmaßes Chi-Square samt seiner Normierung Cramers V ist zulässig.

Wählen Sie dazu Analysieren > Deskriptive Statistiken > Kreuztabellen. Ziehen sie die neue Links-Rechts-Variable in die Spalten, die Position zu TTIP in die Zeilen. Klicken Sie auf „Statistiken“ und wählen Sie alle diese Maße. Stellen Sie unter „Zellen“ sicher, dass die Spaltenprozentuierung ausgegeben wird. Lassen Sie sich dann die Kreuztabelle ausgeben.

Wählen Sie alternativ folgende Syntax:

```
CROSSTABS /TABLES=PostTTIPGrup BY LRGr /FORMAT=AVALUE TABLES
/STATISTICS=CHISQ CC PHI GAMMA D BTAU
/CELLS=COUNT COLUMN /COUNT ROUND CELL.
```

Sie erhalten:

Chi-Quadrat-Tests			
	Wert	df	Asymptotische Signifikanz (zweiseitig)
Pearson-Chi-Quadrat	60,562 ^a	4	<,001
Likelihood-Quotient	49,331	4	<,001
Zusammenhang linear-mit-linear	42,467	1	<,001
Anzahl der gültigen Fälle	165		

a. 5 Zellen (55,6%) haben eine erwartete Häufigkeit kleiner 5. Die minimale erwartete Häufigkeit ist ,34.

Richtungsmaße

			Wert	Asymptotischer Standardfehler ^a	Näherungsweise ^{t^b}	Näherungsweise Signifikanz
Ordinal- bzgl. Ordinalmaß	Somers-d	Symmetrisch	,496	,067	6,946	<,001
		Gruppierte Position TTIP abhängig	,503	,068	6,946	<,001
		Links Rechts in drei Gruppen abhängig	,488	,068	6,946	<,001

a. Die Null-Hyphothese wird nicht angenommen.

b. Unter Annahme der Null-Hyphothese wird der asymptotische Standardfehler verwendet.

Symmetrische Maße

		Wert	Asymptotischer Standardfehler ^a	Näherungsweise ^{t^b}	Näherungsweise Signifikanz
Nominal- bzgl. Nominalmaß	Phi	,606			<,001
	Cramer-V	,428			<,001
Ordinal- bzgl. Ordinalmaß	Kendall-Tau-b	,496	,067	6,946	<,001
	Gamma	,767	,069	6,946	<,001
Anzahl der gültigen Fälle		165			

a. Die Null-Hyphothese wird nicht angenommen.

b. Unter Annahme der Null-Hyphothese wird der asymptotische Standardfehler verwendet.

Chi-Square beträgt 60,52. Es kann also von einem Zusammenhang ausgegangen werden. Cramers V beträgt 0,43, d.h. es kann von einem mittleren Zusammenhang ausgegangen werden. Auch Kendall's τ_b zeigt mit 0,496 einen mittleren Zusammenhang an. Somer's d lässt zwar auf einen starken Zusammenhang schließen, der Wert von 0,5 liegt jedoch nur geringfügig über den anderen. Lediglich Goodman und Kruskal's γ überschätzt den Zusammenhang stark (0,77).