



Mehrfachantworten

Tabellierung von Fragen mit Mehrfachantworten

Kurt Holm

Almo Statistik-System

www.almo-statistik.de

holm@almo-statistik.de

kurt.holm@jku.at

Siehe auch die beiden Almo-Dokumente Nr. 1a "Eindimensionale Tabellierung" und Nr. 1b "Zwei- und drei-dimensionale Tabellierung".

Im Text wird häufig auf das Dokument **P0** Bezug genommen. Dabei handelt es sich um das Almo-Dokument "Arbeiten mit Almo.PDF" (Dokument 0).

Weitere Almo-Dokumente

Die folgenden Dokumente können alle kostenlos von der Handbuchseite in <http://www.almo-statistik.de/> heruntergeladen werden

0. Arbeiten_mit_Almo.PDF (1 MB)
- 1a. Eindimensionale Tabellierung
- 1b. Zwei- und drei-dimensionale Tabellierung.PDF (1.1 MB)
2. Beliebig-dimensionale Tabellierung.PDF (1.7 MB)
3. Nicht-parametrische Verfahren.PDF (0.9 MB)
4. Kanonische Analysen.PDF (1.8 MB)
Diskriminanzanalyse.PDF (1.8 MB)
enthält: Kanonische Korrelation, Diskriminanzanalyse, bivariate Korrespondenzanalyse, optimale Skalierung
5. Korrelation.PDF (1.4 MB)
6. Allgemeine multiple Korrespondenzanalyse.PDF (1.5 MB)
7. Allgemeines ordinales Rasch-Modell.PDF (0.6 MB)
- 7a. Wie man mit Almo ein Rasch-Modell rechnet.PDF (0.2 MB)
8. Tests auf Mittelwertsdifferenz, t-Test.PDF (1,6 MB)
9. Logitanalyse.pdf (1,2MB) enthält Logit- und Probitanalyse
10. Koeffizienten der Logitanalyse.PDF (0,06 MB)
11. Daten-Fusion.PDF (1,1 MB)
12. Daten-Imputation.PDF (1,3 MB)
13. ALM Allgemeines Lineares Modell.PDF (2.3 MB)
- 13a. ALM Allgemeines Lineares Modell II.PDF (2.7 MB)
14. Ereignisanalyse: Sterbetafel-Methode, Kaplan-Meier-Schätzer, Cox-Regression.PDF (1,5 MB)
15. Faktorenanalyse.PDF (1,6 MB)
16. Konfirmatorische Faktorenanalyse.PDF (0,3 MB)
17. Clusteranalyse.PDF (3 MB)
18. Pisa 2012 Almo-Daten und Analyse-Programme.PDF (17 KB)
19. Guttman- und Mokken-Skalierung.PFD (0.8 MB)
20. Latent Structure Analysis.PDF (1 MB)
21. Statistische Algorithmen in C (80 KB)
22. Conjoint-Analyse (PDF 0,8 MB)
23. Ausreisser entdecken (PDF 170 KB)
24. Statistische Datenanalyse Teil I, Data Mining I
25. Statistische Datenanalyse Teil II, Data Mining II
26. Statistische Datenanalyse Teil III, Arbeiten mit Almo-Datenanalyse-System
27. Mehrfachantworten, Tabellierung von Fragen mit Mehrfachantworten
28. Metrische multidimensionale Skalierung (MDS) (0,4 MB)
29. Metrisches multidimensionales Unfolding (MDU) (0,6 MB)
30. Nicht-metrische multidimensionale Skalierung (MDS) (0,4 MB)
31. Pfadanalyse als wiederholte Regressionsanalyse (0,7 MB)
32. Datei-Operationen mit Almo (1,1 MB)
33. Wählerstromanalyse und Wahlhochrechnung (1,6 MB)
34. Soziometrie. Auswertung soziometrischer Daten (0,5 MB)

INHALTSVERZEICHNIS

P4.11 Auszählung und Tabellierung von Variablen mit Mehrfach-Antworten	4
P4.11.1 Dichotome Methode: Auszählung von "Mehrfach-Fragen" mit Prog00mm	4
4.11.1.1 Erläuterungen zu den Eingabeboxen von Prog00mm	8
P4.11.1.2 Ergebnisse aus Prog00mm	11
P4.11.2 Dichotome Methode: Kreuztabellierung von Mehrfach-Fragen mit Prog00mn .	13
4.11.2.1 Erläuterungen zu den Eingabeboxen von Prog00mn	14
P4.11.2.2 Ergebnisse aus Prog00mn	15
P4.11.3 Kategoriale Methode: Auszählung von Mehrfach-Fragen mit Prog00mi.....	18
4.11.3.1 Erläuterungen zu den Eingabeboxen von Prog00mi	24
P4.11.3.2 Ergebnisse aus Prog00mi	31
P4.11.4 Kategoriale Methode: Kreuztabellierung von Mehrfach-Fragen mit Prog00mo .	33
P4.11.4.1 Ergebnisse aus Prog00mo	34

P4.11 Auszählung und Tabellierung von Variablen mit Mehrfach-Antworten

P4.11.1 Dichotome Methode: Auszählung von "Mehrfach-Fragen" mit Prog00mm

Betrachten wir ein Beispiel.

In einer Umfrage wurde die Frage gestellt:

Was trinken Sie, wenn Sie Durst haben ?

Die Befragten können dabei mehrere verschiedene Getränke angeben. Ein Befragter antwortet z.B.: Limonade, Apfelsaft, Bier

Insgesamt, so wollen wir annehmen, wurden von allen Befragten 5 verschiedene Getränke genannt. Beim Schreiben der Daten werden deswegen 5 dichotome Variable verwendet, in die die Getränke eingeschrieben werden.

Allgemein formuliert:

1. Eine Frage besitzt x nicht-alternative Antwortmöglichkeiten.
In unserem Beispiel: 5
2. Beim Schreiben der Daten werden deswegen x dichotome Variable verwendet, d.h. in den Datensatz eingesetzt

Die Datenmatrix für das Beispiel hat folgendes Aussehen:

Lebens- alter V1	Geschl. V2	Limonade V3	Apfelsaft V4	Bier V5	Milch V6	Wasser V7
27	1	1	0	0	1	0
43	2	0	0	1	0	1
61	2	1	1	0	1	0
81	1	1	1	0	0	0
46	1	0	0	1	1	1
93	1	-1	-1	-1	-1	-1
19	2	1	1	0	1	0
33	2	0	0	0	0	1

Diese Datei ist unter dem Namen "Getraenk.fre" im Verzeichnis Testdat enthalten

Die 1. Person ist 27 Jahre alt, männlich und trinkt Limonade und Milch
Die 2. Person ist 43 Jahre alt, weiblich und trinkt Bier und Wasser
Die 6. Person ist 93 Jahre alt, männlich und hat die Antwort verweigert

Beachte: Die dichotomen Variablen für die Getränke können im Datensatz an beliebiger Stelle stehen. Vor, hinter und zwischen diesen Variablen können beliebige andere Variable stehen. In unserem Beispiel steht z.B. vor diesen Variablen das Lebensalter und das Geschlecht

Wir verwenden folgende Codes für die dichotomen "Getränke"-Variablen:

```
keine Angabe      = 0
genannt           = 1
Antwort verweigert = -1
```

Eine Auszählung der 5 dichotomen "Getränke"-Variablen mit Prog04m1 oder Prog04m2 ist unproblematisch. Es entstehen 5 Tabellen folgender Art (z.B. für "Limonade"). Die Codeziffer -1 haben wir als Kein-Wert-Code deklariert.

Variable 3 Limonade

Wert	Faelle	%	% ohne KW	% kumuliert
Kein Wert	1	12.50	-	-
0 nicht genannt	3	37.50	42.86	42.86
1 genannt	4	50.00	57.14	100.00
Summe	8	100%		
Summe ohne KW	7		100%	

Wir wünschen jedoch eine gemeinsame Tabelle für alle 5 Getränke, um auf einen Blick feststellen zu können, welches Getränk das beliebteste ist. Also soll also eine Häufigkeitsverteilung für alle Antwortmöglichkeiten erstellen. Es soll folgende gemeinsame Tabelle erzeugt werden:

		Häufigkeit		Summe
		nicht genannt	genannt	
Antwort	Limonade	4	4	8
	Apfelsaft	5	3	8
	Bier	6	2	8
	Milch	4	4	8
	Wasser	5	3	8
Summe		24	16	40

Dies wird durch das nachfolgende Programm **Prog00mm** ermöglicht

Prog00mm.Msk
 Auszählen einer Frage mit Mehrfach-Antworten 1
 dichotomer Fall

Das Programm erstellt eine Häufigkeitsverteilung für eine Frage mit mehreren Antwortmöglichkeiten

Beispiel:

In einer Umfrage wurde die Frage gestellt:

Was trinken Sie, wenn Sie Durst haben ?

Die Befragten können dabei mehrere verschiedene Getränke angeben. Ein Befragter antwortet z.B.: Limonade, Apfelsaft, Bier

Insgesamt wurden von allen Befragten 5 verschiedene Getränke genannt. Beim Schreiben der Daten werden deswegen 5 dichotome Variable verwendet, in die die Getränke eingeschrieben werden.

Allgemein formuliert:

1. Eine Frage besitzt x nicht-alternative Antwortmöglichkeiten.
 In unserem Beispiel: 5
2. Beim Schreiben der Daten werden deswegen x dichotome Variable verwendet, d.h. in den Datensatz eingesetzt
3. Also erstellt eine Häufigkeitsverteilung für die gesamten x Antwortmöglichkeiten

Die Datenmatrix für das Beispiel hat folgendes Aussehen:

Lebens- alter U1	Geschl. U2	Limonade U3	Apfelsaft U4	Bier U5	Milch U6	Wasser U7
27	1	1	0	0	1	0
43	2	0	0	1	0	1
61	2	1	1	0	1	0
81	1	1	1	0	0	0
46	1	0	0	1	1	1
93	1	-1	-1	-1	-1	-1
19	2	1	1	0	1	0
33	2	0	0	0	0	1

Die 1. Person ist 27 Jahre alt, männlich und trinkt Limonade und Milch

Die 2. Person ist 43 Jahre alt, weiblich und trinkt Bier und Wasser

Die 6. Person ist 93 Jahre alt, männlich und hat die Antwort verweigert

Beachte: Die dichotomen Variablen für die Getränke können im Datensatz an beliebiger Stelle stehen. Vor, hinter und zwischen diesen Variablen können beliebige andere Variable stehen. In unserem Beispiel steht z.B. vor diesen Variablen das Lebensalter und das Geschlecht

Wir verwenden folgende Codes:

keine Angabe = 0
 genannt = 1
 Antwort verweigert = -1

Der Benutzer kann entscheiden, ob eine Verweigerung wie "keine Angabe" gezählt werden soll oder ob eine Person, die verweigert, aus der Analyse ausgeschlossen werden soll

Was ist ein Kurzprogramm ? --> Hilfe
 Bedienung --> Hilfe

Speicher fuer x Variable
Vereinbaren Sie mindestens 2 Variable mehr als notwendig
Vereinbare Variable= ;

Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

Datei der Variablennamen
 "C:\Almo7\TESTDAT\Getraenk.nam"
 zeige zeige = Namensdatei in Output zeigen
leer = nicht

Freie Namensfelder

 erzeuge zusätzliche Namensfelder

Datei aus der gelesen wird
bei Datei-Problemen
 "C:\Almo7\TESTDAT\Getraenk.fre"
 frei Format der Daten
 U1:? der Datensatz enthält diese Variablen
Bei Format DIREKT schreiben Sie: alle_U

Wenn Dateiformat FIX oder Nicht-Standard-FREI

Option: Ein- und Ausschliessen von Untersuchungseinheiten

Option: Umkodierungen und Kein-Wert-Angaben

die im Datensatz enthaltenen "dichotomen Variablen"
 Limonade,Apfelsaft,Bier,Milch,Wasser
 0 Kodeziffer für "keine Angabe"
 -1 Kodeziffer für "Antwort verweigert"
bzw. für "Kein-Wert"
 0 0= "Antwort verweigert" bzw. "Kein-Wert"
wie "keine Angabe" behandeln
1= Datensatz ausschließen,
wenn "Antwort verweigert" bzw. "Kein-Wert"



4.11.1.1 Erläuterungen zu den Eingabeboxen von Prog00mm

Box "Speicher für x Variable"

Siehe Almo-Dokument Nr.0 "Arbeiten mit Almo", Abschnitt P0.1

Box "Option: Weitere Vereinbarungen"

Siehe Almo-Dokument Nr.0 "Arbeiten mit Almo", Abschnitt P0.2

Box "Datei der Variablennamen"

Box "Freie Namensfelder"

Siehe P0.3

Box "Datei aus der gelesen wird"

Box " Wenn Dateiformat FIX oder Nicht-Standard-FREI"

Siehe P0.4

Box "Option: Ein- und Ausschliessen von Untersuchungseinheiten"

Siehe P0.7.1

Box "Option: Umkodierungen und Kein-Wert-Angaben"

Siehe P0.5

Box "die im Datensatz enthaltenen "dichotomen Variablen"



1. Eingabefeld: Geben Sie hier die dichotomen Variablen an, in die die verschiedenen

Antwortmöglichkeiten eingeschrieben worden sind.

2. *Eingabefeld*: Geben Sie hier die Kodeziffer an, die für "keine Angabe" geschrieben wurde. Üblicherweise wird die 0 als Codeziffer verwendet, wenn der Befragte auf eine dichotome Frage keine Nennung abgibt. In unserem Beispiel trinkt der 1. Befragte keinen Apfelsaft, kein Bier und kein Wassern. Für diese dichotomen Variablen wird also eine 0 für ihn kodiert.

Alle anderen Zahlenwerte

ausser der Zahl, die in das Eingabefeld für
"Antwort verweigert bzw. Kein_Wert" eingetragen wurde

werden Almo-intern auf 1 umkodiert und als Nennung interpretiert.

In das Eingabefeld für "keine Angabe" können mehrere durch Beistrich getrennte Zahlen eingetragen werden. Möglich ist auch ein durch Doppelpunkt gekennzeichnetes Werte-Intervall. Beispiel: Wenn "keine Angabe" mit 0 und 91 bis 99 kodiert wurde, dann wird in das Eingabefeld geschrieben:

0, 91:99

Beachte: Es wird unterstellt, dass der Code für "keine Angabe" für alle dichotomen Fragen derselbe ist. Ist das nicht der Fall - was sehr ungewöhnlich ist - dann müssen die Codeziffern für "keine Angabe" für die verschiedenen dichotomen Fragen so umkodiert werden, dass sie alle gleich sind. Das geschieht in der Optionsbox "Umkodierungen und Kein-Wert-Angaben".

Beispiel: Bei der Variablen "trinkt Apfelsaft" ist, abweichend von den anderen dichotomen Variablen, "keine Angabe" mit -9 kodiert. Bei den anderen dichotomen Variablen wurde "keine Angabe" mit 0 kodiert. In diesem Fall schreiben Sie in die Box "Option: Umkodierungen und Kein-Wert-Angaben"

Apfelsaft (-9 = 0)

Damit erhält diese Variable denselben Code für "keine Angabe" wie die anderen Variablen

3. *Eingabefeld*: Geben Sie hier die Kodeziffer an, die für "Antwort verweigert" bzw. für "Kein-Wert" geschrieben wurde. Es können auch mehrere durch Beistrich getrennte Codes angegeben werden. Möglich ist auch ein durch Doppelpunkt gekennzeichnetes Werte-Intervall. Beispiel:

-1, 91:99

Ist in der auszuzählenden Datei bereits der Almo-interne Kein-Wert-Code enthalten, dann kann das Eingabefeld leer bleiben

Beachte: Es wird unterstellt, dass der Kein-Wert-Code für alle Mehrfach-Fragen derselbe ist. Ist das nicht der Fall - was sehr ungewöhnlich ist - dann müssen die Codeziffern für "Kein_Wert" für die verschiedenen Mehrfach-Fragen so umkodiert werden, dass sie alle gleich sind. Das geschieht in der Optionsbox "Umkodierungen und Kein-Wert-Angaben".

Beispiel: Bei der Variablen "trinkt Apfelsaft" ist, abweichend von den anderen dichotomen

Variablen, "Antwort verweigert" bzw. KeinWert mit 0 kodiert. Bei den anderen dichotomen Variablen wurde -1 kodiert. In diesem Fall schreiben Sie in die Box "Option: Umkodierungen und Kein-Wert-Angaben"

Apfelsaft (0 = -1)

Damit erhält diese Variable denselben Code für KeinWert wie die anderen Variablen

4. *Eingabefeld*: Der Benutzer kann entscheiden, ob eine Antwortverweigerung wie "keine Angabe" gezählt werden soll oder ob eine Person, die verweigert, aus der Analyse ausgeschlossen werden soll. Schreiben Sie

- 0 wenn "Antwort verweigert" bzw. "Kein-Wert" wie "keine Angabe" behandelt werden soll
- 1 wenn der Datensatz ausgeschlossen werden soll, wenn "Antwort verweigert" bzw. "Kein-Wert" vorliegt

Box "Namen für die Antwortmöglichkeiten"



Den verschiedenen Wahlmöglichkeiten können Namen gegeben werden, die Almo in der Ergebnis-Tabelle verwendet. Diese Namen können (müssen aber nicht) anders sein als die Variablenamen der dichotomen Variablen.

Werden keine Namen vergeben, dann schreibt Almo in der Ergebnis-Tabelle nur die fortlaufenden Nummern 1, 2, 3,

Wenn Sie jedoch Namen vergeben, dann ist folgendes zu beachten.

1. Die Reihenfolge der Namen für die Wahlmöglichkeiten muss mit der Reihenfolge der Variablen in der Box "die im Datensatz enthaltenen dichotomen Variablen" übereinstimmen.
2. Die Namen werden untereinander je in ein Eingabefeld geschrieben
3. Die Namen können (müssen aber nicht) mit Codeziffern in Klammern geschrieben werden. Es gibt also 2 Möglichkeiten

a. Ohne Codeziffern

Limonade
Apfelsaft
Bier
Milch
Wasser

b. Mit Codeziffern

Die Codeziffern müssen mit 1 beginnen und fortlaufend sein

- (1) Limonade
- (2) Apfelsaft
- (3) Bier
- (4) Milch
- (5) Wasser

4. Werden die Namen mit Codeziffern geschrieben, dann dürfe mitten drin Namen fehlen. Beispiel:

- (1) Limonade
- (2) Apfelsaft
- (5) Wasser

Die Namen (3) Bier und (4) Milch fehlen. Also kann in der Ergebnis-Tabelle die Namen richtig zuordnen, da die Codeziffern mit den Namen verknüpft sind

5. Insgesamt können 88 Namen vergeben werden. Sind mehr Wahlmöglichkeiten vorhanden, z.B. 100, dann kann den letzten 12 kein Name gegeben werden. Also verwendet dann für diese die Ausprägungsnummern 89, 90, 91,99,100

Box "Option: "Aussehen" der auszugebenden Tabelle bzw. Matrix"
Siehe P0.9

Box "Grafik-Optionen"
Siehe P0.10

P4.11.1.2 Ergebnisse aus Prog00mm

Zahl der eingelesenen Datensätze: 40

***** MITTEILUNG
Bei Auszahlung von Variablen mit Mehrfach-Antworten entsteht die oben angegebene "Zahl der eingelesenen Datensätze" aus der Zahl der tatsächlich eingelesenen Datensätze mal Zahl der Auswahlmöglichkeiten

die Zahl der tatsächlich eingelesenen Datensätze ist: 8

***** MITTEILUNG
Bedingt durch KEIN_WERT-Fälle können für die Tabellen unterschiedliche Summen entstehen und auch Summen, die von der oben angegebenen Zahl der Datensätze abweichen

***** MITTEILUNG
Also fasst die mehrfachen Antwortmöglichkeiten zusammen in V20 und verwendet zur Bildung der Tabelle die beiden Variablen V19 und V20

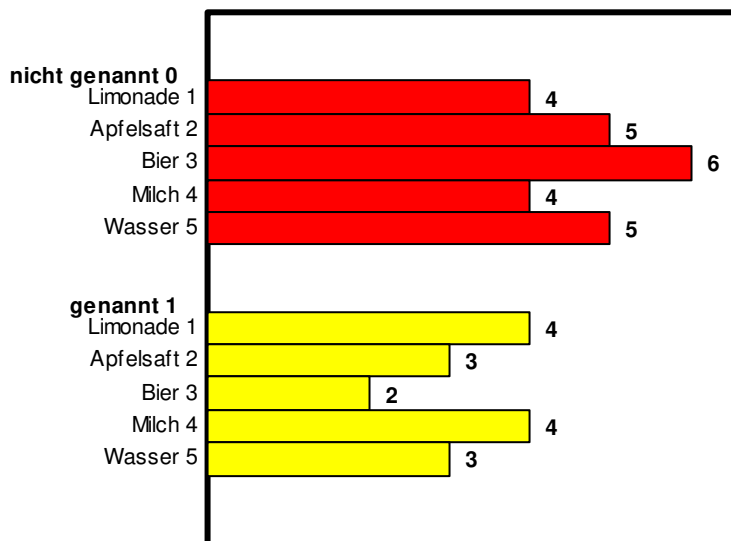
		Häufigkeit		Summe
		nicht genannt	genannt	
Antwort	Limonade	4	4	8
	Apfelsaft	5	3	8
	Bier	6	2	8
	Milch	4	4	8
	Wasser	5	3	8
Summe		24	16	40

******* Erläuterung**

In der Spaltensumme steht jeweils 8. Das ist die Zahl der Befragten. Von diesen 8 Befragten haben 4 Limonade genannt und 4 nicht, 3 haben Apfelsaft genannt und 5 nicht etc. Die 5 Getränke wurden von den 8 Befragten insgesamt 16 mal genannt und 24 mal nicht genannt. Das ist der unteren Spaltensumme zu entnehmen.

Almo zeichnet für diese Tabelle dann noch ein Säulendiagramm.

2-dimensionale Verteilung
Antwort
und
Häufigkeit



Zeilenweise prozentuiert

		Häufigkeit		Summe
		nicht genannt	genannt	
Antwort	Limonade	50.0000	50.0000	100.0000
	Apfelsaft	62.5000	37.5000	100.0000
	Bier	75.0000	25.0000	100.0000
	Milch	50.0000	50.0000	100.0000
	Wasser	62.5000	37.5000	100.0000
Summe		60.0000	40.0000	100.0000

******* Erläuterung**

Von den 8 Befragten haben 50 % Limonade genannt und 50 % nicht, 37.5 % von den 8 Befragten haben Apfelsaft genannt und 62.5 % nicht etc. Die geringste Häufigkeit an Nennungen hat Bier mit 25 %

Almo zeichnet für diese Tabelle dann noch ein Säulendiagramm.

Spaltenweise prozentuiert

		Häufigkeit		Summe
		nicht genannt	genannt	
Antwort	Limonade	16.6667	25.0000	20.0000
	Apfelsaft	20.8333	18.7500	20.0000
	Bier	25.0000	12.5000	20.0000
	Milch	16.6667	25.0000	20.0000
	Wasser	20.8333	18.7500	20.0000
Summe		100.0000	100.0000	100.0000

******* Erläuterung**

Die 5 Getränke wurden 16 mal genannt (siehe 1. Tabelle). Von diesen 16 Nennungen entfielen 25 % auf Limonade, 18.75 % auf Apfelsaft etc. Die geringste Häufigkeit an Nennungen hat Bier mit 12.5 %

Almo zeichnet für diese Tabelle dann noch ein Säulendiagramm.

P4.11.2 Dichotome Methode: Kreuztabellierung von Mehrfach-Fragen mit Prog00mn

Die 5 dichotomen "Getränke"-Variablen sollen gegen das Geschlecht der Befragten tabelliert werden. Dies ist problemlos möglich mit den Tabellierungsprogrammen in Almo, z.B. mit Prog10m1, Prog10m2 oder Prog11m oder Prog11m1.

Dabei entstehen 5 Tabellen folgender Art (z.B. für "Limonade" mit Geschlecht). Die Codeziffer -1 haben wir als Kein-Wert-Code deklariert.

	Limonade		Summe
	nicht	genannt	
Geschl männlich	1	2	3
weiblich	2	2	4
Summe	3	4	7

Wir wünschen jedoch eine gemeinsame Tabelle, in der alle 5 genannten Getränke gegen die Variable des Geschlecht tabelliert sind. Almo soll also eine Tabelle für alle Antwortmöglichkeiten gegenüber dem Geschlecht erstellen. Es soll folgende gemeinsame Tabelle erzeugt werden:

Genanntes Getränk mit Geschlecht

		Geschlecht		Summe
		männlich	weiblich	
genannt	Limonade	2	2	4
	Apfelsaft	1	2	3
	Bier	1	1	2
	Milch	2	2	4
	Wasser	1	2	3
Summe		7	9	16

Diese Tabelle zeigt uns, wie sich die genannten Getränke auf die beiden Geschlechter verteilen.

Eine 2. Tabelle soll dann zeigen, wie sich die Nicht-Nennungen auf die beiden Geschlechter verteilen.

Dies wird durch das Programm Prog00mn ermöglicht

4.11.2.1 Erläuterungen zu den Eingabeboxen von Prog00mn

Das Programm ist identisch mit dem oben dargestellten Prog00mm. Es besitzt folgende zwei zusätzliche Boxen:

Box "Die Eine-Antwort-Variablen"



Geben Sie hier die Variable an, gegen die die verschiedenen Antwortmöglichkeiten der Mehrfach-Variablen tabelliert werden sollen.

Box "Option: Koeffizienten berechnen"



Wird die Optionsbox geöffnet, dann sieht man folgende Box



Almo bildet zunächst, wie bei den Tabellierungsprogrammen Prog10m1 oder Prog10m2, 2-dimensionale Tabellen zwischen den einzelnen Antwortmöglichkeiten und der einen Einfach-Variablen. Im Beispiel:

"Trinkt Limonade"	mit	"Geschlecht"
"Trinkt Apfelsaft"	mit	"Geschlecht"
"Trinkt Bier"	mit	"Geschlecht"
"Trinkt Milch"	mit	"Geschlecht"
"Trinkt Wasser"	mit	"Geschlecht"

Für diese 2-dimensionalen Tabellen werden dann Chi-Quadrat und verschiedene Korrelationskoeffizienten berechnet.

In der Optionsbox wird dem Benutzer angeboten, auf verschieden Teile der Ausgabe zu verzichten. Wird auf alles verzichtet, dann wird lediglich ein Chi-Quadrat-Test gerechnet. In der Regel genügt dies auch.

P4.11.2.2 Ergebnisse aus Prog00mn

Tabelle

1

***** MITTEILUNG

Almo fasst die mehrfachen Antwortmoeglichkeiten zusammen in V20
und verwendet zur Bildung der Tabelle die beiden Variablen
V2 und V20

***** Erläuterung

Almo teilt mit, dass es die 5 Antwortmöglichkeiten in einer eigens generierten Variablen V20 zusammengefasst hat. V2 ist die "Einfach-Variable" des Geschlechts.

Variable 20 genannt
mit
Variable 2 Geschlecht

			Geschlecht		Summe
			männlich 1	weiblich 2	
genannt	Limonade	1	2	2	4
	Apfelsaft	2	1	2	3
	Bier	3	1	1	2
	Milch	4	2	2	4
	Wasser	5	1	2	3
Summe			7	9	16

***** Erläuterung

Almo gibt zunächst die Tabellen für die Nennungen und später für die Nicht-Nennungen aus. Die Männer geben insgesamt 7 Nennungen ab, die Frauen 9. Dies ist aus der unteren Summen-Zeile ablesbar. Aus der hinteren Summen-Spalte ergibt sich, dass Limonade 4 mal genannt wurde, Apfelsaft 3 mal etc.

Almo zeichnet für diese Tabelle dann noch ein Säulendiagramm.

Zeilenweise prozentuiert

			Geschlecht		Summe
			männlich 1	weiblich 2	
genannt	Limonade	1	50.0000	50.0000	100.0000
	Apfelsaft	2	33.3333	66.6667	100.0000
	Bier	3	50.0000	50.0000	100.0000
	Milch	4	50.0000	50.0000	100.0000
	Wasser	5	33.3333	66.6667	100.0000
Summe			43.7500	56.2500	100.0000

***** Erläuterung

Hier wird das jeweilige Getränk auf 100 % gesetzt. So wird beispielweise Apfelsaft zu 33.33 % von Männern, aber zu 66.67 % von Frauen genannt.

Spaltenweise prozentuiert

			Geschlecht		Summe
			männlich 1	weiblich 2	
genannt	Limonade	1	28.5714	22.2222	25.0000
	Apfelsaft	2	14.2857	22.2222	18.7500
	Bier	3	14.2857	11.1111	12.5000
	Milch	4	28.5714	22.2222	25.0000
	Wasser	5	14.2857	22.2222	18.7500
Summe			100.0000	100.0000	100.0000

***** **Erläuterung**

Die Nennungen, die die Männer und die Frauen abgegeben haben, werden hier jeweils auf 100 % gesetzt. Mann erkennt, dass 28.5714 % der Nennungen von Männern für Limonade und Milch abgegeben wurden etc.

Koeffizienten fuer Tabelle 1

Signifikanztest

Chi-Quadrat = 0.4233 df = 4 Signifikanz p = 0.981
Signifikanz (1-p)*100 = 1.948%

***** **Erläuterung**

Almo rechnet noch einen Chi-Quadrat-Test, der lediglich pauschal aussagt, ob in der Tabelle ein Zusammenhang zwischen Getränken und Geschlecht besteht oder nicht. Ob beispielsweise Männer signifikant eher "Bier" angeben als Frauen ist erst den nachfolgend ausgegebenen Tabellen zu entnehmen.

Almo tabelliert dann in gleicher Weise die Nicht-Nennungen mit dem Geschlecht. Üblicherweise sind diese Tabellen weniger interessant. Wir bilden sie deswegen hier auch nicht ab.

Wurde in Prog00mn die Optionsbox "Koeffizienten berechnen" geöffnet, dann gibt Almo 5 zwei-dimensionale Tabellen aus, die jeweils den Zusammenhang zwischen einem Getränk und dem Geschlecht darstellen. Wir betrachten hier die 2. Tabelle "Apfelsaft" mit Geschlecht.

Tabelle 2

Variable 4 Apfelsaft
mit
Variable 2 Geschlecht

		Geschlecht		Summe	
		männlich 1	weiblich 2		
Apfelsaft	nicht genannt	0	3	2	5
	genannt	1	1	2	3
Summe			4	4	8

***** Erläuterung

Die Tabelle zeigt, wie Apfelsaft von den befragten Männern und Frauen genannt bzw. nicht genannt wurde. Die nachfolgenden Koeffizienten drücken dann aus, ob zwischen dem Geschlecht und dem Trinken von Apfelsaft ein Zusammenhang besteht. Diese Koeffizienten wurden ausführlich in Handbuch, Teil 3, Abschnitt P10.2 und P10.4. beschrieben. In unserem Beispiel sind die Häufigkeiten sehr gering - so dass ein Chi-Quadrat-Test für dieses Beispiel eigentlich nicht zulässig ist.

Koeffizienten fuer Tabelle 2

Signifikanztest

Chi-Quadrat = 0.5333 df = 1 Signifikanz p = 0.465
Signifikanz (1-p)*100 = 53.477%

Ein Wert ueber ca. 95% bedeutet:
Zwischen den Variablen besteht
ein signifikanter Zusammenhang

Erwartungswerte kleiner 1 = 0%
kleiner 5 =100% der Zellen

Korrelationskoeffizienten

Messniveau		Korrelationskoeffizient		zweiseitige Signifikanz	
der einen Variablen	der anderen Variablen			p	(1-p)*100
dichotom	dichotom	Phi	= 0.2582	0.46523	53.4774
dichotom	polytom	Phi'	= 0.2582	0.46523	53.4774
polytom	polytom	Kontingenzk.C(cor)	= 0.3536	0.46523	53.4774
polytom	polytom	Tschuprow's T	= 0.2582	0.46523	53.4774
polytom	polytom	Cramer's V	= 0.2582	0.46523	53.4774
polytom	polytom	Lambda (asymm.)	= 0.2500	0.66501	33.4985
ordinal	ordinal	Gamma	= 0.5000	0.38275	61.7246
ordinal	ordinal	tau-b	= 0.2582	0.51649	48.3510
dichotom	ordinal	biseriales tau-b	= 0.2582	0.51649	48.3510
ordinal	ordinal	Rho	= 0.2582	0.53696	46.3037
dichotom	quantit.	punktbiseriales r	= 0.2582	0.53696	46.3037
quantit.	quantit.	Produkt-Moment r	= 0.2582	0.53696	46.3037

Zeilen- variable	Spalten- variable				
quantit.	ordinal	Gross-Gamma I	= 0.2582	0.49460	50.5396
ordinal	quantit.	Gross-Gamma II	= 0.2582	0.49460	50.5396
polytom	quantit.	Eta	= 0.2582	0.53696	46.3037
quantit.	quantit.	nichtlineares Eta	= 0.2582	0.53696	46.3037
		Signifik. der Nichtlinearitaet		1.00000	0.0000
dichotom	dichotom	tetrachorisches r	= 0.4086	0.47018	52.9819

P4.11.3 Kategoriale Methode: Auszählung von Mehrfach-Fragen mit Prog00mi

Betrachten wir ein Beispiel.

In einer Umfrage wurde die Frage gestellt:

Welche Zeitungen lesen Sie?

Insgesamt, so wollen wir annehmen, wurden von allen Befragten 10 verschiedene Zeitungen genannt. Kein Befragter gab aber mehr als 3 Zeitungen an, die er liest. Beim Kodieren und Schreiben der Daten wurden deswegen, um Schreibarbeit zu sparen, nur 3 Variable verwendet, in die die Zeitungen eingeschrieben wurden.

Allgemein formuliert:

1. Eine Frage besitzt x nicht-alternative Antwortmöglichkeiten.
In unserem Beispiel: 10
2. Kein Befragter nutzt jedoch mehr als y Antwortmöglichkeiten.
In unserem Beispiel: 3. Im Grenzfall kann y gleich x sein
Demzufolge werden beim Schreiben der Daten nur y Variable für die Frage mit x Antwortmöglichkeiten angelegt.

Die Datenmatrix für das Beispiel hat folgendes Aussehen:

Lebensalter	Geschlecht	genannte Zeitung 1	genannte Zeitung 2	genannte Zeitung 3
V1	V2	V3	V4	V5
27	männl	Spiegel	Profil	Stern
43	weibl	FAZ	Spiegel	-
61	männl	Krone	Profil	Spiegel
81	weibl	-	-	-
46	männl	verweigert	verweigert	verweigert
:	:	:	:	:

Die 1. Person ist 27 Jahre alt, männlich und liest den Spiegel, das Profil und den Stern.

Die 4. Person ist 81 Jahre alt, weiblich und liest überhaupt keine Zeitung

Die 5. Person ist 46 Jahre alt, männlich und hat die Antwort verweigert

Beachte: Die Variablen für die genannten Zeitungen 1 bis 3 können im Datensatz an beliebiger Stelle stehen. Vor, hinter und zwischen diesen Variablen können beliebige andere Variable stehen. In unserem Beispiel steht z.B. vor diesen Variablen das Lebensalter und das Geschlecht des Befragten.

Wir verwenden folgende Codes:

```
keine Angabe = 0
FAZ           = 1
Stern        = 2
Spiegel      = 3
Profil       = 4
....
....
Antwort verweigert= -1
```

Der Benutzer kann entscheiden, ob eine Verweigerung wie "keine Angabe" gezählt werden soll oder ob eine Person, die verweigert, aus der Analyse ausgeschlossen werden soll

Die Datenmatrix ist also folgende:

	V1	V2	V3	V4	V5
	--	--	--	--	--
	27	1	3	4	2
	43	2	1	3	0
	61	1	5	4	3
	81	2	0	0	0
	46	1	-1	-1	-1

Diese Datei ist unter dem Namen "Zeitung2.fre" im Verzeichnis Testdat enthalten

Almo soll für diese Daten eine Häufigkeitsverteilung für alle Antwortmöglichkeiten erstellen. Es soll folgende gemeinsame Tabelle erzeugt werden:

		Häufigkeit		Summe
		nicht genannt	genannt	
Antwort	FAZ	10	2	12
	Stern	11	1	12
	Spiegel	5	7	12
	Profil	7	5	12
	Krone	10	2	12
	Bild	10	2	12
	Bunte	9	3	12
	Frau	10	2	12
	Focus	11	1	12
	Kurier	10	2	12
Summe		93	27	120

Dies wird durch das nachfolgende Programm Prog00mi ermöglicht

Almo generiert automatisch x neue Variable und fügt diese am Ende des Datensatzes an. In unserem Beispiel sind das 10 neue Variable - für jede Zeitung eine. Diese Variable sind dichotom. Sie können die Werte 1 (=genannt) und 0 (=nicht genannt) annehmen. In unserem Beispiel erhalten diese 10 dichotomen Variablen die Nummern V6 bis V15

Als Option wird im Programm Prog00mi angeboten, die neue Datei zu speichern. In unserem Beispiel enthält die neue Datei folgende Variable:

alte Variable					neue dichotome Variable									
V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15
--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
27	1	3	4	2	0	1	1	1	0	0	0	0	0	0
43	2	1	3	0	1	0	1	0	0	0	0	0	0	0
61	1	5	4	3	0	0	1	1	1	0	0	0	0	0
81	2	0	0	0	0	0	0	0	0	0	0	0	0	0
46	1	-1	-1	-1	0	0	0	0	0	0	0	0	0	0
18	1	5	0	0	0	0	0	0	1	0	0	0	0	0
34	2	1	3	4	1	0	1	1	0	0	0	0	0	0
.
.

Terminologie: Wir nennen
 die Variablen V2,V3,V4 aus der Datenmatrix: "Mehrfach-Variable"
 und die 10 neuen Variablen V6 bis V15: "dichotome Variable"

Prog00mi.Msk
Auszählen einer Frage mit Mehrfach-Antworten
kategorialer Fall

Das Programm erstellt eine Häufigkeitsverteilung für eine Frage mit mehreren Antwortmöglichkeiten

Außerdem bietet es die Option an, aus der Mehrfachantworten-Frage so viele dichotome Variable wie Antwortmöglichkeiten vorhanden sind zu erzeugen und diese gemeinsam mit den alten Fragen in eine neue Datei zu speichern.

Beispiel:
 In einer Umfrage wurde die Frage gestellt:

Welche Zeitungen lesen Sie?

Insgesamt wurden von allen Befragten 10 verschiedene Zeitungen genannt. Kein Befragter gibt aber mehr als 3 Zeitungen an, die er liest. Beim Kodieren und Schreiben der Daten wurden deswegen, um Schreibarbeit zu sparen, nur 3 Variable verwendet, in die die Zeitungen eingeschrieben wurden.

Allgemein formuliert:

1. Eine Frage besitzt x nicht-alternative Antwortmöglichkeiten.
 In unserem Beispiel: 10
2. Kein Befragter nutzt jedoch mehr als y Antwortmöglichkeiten.
 In unserem Beispiel: 3. Im Grenzfall kann y gleich x sein
 Demzufolge werden beim Schreiben der Daten nur y Variable für die Frage mit x Antwortmöglichkeiten angelegt.
3. Also erstellt eine Häufigkeitsverteilung für die gesamten x Antwortmöglichkeiten
4. Als Option wird angeboten: Also generiert automatisch x neue Variable und fügt diese am Ende des Datensatzes an. In unserem Beispiel sind das 10 neue Variable. Diese Variable sind dichotom. Sie können die Werte 1 (=genannt) und 0 (=nicht genannt) annehmen.

Die Datenmatrix für das Beispiel hat folgendes Aussehen:

Lebens- alter U1	Geschlecht U2	genannte Zeitung 1 U3	genannte Zeitung 2 U4	genannte Zeitung 3 U5
27	männl	Spiegel	Profil	Stern
43	weibl	FAZ	Spiegel	-
61	männl	Krone	Profil	Spiegel
81	weibl	-	-	-
46	männl	verweigert	verweigert	verweigert
:	:	:	:	:

Die 1. Person ist 27 Jahre alt, männlich und liest den Spiegel, das Profil und den Stern.
 Die 4. Person ist 81 Jahre alt, weiblich und liest überhaupt keine Zeitung
 Die 5. Person ist 46 Jahre alt, männlich und hat die Antwort verweigert

Beachte: Die Variablen für die genannten Zeitungen 1 bis 3 können im Datensatz an beliebiger Stelle stehen. Vor, hinter und zwischen diesen Variablen können beliebige andere Variable stehen. In unserem Beispiel steht z.B. vor diesen Variablen das Lebensalter und das Geschlecht des Befragten.

Wir verwenden folgende Codes:

```

keine Angabe = 0
FAZ           = 1
Stern        = 2
Spiegel      = 3
Profil       = 4
....
Antwort verweigert= -1

```

Der Benutzer kann entscheiden, ob eine Verweigerung wie "keine Angabe" gezählt werden soll oder ob eine Person, die verweigert, aus der Analyse ausgeschlossen werden soll

Die Datenmatrix ist also folgende:

	V1	V2	V3	V4	V5
--	--	--	--	--	--
27	1	3	4	2	
43	2	1	3	0	
61	1	5	4	3	
81	2	0	0	0	
46	1	-1	-1	-1	
.
.

Option: Neue dichotome Variable generieren.

Es müssen 10 neue Variable generiert werden, die an das Ende des Datensatzes gestellt werden. Sie erhalten die Variablen-Nr. V6 bis V15. Diese neuen Variablen werden von ALMO selbst erzeugt und automatisch an das Ende des Datensatzes gestellt. Die neue Datei wird gespeichert. In unserem Beispiel enthält die neue Datei folgende Variable:

	alte Variable					neue dichotome Variable									
	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15
27	1	3	4	2		0	1	1	1	0	0	0	0	0	0
43	2	1	3	0		1	0	1	0	0	0	0	0	0	0
61	1	5	4	3		0	0	1	1	1	0	0	0	0	0
81	2	0	0	0		0	0	0	0	0	0	0	0	0	0
46	1	-1	-1	-1		0	0	0	0	0	0	0	0	0	0
18	1	5	0	0		0	0	0	0	1	0	0	0	0	0
34	2	1	3	4		1	0	1	1	0	0	0	0	0	0

Terminologie:

Wir nennen
 die Variablen V2,V3,V4 aus der Datenmatrix: "Mehrfach-Variable"
 und die 10 neuen Variablen V6 bis V15: "dichotome Variable"

siehe Handbuch, Teil 3, Abschnitt P4.11

Programm-Bedienung ---> Hilfe

Speicher für x Variable

Im Beispiel wird mindestens Speicher für 15 Variable benötigt:
 Der eingelesene Datensatz umfasst die Variablen V1 bis V5
 ALMO generiert die neuen "dichotomen Variablen" V6 bis V15

Vereinbaren Sie jedoch ca. 6 bis 10 Variable mehr

Vereinbare Variable= 20

Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

Datei der Variablennamen [Hilfe](#)
"C:\Almo7\TESTDAT\ZEITUNG2.NAM"
zeige zeige = Namensdatei in Output zeigen
leer = nicht

Freie Namensfelder [Hilfe](#)
erzeuge zusätzliche Namensfelder

Datei aus der gelesen wird [Hilfe](#)
bei Datei-Problemen
"C:\Almo7\TESTDAT\ZEITUNG2.FRE"
frei Format der Daten [Hilfe](#)
U1:5 der Datensatz enthält diese Variablen
Bei Format DIREKT schreiben Sie: alle_U

Wenn Dateiformat FIX oder Nicht-Standard-FREI [Hilfe](#)

Option: Ein- und Ausschliessen von Untersuchungseinheiten

Option: Umkodierungen und Kein-Wert-Angaben

die im Datensatz enthaltenen "Mehrfach-Variablen"
Zeitung1, Zeitung2, Zeitung3
10 Zahl der Antwortmöglichkeiten
0 Kodeziffer für "keine Angabe" [Hilfe](#)
-1 Kodeziffer für "Antwort verweigert" bzw. für "Kein-Wert" [Hilfe](#)
0 0= "Antwort verweigert" bzw. "Kein-Wert" [Hilfe](#)
wie "keine Angabe" behandeln
1= Datensatz ausschließen,
wenn "Antwort verweigert" bzw. "Kein-Wert"



4.11.3.1 Erläuterungen zu den Eingabeboxen von Prog00mi

Box "Speicher fuer x Variable"

Siehe Almo-Dokument Nr.0 "Arbeiten mit Almo", Abschnitt P0.1

Box "Option: Weitere Vereinbarungen"

Siehe Almo-Dokument Nr.0 "Arbeiten mit Almo", Abschnitt P0.2

Box "Datei der Variablennamen"

Box "Freie Namensfelder"

Siehe P0.3

Box "Datei aus der gelesen wird"

Box " Wenn Dateiformat FIX oder Nicht-Standard-FREI"

Siehe P0.4

Box "Option: Ein- und Ausschliessen von Untersuchungseinheiten"

Siehe P0.7.1

Box "Option: Umkodierungen und Kein-Wert-Angaben"

Siehe P0.4

Box "die im Datensatz enthaltenen "Mehrfach-Variablen"

1. *Eingabefeld*: Geben Sie hier die Variablen mit Mehrfach-Antworten an. In unserem Beispiel sind dies die 3 Variablen Zeitung1, Zeitung2 und Zeitung3.

2. *Eingabefeld*: Geben Sie hier die Zahl der Antwortmöglichkeiten an. In unserem Beispiel wurden von den Befragten maximal 10 Zeitungen genannt. In das Eingabefeld wird somit 10 geschrieben.

3. *Eingabefeld*: Kodeziffer für "keine Angabe"
Es wird unterstellt, dass die verschiedenen Antwortmöglichkeiten

- (a) mit 1 beginnend aufsteigend kodiert sind und
- (b) in allen Mehrfach-Fragen gleich kodiert sind

In unserem Beispiel sind die 10 Zeitungen von 1 bis 10 kodiert und dies in den 3 Mehrfach-Fragen gleich. So ist etwa die Zeitung "Spiegel" in allen 3 Mehrfach-Fragen mit 3 kodiert.

Sind diese beiden Bedingungen nicht erfüllt - was sehr ungewöhnlich ist - dann müssen die Codeziffern so umkodiert werden, dass diese Bedingungen erfüllt sind. Das geschieht in der Optionsbox "Umkodierungen und Kein-Wert-Angaben".

Beispiel: Nehmen wir an, bei der Variablen "Zeitung3" sei, abweichend von den anderen beiden ZeitungsvARIABLEN, "Spiegel" mit 4 und "Profil" mit 3 kodiert. Bei den anderen beiden ZeitungsvARIABLEN, "Zeitung1" und "Zeitung2" wurde umgekehrt kodiert. In diesem Fall schreiben Sie in die Box "Option: Umkodierungen und Kein-Wert-Angaben"

Zeitung3 (4 = 3; 3 = 4)

(aus 4 soll 3 werden und aus 3 soll 4 werden). Damit erhält diese Variable denselben Code für Spiegel und Profil wie die anderen Variablen

Üblicherweise wird die 0 als Codeziffer verwendet, wenn der Befragte auf eine Frage mit Mehrfach-Antwort keine konkrete Nennung abgibt. In unserem Beispiel liest der 2. Befragte nur 2 Zeitungen. Für die 3. Zeitung gibt er keine Nennung ab. Für "Zeitung3" wird deswegen 0 kodiert.

In das Eingabefeld können mehrere durch Beistrich getrennte Zahlen eingetragen werden. Möglich ist auch ein durch Doppelpunkt gekennzeichnetes Werte-Intervall. Beispiel: Wenn "keine Angabe" mit 0 und 91 bis 99 kodiert wurde, dann wird in das Eingabefeld geschrieben:

0, 91:99

Beachte:

Es wird unterstellt, dass der Code für "keine Angabe" für alle Mehrfach-Fragen derselbe ist. Ist das nicht der Fall - was sehr ungewöhnlich ist - dann müssen die Codeziffern für "keine Angabe" für die verschiedenen Mehrfach-Fragen so umkodiert werden, dass sie alle gleich sind. Das geschieht in der Optionsbox "Umkodierungen und Kein-Wert-Angaben".

4. *Eingabefeld:* Geben Sie hier die Kodeziffer an, die für "Antwort verweigert" bzw. für "Kein-Wert" geschrieben wurde. Es können auch mehrere durch Beistrich getrennte Codes angegeben werden. Möglich ist auch ein durch Doppelpunkt gekennzeichnetes Werte-Intervall. Beispiel:

-1, 11:19

Ist in der auszuzählenden Datei bereits der Algo-interne Kein-Wert-Code enthalten, dann kann das Eingabefeld leer bleiben

Beachte: Es wird unterstellt, dass der Kein-Wert-Code für alle Mehrfach-Fragen derselbe ist. Ist das nicht der Fall - was sehr ungewöhnlich ist - dann müssen die Codeziffern für "Kein_Wert" für die verschiedenen Mehrfach-Fragen so umkodiert werden, dass sie alle gleich sind. Das geschieht in der Optionsbox "Umkodierungen und Kein-Wert-Angaben".

Beispiel: Bei der Variablen "Zeitung3" ist, abweichend von den anderen beiden Variablen, "Antwort verweigert" bzw. KeinWert mit 0 kodiert. Bei den anderen Variablen wurde -1 kodiert. In diesem Fall schreiben Sie in die Box "Option: Umkodierungen und Kein-Wert-Angaben"

Zeitung3 (0 = -1)

Damit erhält diese Variable denselben Code für KeinWert wie die anderen Variablen

5. *Eingabefeld:* Der Benutzer kann entscheiden, ob eine Antwortverweigerung wie "keine Angabe" gezählt werden soll oder ob eine Person, die verweigert, aus der Analyse ausgeschlossen werden soll.

Schreiben Sie

- 0 wenn "Antwort verweigert" bzw. "Kein-Wert" wie "keine Angabe" behandelt werden soll
- 1 wenn der Datensatz ausgeschlossen werden soll, wenn "Antwort verweigert" bzw. "Kein-Wert" vorliegt

Box "Namen für die Antwortmöglichkeiten"



Den neuen generierten dichotomen Variablen können Namen gegeben werden Das ist nicht obligatorisch. Werden keine Namen vergeben, dann schreibt Almo in der Ergebnis-Tabelle nur die fortlaufenden Nummern 1, 2, 3,

Wenn Sie jedoch Namen vergeben, dann ist folgendes zu beachten.

1. Die Reihenfolge der Namen muss mit den Codeziffern der Wahlmöglichkeiten (im Beispiel: der Zeitungen) bzw. der Reihenfolge der neuen generierten dichotomen Variablen übereinstimmen.

Wird "FAZ" mit 1 kodiert und "Stern" mit 2, dann steht in der 1. neuen dichotomen Variablen die FAZ und in der 2. der Stern. Dann muss als 1. Name "FAZ" und als 2. "Stern" geschrieben werden etc....

2. Die Namen werden untereinander je in ein Eingabefeld geschrieben

3. Die Namen können (müssen aber nicht) mit den Codeziffern in Klammern geschrieben werden. Es gibt also 2 Möglichkeiten

1. Möglichkeit: Ohne Codeziffern

FAZ
Stern
Spiegel
Profil
Krone
Bild
Bunte
Frau
Focus
Kurier

2. Möglichkeit: Mit Codeziffern

(1) FAZ
(2) Stern
(3) Spiegel
(4) Profil
(5) Krone
(6) Bild
(7) Bunte
(8) Frau
(9) Focus
(10) Kurier

4. Werden die Namen mit Codeziffern geschrieben, dann dürfen mitten drin Namen fehlen. Beispiel:

- (1) FAZ
- (2) Stern
- (3) Spiegel
- (4) Profil
- (5) Krone
- (8) Frau
- (9) Focus
- (10) Kurier

Die Namen (6) Bild und (7) Bunte fehlen. Almo kann in der Ergebnis-Tabelle Namen richtig zuordnen, da die Codeziffern mit den Namen verknüpft sind. Für "Bild" und "Bunte" schreibt Almo in der Ergebnistabelle dann nur "6" und "7".

BEACHTEN: Diese Art der Namensgebung ist jedoch nicht möglich, wenn (in der folgenden Box "Erzeuge neue Datei (inkl. der "dichotomen Variablen")" eine neue Namensdatei angefordert wird. Siehe nachfolgend die Erläuterungen zum Eingabefeld 2.

5. Insgesamt können 88 Namen vergeben werden. Sind mehr Wahlmöglichkeiten vorhanden, z.B. 100, dann kann den letzten 12 kein Name gegeben werden. Almo verwendet dann für diese die Ausprägungsnummern 89, 90, 91,99,100

Box "Option: Erzeuge neue Datei (inkl. der "dichotomen Variablen")"

Option: Erzeuge neue Datei (inkl. der "dichotomen Variablen")

Optionsbox geöffnet:

Loesche wieder diese Box

Option: Erzeuge neue Datei (inkl. der "dichotomen Variablen")

Wollen Sie keine neue Datei anlegen, dann löschen Sie den Dateinamen durch Klick auf den doppelköpfigen Pfeil

↔ "C:\Almo7\Progs\Neudat"

Geben Sie einen neuen Dateinamen ohne Erweiterung an

Almo erzeugt 2 Dateien:

1. eine nicht lesbare Almo-Arbeitsdatei mit der Erweiterung `__.dir`
2. eine anschaulichere Datei im freien Format mit der Erweiterung `__.fre`

In den unter 1. und 2. angegebenen neuen Dateien sind nun enthalten:

- die Variablen aus der alten Datei im Beispiel: U1:4
- die generierten dichotomen Variablen, die als letzte Variable hinter die Variablen der alten Datei gestellt wurde im Beispiel: U5:14 Hilfe

↑↓ 1

1 = Almo erzeugt auch eine Datei der Variablen- und Ausprägungsnamen mit der Erweiterung `__.nam`

bitte lesen ---> Hilfe

0 = keine Namensdatei erzeugen

Die Namensdatei enthält die alten Namen

- (im Beispiel: die Namen 1 bis 4)

und die Namen der generierten dichotomen Variablen

- (im Beispiel: die Namen 5 bis 14)

Der Benutzer kann eine neue Datei erzeugen. Sie enthält dann die seitherigen alten Variablen und die neuen, generierten, dichotomen Variablen. In unserem Beispiel sieht ein neuer Datensatz z.B. folgendermaßen aus:

alte Variable					neue dichotome Variable									
V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15
43	1	1	3	-1	1	0	1	0	0	0	0	0	0	0

V1 ist das Lebensalter (=43)

V2 ist das Geschlecht (=1)

V3,4,5 sind die "Mehrfach-Variablen", also Zeitung1 bis Zeitung3
Für Zeitung3 wurde von diesem Befragten die Antwort verweigert (= -1)

V6:15 sind die 10 neuen, generierten dichotomen Variablen
Für "Zeitung genannt" wurde =1 gesetzt, für "nicht genannt" =0

Mit der neuen Datei können jetzt alle in Almo möglichen Analysen gerechnet werden

1. *Eingabefeld*: Geben Sie einen beliebigen neuen Dateinamen an. Der Dateiname darf keine Erweiterung besitzen (also z.B. ____fre) Almo erzeugt dann 2 Dateien:

1. eine nicht lesbare Almo-Arbeitsdatei mit der Erweiterung ____dir
2. eine anschauliche Datei im freien Format mit der Erweiterung ____fre

2. *Eingabefeld*: Wenn Sie hier eine "1" eingeben, dann erzeugt Almo auch eine Datei der Variablen- und Ausprägungsnamen mit der Erweiterung ____nam. Die Namensdatei enthält dann die alten Namen, in unserem Beispiel die Namen der Variablen V1 bis V5 und die Namen der neuen generierten dichotomen Variablen in unserem Beispiel die Namen der Variablen V6 bis V15.

In unserem Beispiel wird folgende Namensdatei erzeugt:

```
N1=Lebensalter;
N2=Geschlecht:männlich,weiblich;
N3=Zeitung1;
N4=Zeitung2;
N5=Zeitung3;

N6 =FAZ:      (0)nicht genannt, (1)genannt;
N7 =Stern:    (0)nicht genannt, (1)genannt;
N8 =Spiegel:  (0)nicht genannt, (1)genannt;
N9 =Profil:   (0)nicht genannt, (1)genannt;
N10 =Krone:   (0)nicht genannt, (1)genannt;
N11=Bild:     (0)nicht genannt, (1)genannt;
N12=Bunte:    (0)nicht genannt, (1)genannt;
N13=Frau:     (0)nicht genannt, (1)genannt;
N14=Focus:    (0)nicht genannt, (1)genannt;
N15=Kurier:   (0)nicht genannt, (1)genannt;
```

Die Namen N1 bis N5 sind die alten Namen. Die Namen N6 bis N15 sind die Namen der neu generierten dichotomen Variablen.

Folgende Probleme können auftreten:

1. Der Benutzer hat beispielsweise nur für die ersten 5 Wahlmöglichkeiten Namen angegeben. Etwa so:

- (1) FAZ
- (2) Stern
- (3) Spiegel
- (4) Profil
- (5) Krone

In diesem Fall schreibt Almo selbstverständlich nur die neuen Namen N6 bis N10. Die restlichen neu generierten dichotomen Variablen erhalten keine Variablennamen:

2. Der Benutzer hat einige Wahlmöglichkeiten nicht benannt.
Beispiel:

- (1) FAZ
- (2) Stern
- (3) Spiegel
- (4) Profil
- (5) Krone
- (6) Bild

- (8) Frau
- (9) Focus
- (10) Kurier

Die 7. Wahlmöglichkeit wurde nicht benannt. Almo wird dann ab der 7. neu generierten dichotomen Variablen falsche Variablennamen zuordnen. Almo bringt folgende Fehlermeldung

```

***** FEHLER
      Zahl der geschriebenen Variablennamen
und Zahl der Variablennummern hinter "Name..."
      stimmen nicht überein

```

EMPFEHLUNG: Vermeiden Sie die oben in 2 angegebene Namensgebung. Laden Sie die erzeugte Namensdatei in ein Almo-Fenster und kontrollieren Sie dann die Namensgebungen der neuen generierten dichotomen Variablen.

Box "Option: "Aussehen" der auszugebenden Tabelle bzw. Matrix"
 Siehe P0.9

Box "Grafik-Optionen"
 Siehe P0.10

P4.11.3.2 Ergebnisse aus Prog00mi

Zahl der eingelesenen Datensätze: 120

```

***** MITTEILUNG
      Bei Variablen mit Mehrfach-Antworten entsteht die oben
angegebene "Zahl der eingelesenen Datensätze" aus der Zahl der
tatsächlich eingelesenen Datensätze mal Zahl der Auswahlmöglichkeiten

      die Zahl der tatsächlich eingelesenen Datensätze ist: 12

```

```

***** MITTEILUNG
      Bedingt durch KEIN_WERT-Fälle können für die Tabellen
      unterschiedliche Summen entstehen und auch Summen, die
      von der oben angegebenen Zahl der Datensätze abweichen

```

```

***** MITTEILUNG
      Almo fasst die mehrfachen Antwortmöglichkeiten zusammen in V20
      und verwendet zur Bildung der Tabelle die beiden Variablen
      V19 und V20

```

			Häufigkeit		Summe
			nicht genannt	genannt	
Antwort	FAZ	1	10	2	12
	Stern	2	11	1	12
	Spiegel	3	5	7	12
	Profil	4	7	5	12
	Krone	5	10	2	12
	Bild	6	10	2	12
	Bunte	7	9	3	12
	Frau	8	10	2	12
	Focus	9	11	1	12
	Kurier	10	10	2	12
Summe			93	27	120

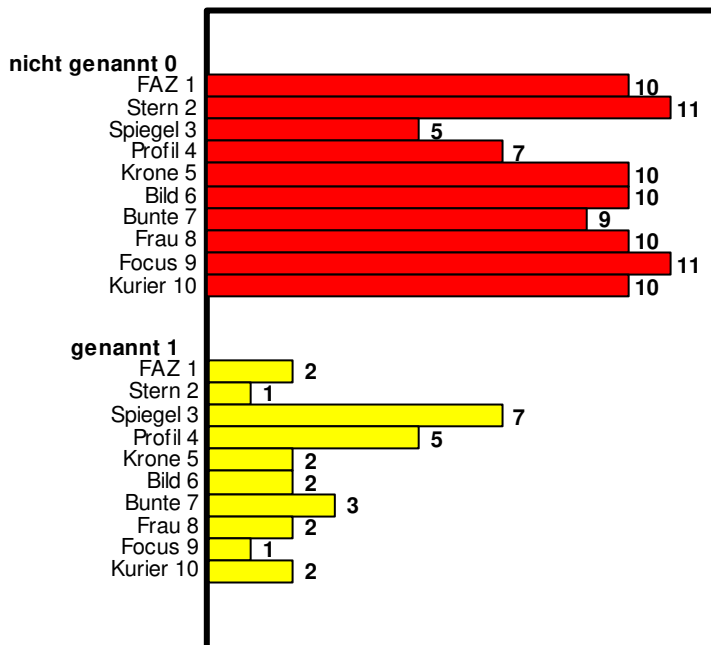
***** Erläuterung

In der Spaltensumme steht jeweils 12. Das ist die Zahl der Befragten. Von diesen 12 Befragten haben 2 die FAZ genannt und 10 nicht etc. Die 10 Zeitungen wurden von den 12 Befragten insgesamt 27 mal genannt und 93 mal nicht genannt. Das ist der unteren Spaltensumme zu entnehmen.

Almo zeichnet dann noch folgendes Säulendiagramm.

2-dimensionale Verteilung

Antwort
und
Häufigkeit



Zeilenweise prozentuiert

			Häufigkeit		Summe
			nicht genannt	genannt	
Antwort	FAZ	1	83.33	16.67	100.00
	Stern	2	91.67	8.33	100.00
	Spiegel	3	41.67	58.33	100.00
	Profil	4	58.33	41.67	100.00
	Krone	5	83.33	16.67	100.00
	Bild	6	83.33	16.67	100.00
	Bunte	7	75.00	25.00	100.00
	Frau	8	83.33	16.67	100.00
	Focus	9	91.67	8.33	100.00
	Kurier	10	83.33	16.67	100.00
Summe			77.50	22.50	100.00

***** **Erläuterung**

Von den 12 Befragten haben 16.67 % die FAZ genannt und 83.33 % nicht etc. Die geringste Häufigkeit an Nennungen haben der Stern und der Focus mit 8.33 %

Almo zeichnet für diese Tabelle dann noch ein Säulendiagramm.

Spaltenweise prozentuiert

			Häufigkeit		Summe
			nicht genannt 0	genannt 1	
Antwort	FAZ	1	10.75	7.41	10.00
	Stern	2	11.83	3.70	10.00
	Spiegel	3	5.38	25.93	10.00
	Profil	4	7.53	18.52	10.00
	Krone	5	10.75	7.41	10.00
	Bild	6	10.75	7.41	10.00
	Bunte	7	9.68	11.11	10.00
	Frau	8	10.75	7.41	10.00
	Focus	9	11.83	3.70	10.00
	Kurier	10	10.75	7.41	10.00
Summe			100.00	100.00	100.00

***** **Erläuterung**

Die 10 Zeitungen wurden 27 mal genannt (siehe 1. Tabelle). Von diesen 27 Nennungen entfielen 7.41 % auf die FAZ etc. Die geringste Häufigkeit an Nennungen (von den 27) haben der Stern und der Focus mit 3.7 %

Almo zeichnet für diese Tabelle dann noch ein Säulendiagramm.

P4.11.4 Kategoriale Methode: Kreuztabellierung von Mehrfach Fragen mit Prog00mo

Die 3 "Mehrfach-Variablen" Zeitung1 bis Zeitung3 sollen gegen das Geschlecht der

Befragten tabelliert werden. Dazu ist es unumgänglich diese 3 "Mehrfach-Variablen" zuvor in 10 dichotome Variable aufzulösen und dann diese gegen das Geschlecht zu tabellieren. Dies leistet Prog00mo.

Prog00mo ist mit dem in P4.11.3 beschriebenen Prog00mi identisch. Es kommen lediglich noch 2 Boxen hinzu:

- Box "Die Eine-Antwort-Variablen"
- Box "Option: Koeffizienten berechnen"

Diese beiden Boxen haben wir schon oben in Abschnitt P4.11.2 erläutert.

P4.11.4.1 Ergebnisse aus Prog00mo

Zahl der eingelesenen Datensätze: 120

***** MITTEILUNG
 Bei Variablen mit Mehrfach-Antworten entsteht die oben angegebene "Zahl der eingelesenen Datensätze" aus der Zahl der tatsächlich eingelesenen Datensätze mal Zahl der Auswahlmöglichkeiten

die Zahl der tatsächlich eingelesenen Datensätze ist: 12

***** MITTEILUNG
 Bedingt durch KEIN_WERT-Fälle können für die Tabellen unterschiedliche Summen entstehen und auch Summen, die von der oben angegebenen Zahl der Datensätze abweichen

#####

Tabelle 1

***** MITTEILUNG
 Also fasst die mehrfachen Antwortmöglichkeiten zusammen in V20 und verwendet zur Bildung der Tabelle die beiden Variablen V2 und V20

Variable 20 genannt
 mit
 Variable 2 Geschlecht

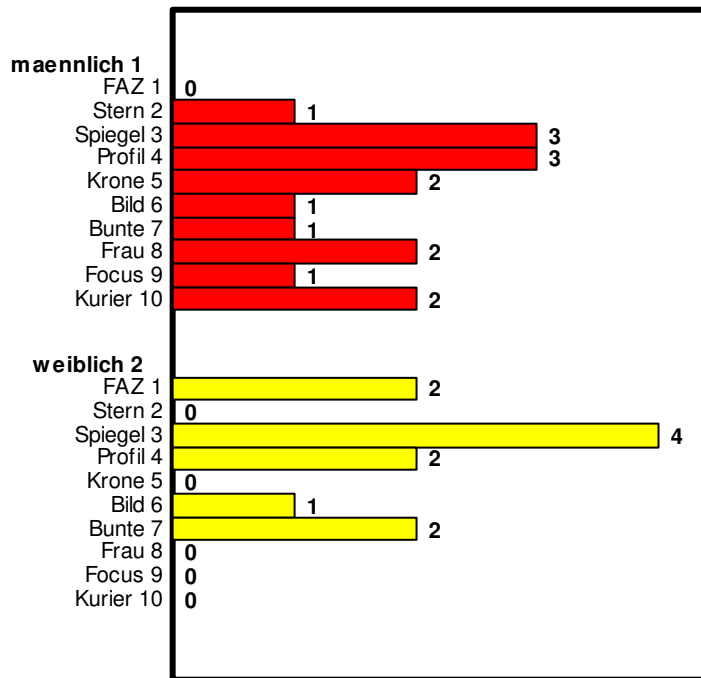
			Geschlecht		Summe
			maennlich	weiblich	
			1	2	
genannt	FAZ	1	0	2	2
	Stern	2	1	0	1
	Spiegel	3	3	4	7
	Profil	4	3	2	5
	Krone	5	2	0	2
	Bild	6	1	1	2
	Bunte	7	1	2	3
	Frau	8	2	0	2
	Focus	9	1	0	1
	Kurier	10	2	0	2
Summe			16	11	27

***** Erläuterung

Almo gibt zunächst die Tabellen für die Nennungen und später für die Nicht-Nennungen aus. Die Männer geben insgesamt 16 Nennungen ab, die Frauen 11. Dies ist aus der unteren Summen-Zeile ablesbar. Aus der hinteren Summen-Spalte ergibt sich, dass FAZ 2 mal genannt wurde etc.

Almo zeichnet für diese Tabelle dann noch ein Säulendiagramm.

2-dimensionale Verteilung
genannt
und
Geschlecht



Zeilenweise prozentuiert

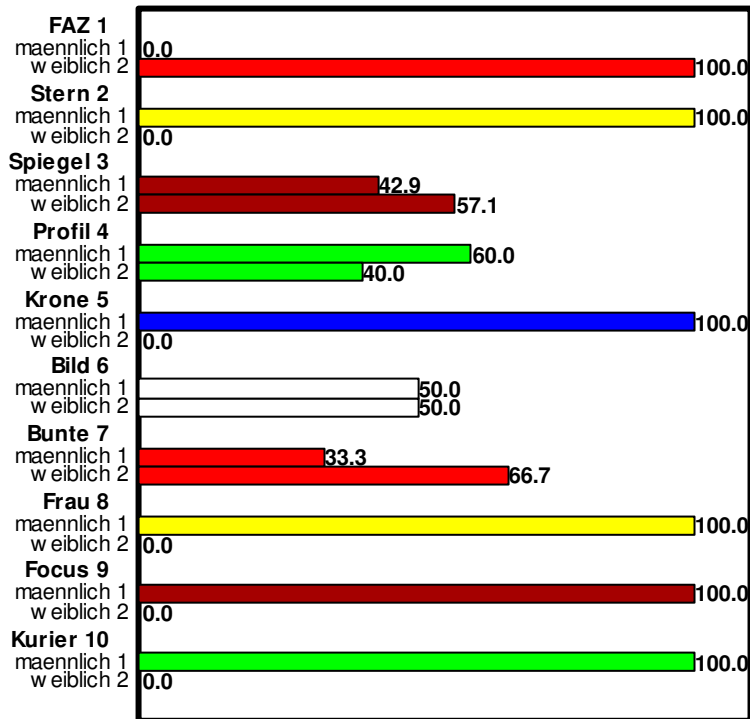
			Geschlecht		Summe
			maennlich 1	weiblich 2	
genannt	FAZ	1	0	100.0000	100.0000
	Stern	2	100.0000	0	100.0000
	Spiegel	3	42.8571	57.1429	100.0000
	Profil	4	60.0000	40.0000	100.0000
	Krone	5	100.0000	0	100.0000
	Bild	6	50.0000	50.0000	100.0000
	Bunte	7	33.3333	66.6667	100.0000
	Frau	8	100.0000	0	100.0000
	Focus	9	100.0000	0	100.0000
	Kurier	10	100.0000	0	100.0000
Summe			59.2593	40.7407	100.0000

***** **Erläuterung**

Hier wird die jeweilige Zeitung auf 100 % gesetzt. So wird beispielweise die "Bunte" zu 33.33 % von Männern, aber zu 66.67 % von Frauen genannt.

Almo zeichnet für diese Tabelle dann noch folgendes Säulendiagramm.

2-dimensionale Verteilung
genannt
und
Geschlecht



Spaltenweise prozentuiert

			Geschlecht		Summe
			maennlich 1	weiblich 2	
genannt	FAZ	1	0	18.1818	7.4074
	Stern	2	6.2500	0	3.7037
	Spiegel	3	18.7500	36.3636	25.9259
	Profil	4	18.7500	18.1818	18.5185
	Krone	5	12.5000	0	7.4074
	Bild	6	6.2500	9.0909	7.4074
	Bunte	7	6.2500	18.1818	11.1111
	Frau	8	12.5000	0	7.4074
	Focus	9	6.2500	0	3.7037
	Kurier	10	12.5000	0	7.4074
Summe			100.0000	100.0000	100.0000

***** Erläuterung

Die Nennungen, die die Männer und die Frauen abgegeben haben, werden hier jeweils auf 100 % gesetzt. Mann erkennt, dass 18.75 % der Nennungen von Männern für den Spiegel und das Profil abgegeben wurden etc.

Koeffizienten fuer Tabelle 1

Signifikanztest

Chi-Quadrat = 10.0965 df = 9 Signifikanz p = 0.343
Signifikanz (1-p)*100 = 65.715%

Ein Wert ueber ca. 95% bedeutet:

Zwischen den Variablen besteht
ein signifikanter Zusammenhang

Erwartungswerte kleiner 1 = 45%
kleiner 5 =100% der Zellen

***** Erläuterung

Almo rechnet noch einen Chi-Quadrat-Test, der lediglich pauschal aussagt, ob in der Tabelle ein Zusammenhang zwischen Zeitung und Geschlecht besteht oder nicht. Ob beispielsweise Frauen signifikant eher die "Bunte" angeben als Männer ist erst den nachfolgend ausgegebenen Tabellen zu entnehmen.

Almo tabelliert dann in gleicher Weise die Nicht-Nennungen mit dem Geschlecht. Üblicherweise sind diese Tabellen weniger interessant. Wir bilden sie deswegen hier auch nicht ab.

Wurde in Prog00mo die Optionsbox "Koeffizienten berechnen" geöffnet, dann gibt Almo 10 zwei-dimensionale Tabellen aus, die jeweils den Zusammenhang zwischen einer Zeitung und dem Geschlecht darstellen. Wir betrachten hier die 3. Tabelle "Spiegel" mit Geschlecht.

Tabelle 3
Variable 8 Spiegel
mit
Variable 2 Geschlecht

		Geschlecht		Summe
		maennlich 1	weiblich 2	
Spiegel	nicht genannt 0	4	1	5
	genannt 1	3	4	7
Summe		7	5	12

***** Erläuterung

Die Tabelle zeigt, wie der "Spiegel" von den befragten Männern und Frauen genannt bzw. nicht genannt wurde. Die nachfolgenden Koeffizienten drücken dann aus, ob zwischen dem Geschlecht und dem Trinken von Apfelsaft ein Zusammenhang besteht. Diese Koeffizienten wurden ausführlich in Handbuch, Teil 3, Abschnitt P10.2 und P10.4. beschrieben. In unserem Beispiel sind die Häufigkeiten sehr gering - so dass ein Chi-Quadrat-Test für dieses Beispiel eigentlich nicht zulässig ist.

Koeffizienten fuer Tabelle 3

Signifikanztest

Chi-Quadrat = 1.6555 df = 1 Signifikanz p = 0.198
Signifikanz (1-p)*100 = 80.176%

Ein Wert ueber ca. 95% bedeutet:
Zwischen den Variablen besteht
ein signifikanter Zusammenhang

Erwartungswerte kleiner 1 = 0%
kleiner 5 =100% der Zellen

Korrelationskoeffizienten

Messniveau		Korrelationskoeffizient	zweiseitige Signifikanz	
der einen Variablen	der anderen Variablen		p	(1-p)*100
dichotom	dichotom	Phi = 0.3714	0.19824	80.1760
dichotom	polytom	Phi' = 0.3714	0.19824	80.1760
polytom	polytom	Kontingenzk.C(cor) = 0.4924	0.19824	80.1760
polytom	polytom	Tschuprow's T = 0.3714	0.19824	80.1760
polytom	polytom	Cramer's V = 0.3714	0.19824	80.1760
polytom	polytom	Lambda (asymm.) = 0.2000	0.76238	23.7625
ordinal	ordinal	Gamma = 0.6842	0.05743	94.2568
ordinal	ordinal	tau-b = 0.3714	0.24368	75.6325
dichotom	ordinal	biseriales tau-b = 0.3714	0.24368	75.6325
ordinal	ordinal	Rho = 0.3714	0.23453	76.5467
dichotom	quantit.	punktbiseriales r = 0.3714	0.23453	76.5467
quantit.	quantit.	Produkt-Moment r = 0.3714	0.23453	76.5467
Zeilenvariable				
Spaltenvariable				
quantit.	ordinal	Gross-Gamma I = 0.3714	0.21814	78.1865
ordinal	quantit.	Gross-Gamma II = 0.3714	0.21814	78.1865
polytom	quantit.	Eta = 0.3714	0.23453	76.5467
quantit.	quantit.	nichtlineares Eta = 0.3714	0.23453	76.5467
		Signifik. der Nichtlinearitaet	1.00000	0.0000
dichotom	dichotom	tetrachorisches r = 0.5823	0.20642	79.3576