

Skript 7 – Kreuztabellen und benutzerdefinierte Tabellen

Ziel: Analysieren und verdeutlichen von Zusammenhängen mehrerer Variablen, wie z.B. Anzahlen pro Kategorien; Mittelwert und Standardabweichung pro Kategorien, ...

Im Gegensatz zu Skript 6 („Häufigkeiten und Deskriptive Statistiken einer Variablen“) geht es hier vor allem um das Zusammenspiel mehrerer Variablen.

Hierbei gibt es beim Auszählen von (nominalen oder ordinalen) Variablen nach einer anderen (nominalen oder ordinalen Variablen) mit dem Dialogfeld „Kreuztabellen“ und dem Dialogfeld „Benutzerdefinierte Tabellen“ noch zwei Möglichkeiten zur Durchführung. Jedoch bietet das Dialogfeld „Benutzerdefinierte Tabellen“ weit mehr Möglichkeiten zur Kombination und Zusammenstellung von Tabellen, so dass **für viele Anwendungen das Dialogfeld „Benutzerdefinierte Tabelle“ die bessere Wahl** ist.¹

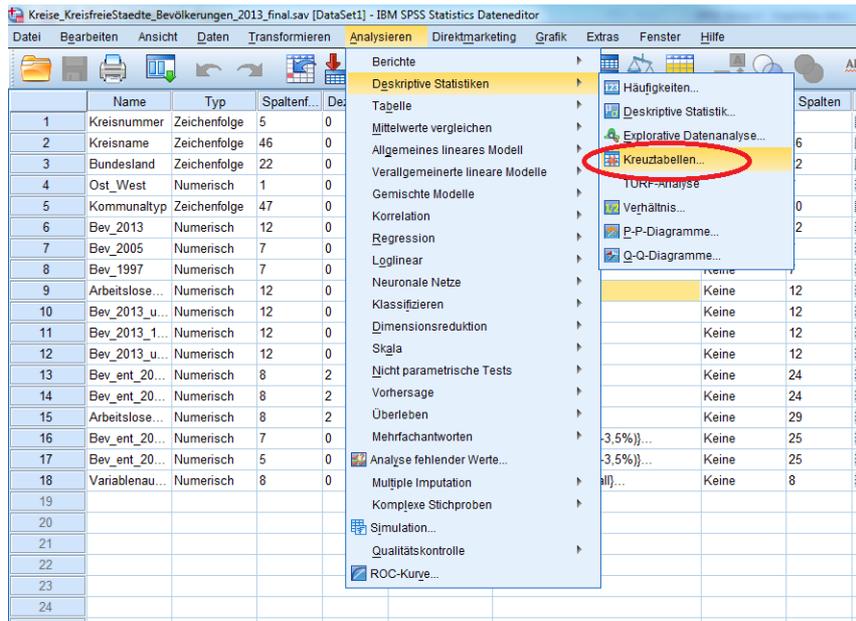
Beispieldatei: Kreise_KreisfreieStaedte_Bevölkerungen_2013_final.sav , verfügbar unter: <https://www.hcu-hamburg.de/it-und-medien/stadtlabor/>: Bevölkerung der Kreise und Kreisfreien Städte Deutschlands 1997, 2005 und 2013; Arbeitslosigkeit 2013 sowie Bevölkerungsentwicklungen und deren Klassifikationen; unterteilt nach Bundesland und Ost/West-Deutschland.

Vorgehen 1: Kreuztabellen

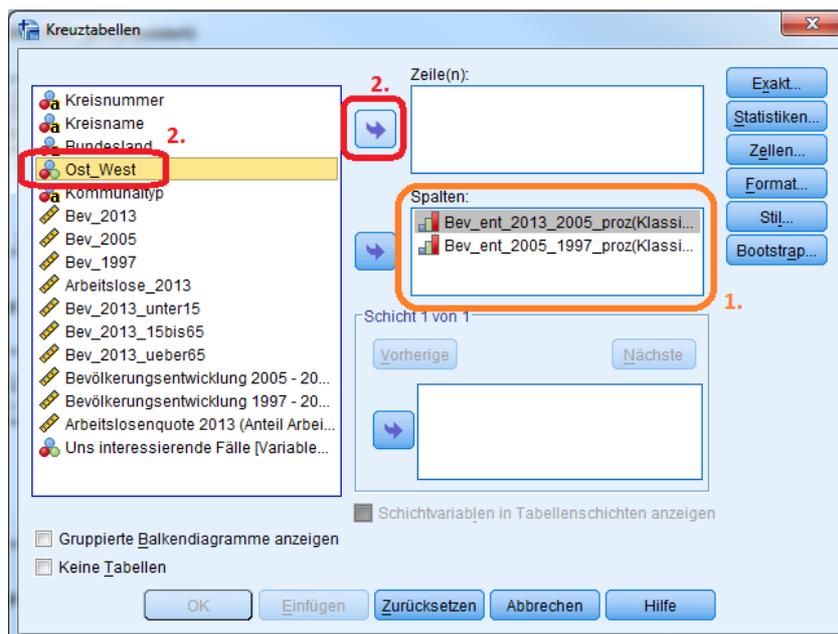
Ziel: Häufigkeitsauszählung der Kategorien einer oder mehrerer Variable, unterteilt nach Kategorien einer anderen Variable. Alle Variablen müssen nominal oder ordinal sein.

1. **Aufrufen des Dialogfelds ‚Kreuztabellen‘** : „Analysieren“ | „Deskriptive Statistiken“ | „Kreuztabellen“

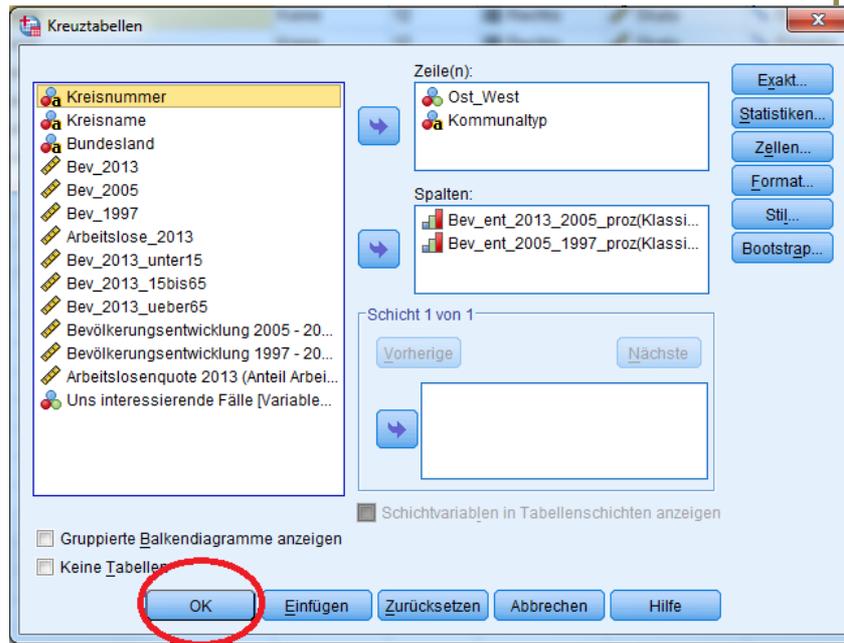
¹ Ausnahme: Es sollen weitergehende statistische Analysen wie z.B. der Chi-Quadrat-Test durchgeführt werden. Hier ist das Dialogfeld ‚Kreuztabellen‘ nötig. Siehe hierfür zum Beispiel das Skript „chi²-Test“ unter <http://www.home.uni-osnabrueck.de/rniketta/method/html/spss-skripte.html>.



2. **Auswahl der auszählenden Variablen („Spalten“, 1.) pro Kategorie der anderen Variablen („Zeilen“, 2.).** Sowohl in Zeilen als auch in Spalten können mehrere Variablen eingegeben werden, es wird dann jede Variable in „Spalten“ mit jeder Variablen in „Zeilen(n)“ kombiniert und jeweils eine Kreuztabelle ausgegeben.



3. **Bestätigung mit ‚OK‘.** Ergebnis:



Kreuztabelle Ost_West*Bev_ent_2013_2005_proz(Klassiert)

Anzahl		Bev_ent_2013_2005_proz(Klassiert)			Gesamtsumme
		schumpfend (<-3,5%)	stabil (>-3,5% &<3,5%)	wachsend (>3,5%)	
Ost_West	West	108	190	26	324
	Berlin	0	1	0	1
	Ost	53	13	4	70
Gesamtsumme		161	204	30	395

Kreuztabelle Ost_West*Bev_ent_2005_1997_proz(Klassiert)

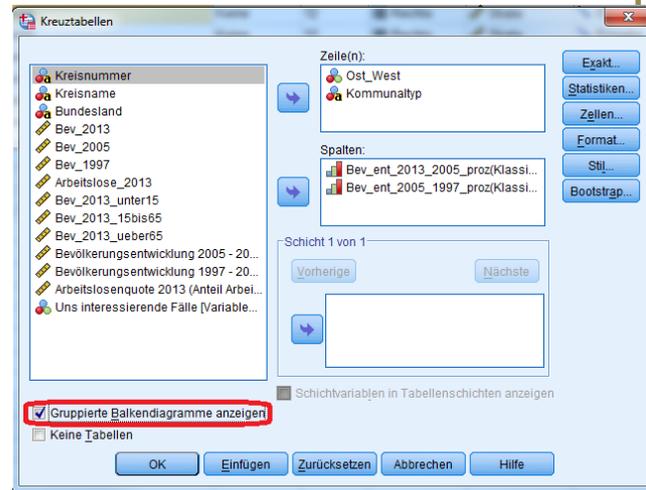
Anzahl		Bev_ent_2005_1997_proz(Klassiert)			Gesamtsumme
		schumpfend (<-3,5%)	stabil (>-3,5% &<3,5%)	wachsend (>3,5%)	
Ost_West	West	26	200	98	324
	Berlin	0	1	0	1
	Ost	43	8	9	60
Gesamtsumme		69	209	107	385

Kreuztabelle Kommunaltyp*Bev_ent_2013_2005_proz(Klassiert)

Anzahl		Bev_ent_2013_2005_proz(Klassiert)			Gesamtsumme
		schumpfend (<-3,5%)	stabil (>-3,5% &<3,5%)	wachsend (>3,5%)	
Kommunaltyp	Kommunalverband besonderer Art	1	1	0	2
	Kreisfreie Stadt	29	56	22	107
	Landkreis	131	147	8	286
Gesamtsumme		161	204	30	395

Kreuztabelle Kommunaltyp*Bev_ent_2005_1997_proz(Klassiert)

Anzahl		Bev_ent_2005_1997_proz(Klassiert)			Gesamtsumme
		schumpfend (<-3,5%)	stabil (>-3,5% &<3,5%)	wachsend (>3,5%)	
Kommunaltyp	Kommunalverband besonderer Art	1	1	0	2
	Kreisfreie Stadt	24	69	13	106
	Landkreis	44	139	94	277
Gesamtsumme		69	209	107	385

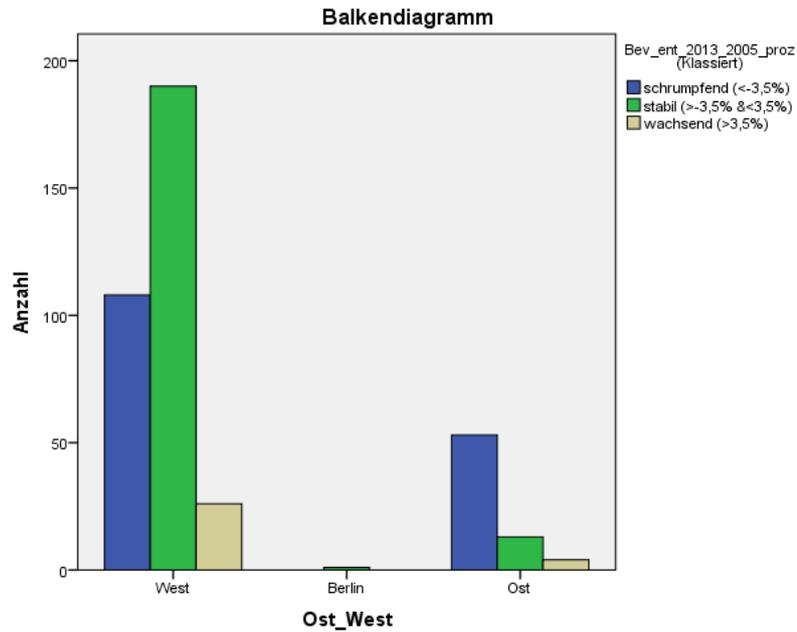


4. Wenn gewünscht, kann vorher noch gewählt werden, dass gruppierte Balkendiagramme ausgegeben werden sollen.

Ergebnis (Ausschnitt):

Kreuztabelle Ost_West*Bev_ent_2013_2005_proz(Klassiert)

Anzahl		Bev_ent_2013_2005_proz(Klassiert)			Gesamtsumme
		schrumpfend (<-3,5%)	stabil (>-3,5% &<3,5%)	wachsend (>3,5%)	
Ost_West	West	108	190	26	324
	Berlin	0	1	0	1
	Ost	53	13	4	70
Gesamtsumme		161	204	30	395



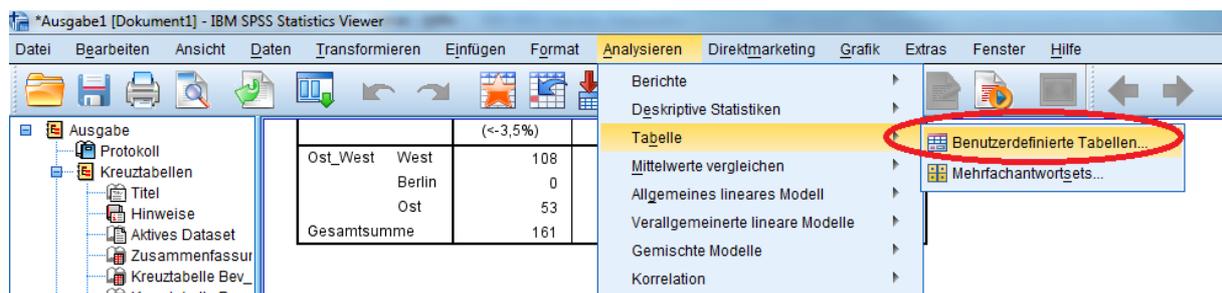
- Die Tabellen und Graphiken können durch Doppelklick bearbeitet werden und es kann beispielsweise ein entsprechender Titel vergeben werden.

Vorgehen 2: Benutzerdefinierte Tabellen

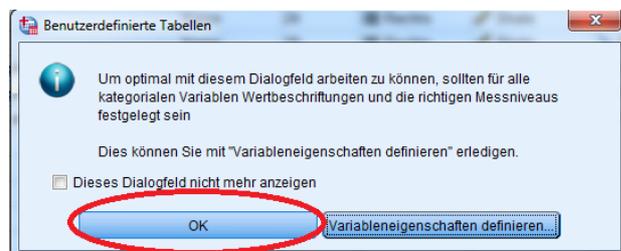
Ziel: Kombination verschiedener Häufigkeitsauszählungen und Statistiken, auch von verschiedenen Variablen, auf Wunsch unterteilt nach diversen Kategorien von anderen Variablen, in einer Tabelle.

Dies geschieht alles im Dialogfeld „Benutzerdefinierte Tabellen“. Dies ist ein ziemlich mächtiges Werkzeug, weshalb hier nur einige Aspekte beispielhaft gezeigt werden können.

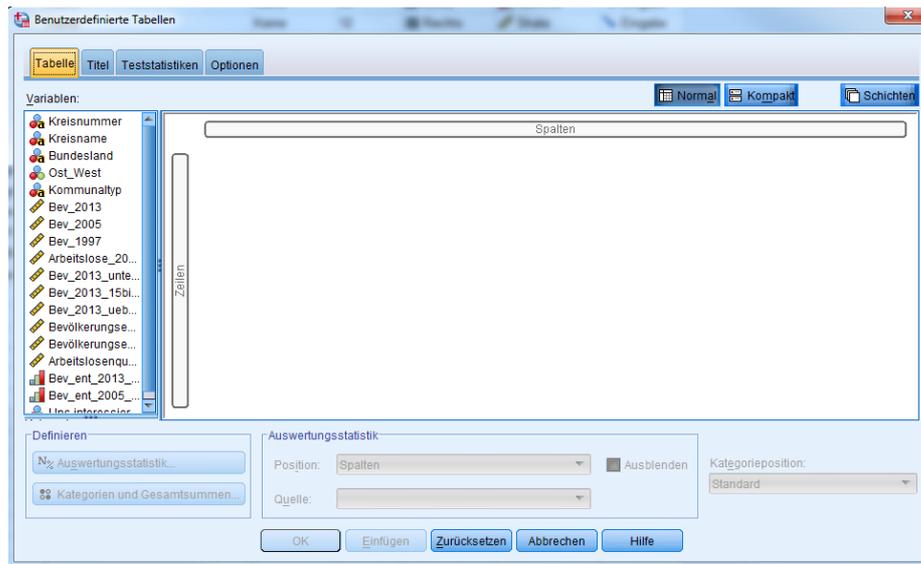
1. Aufrufen des Dialogfelds „Benutzerdefinierte Tabellen“: „Analysieren“ | „Tabelle“ | „Benutzerdefinierte Tabellen“



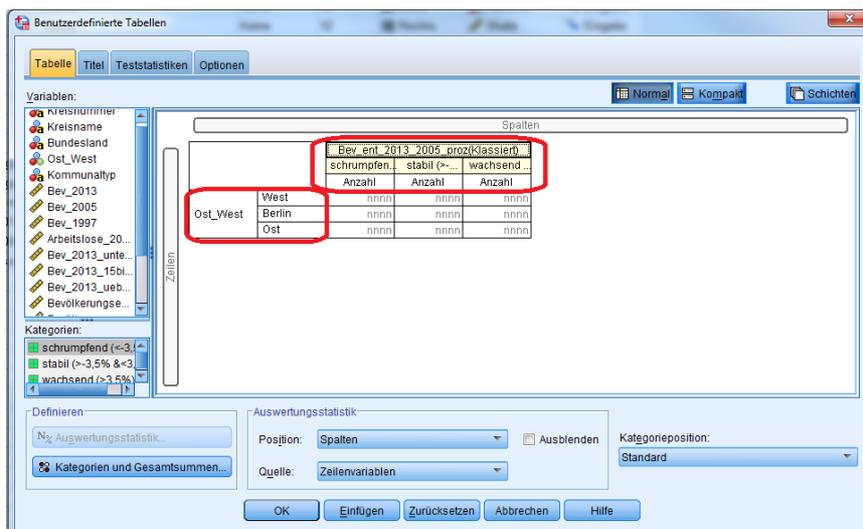
- Bestätigung der Sicherheitsabfrage.** Wenn noch nicht geschehen, sollten allerdings vor Verwenden der Tabelle noch die richtigen Skalenniveaus festgelegt werden (unter: Datenfenster | Variablenansicht). Sonst werden im Folgenden die falschen Einstellungen verwendet. Beispiel einer falschen Einstellung: Eine nominale Variable, in der ‚1‘ für ‚West‘ und ‚2‘ für ‚Ost‘ steht, wurde fälschlich als metrisch deklariert. Hier würde das Dialogfeld standardmäßig versuchen, den Mittelwert zu bilden, was natürlich unsinnig ist.



- Standardansicht des Dialoges „Benutzerdefinierte Tabelle“** . Hier kann jetzt die Variablenauswahl durch Drag&Drop und viele weitere Einstellungen vorgenommen werden.



4. **Beispiel 1: Kreuztabelle.** Im Beispiel wurde die Variable „Ost_West“ auf das Feld „Zeilen“, und die Variable „Bev_ent_2013_2005_klass“ auf das Feld „Spalten“ gezogen. Da es sich ausschließlich um nominale und ordinale Variablen handelt, wird standardmäßig eine Häufigkeitsauszählung ausgegeben.

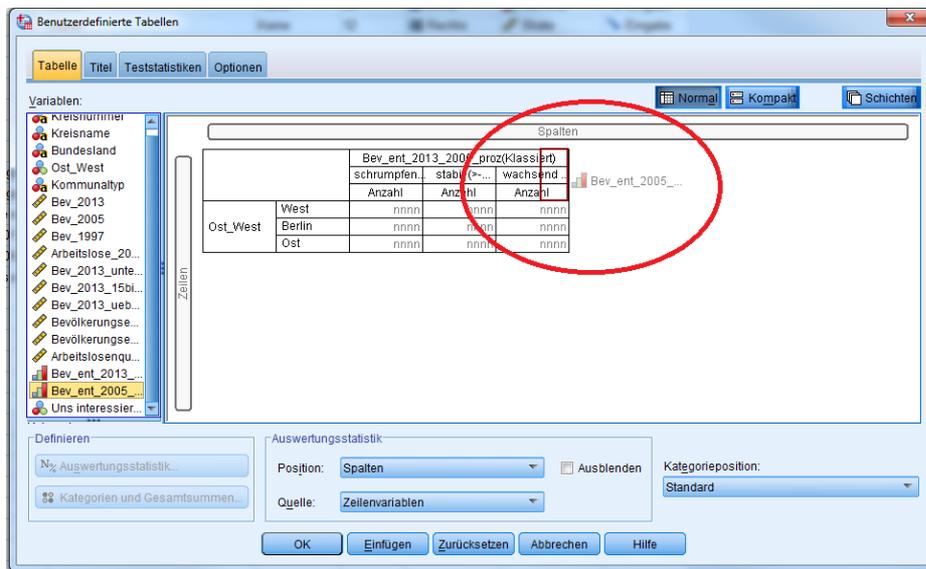


Die eigentliche Tabelle wird hier wie bei allen anderen Beispielen nach Klick auf ‚OK‘ ausgegeben. Ergebnis:

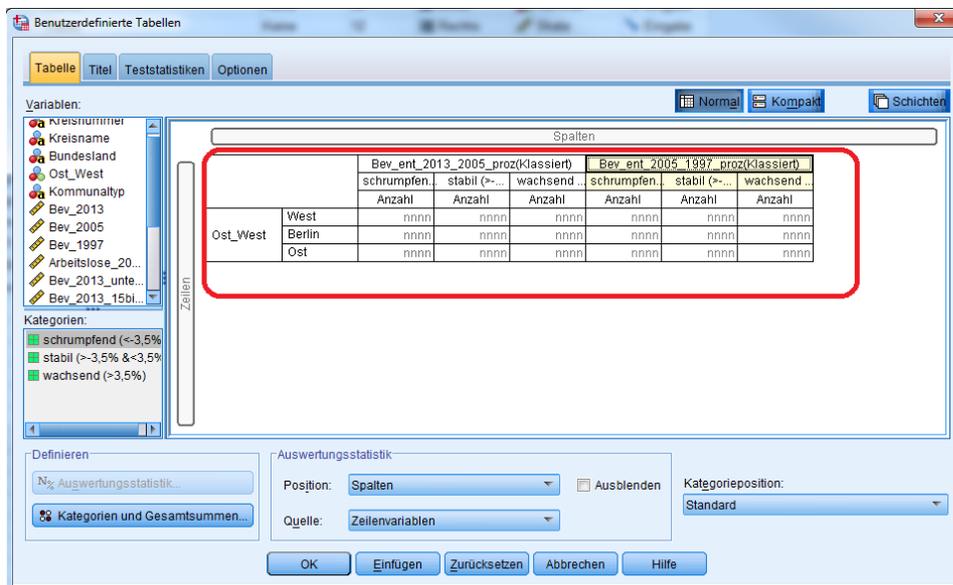
➔ **Benutzerdefinierte Tabellen**

		Bev_ent_2013_2005_proz(klassiert)		
		schrumpfen (<-3,5%)	stabil (>-3,5% &<3,5%)	wachsend (>3,5%)
		Anzahl	Anzahl	Anzahl
Ost_West	West	108	190	26
	Berlin	0	1	0
	Ost	53	13	4

5. **Beispiel 2: hinzufügen einer weiteren ordinalen oder nominalen Variablen zu einer Kreuztabelle.** Im Gegensatz zum Dialogfeld ‚Kreuztabelle‘ können hier mehrere Variablen direkt nebeneinander in einer Tabelle ausgezählt werden. Hierfür muss die zusätzliche Variable (Hier: die Bevölkerungsentwicklung von 1997 bis 2005, klassiert) rechts oder links an den Rand der bisherigen Spalten gezogen werden:



Ergebnis im Dialogfeld:

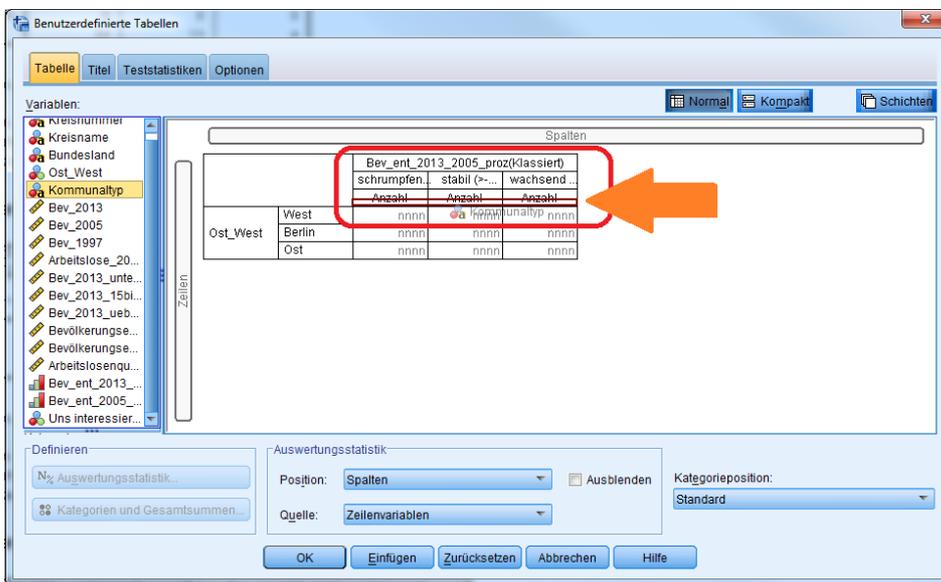


Ergebnis nach Klick auf OK:

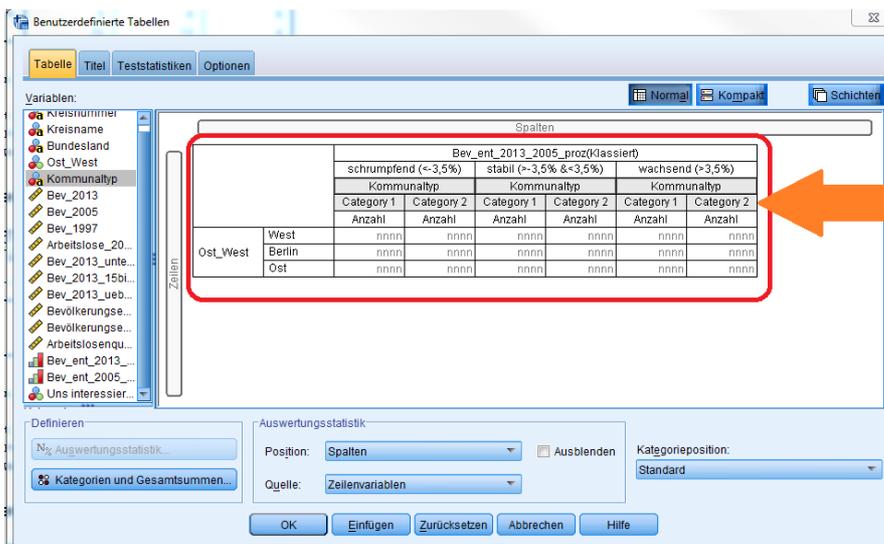
➔ Benutzerdefinierte Tabellen

		Bev_ent_2013_2005_proz(Klassiert)			Bev_ent_2005_1997_proz(Klassiert)		
		schrumpfend (<-3,5%)	stabil (>-3,5% &<3,5%)	wachsend (>3,5%)	schrumpfend (<-3,5%)	stabil (>-3,5% &<3,5%)	wachsend (>3,5%)
		Anzahl	Anzahl	Anzahl	Anzahl	Anzahl	Anzahl
Ost_West	West	108	190	26	26	200	98
	Berlin	0	1	0	0	1	0
	Ost	53	13	4	43	8	9

6. **Beispiel 3: Subunterteilung einer Kreuztabelle nach einer weiteren ordinalen oder nominalen Variablen.** Im Gegensatz zur Vorgehensweise eben können auch weitere ordinale oder nominale Variablen hinzugefügt werden, um bestehende Unterteilungen noch weiter zu unterteilen. Hierzu wird die neue Variable direkt unter die Spalten- bzw. Zeilenvariable gezogen.



Ergebnis im Dialogfeld:



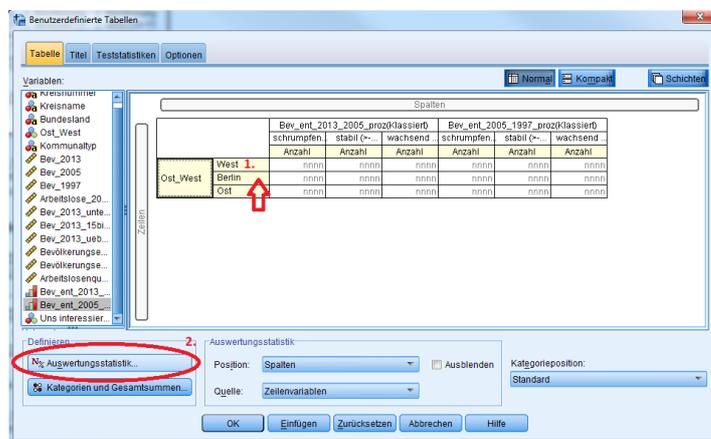
Ergebnis nach Klick auf OK:

➔ Benutzerdefinierte Tabellen

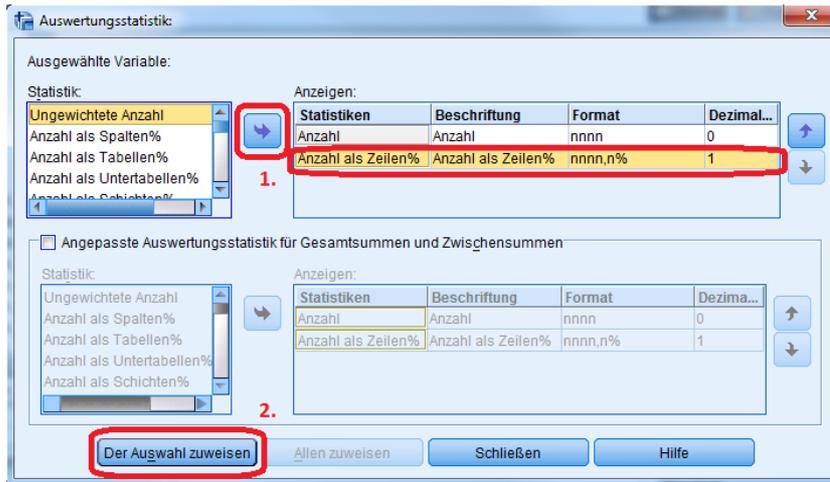
		Bev_ent_2013_2005_proz(Klassiert)								
		schrumpfend (<-3,5%)			stabil (>-3,5% &<3,5%)			wachsend (>3,5%)		
		Kommunaltyp			Kommunaltyp			Kommunaltyp		
		Kommunalver band besonderer Art	Kreisfreie Stadt	Landkreis	Kommunalver band besonderer Art	Kreisfreie Stadt	Landkreis	Kommunalver band besonderer Art	Kreisfreie Stadt	Landkreis
		Anzahl	Anzahl	Anzahl	Anzahl	Anzahl	Anzahl	Anzahl	Anzahl	Anzahl
Ost_West	West	1	21	86	1	49	140	0	18	8
	Berlin	0	0	0	0	1	0	0	0	0
	Ost	0	8	45	0	6	7	0	4	0

(Wird die neue Variable auf die bestehende Spalten, bzw. Zeilenvariable gezogen, wird die alte Variable ersetzt. Wird die neue Variable an den oberen Rand der bestehenden Variable gezogen, dient die neue Variable als Unterteilung oberhalb der bestehenden Variablen).

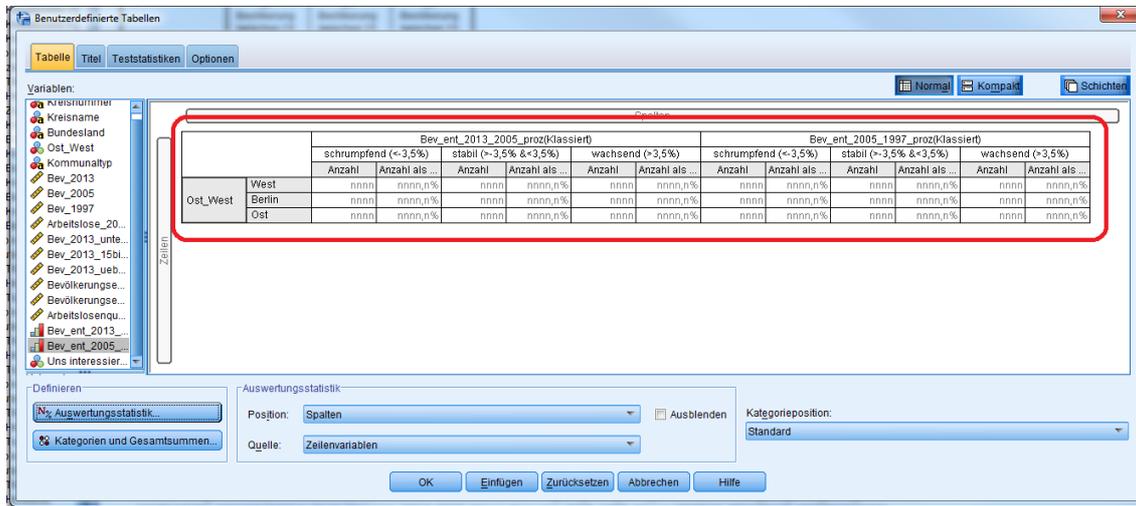
7. **Beispiel 2a: Anpassen der Statistik für eine Häufigkeitstabelle (Kombination ausschließlich ordinaler oder nominaler Variablen).** Sollen neben den Anzahlen noch weitere Statistiken wie beispielsweise Prozentzahlen ausgegeben werden, so geschieht dies über das Feld ‚Auswertungsstatistiken‘ (2.). Vorher muss jedoch noch die gesamte Tabelle durch Klick in das Kategorienfeld der Zeilenvariablen markiert werden (1.). Ausgegangen wird hier vom 2. Beispiel:



Im folgenden Dialogfeld wird im oberen Bereich die passende Statistik ausgewählt (hier: Anzahl als Zeilen %), und per Klick auf den Pfeil in die Auswahl übernommen (1.). Anschließend kann das Dialogfeld per Klick auf „Der Auswahl zuweisen“ geschlossen werden (2.)



Im Dialogfeld sind anschließend für beide Spaltenvariablen neben der Anzahl auch noch die Zeilenprozentfelder zu sehen:



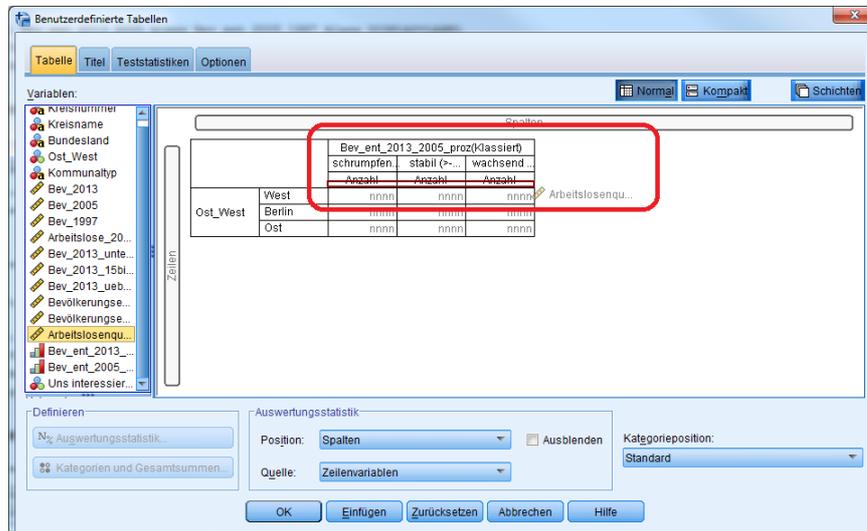
In der endgültigen Tabelle (nach Klick auf OK) sieht man, dass für beiden Variablen die Zeilenprozent getrennt berechnet werden. So sind beispielsweise im Zeitraum 2005-2013 58,6% aller Kommunen in Westdeutschland als ‚stabil‘ klassifiziert worden. Im Zeitraum 1997 bis 2005 sind 61,7% aller Kommunen in Westdeutschland als ‚stabil‘ klassifiziert worden:

➔ Benutzerdefinierte Tabellen

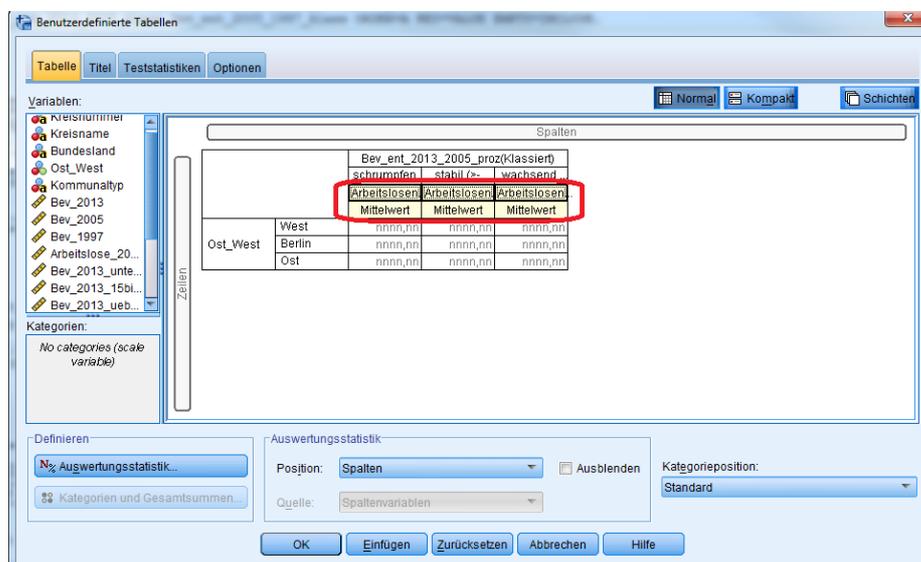
		Bev_ent_2013_2005_proz(Klassiert)						Bev_ent_2005_1997_proz(Klassiert)					
		schrumpfend (<-3,5%)		stabil (~-3,5% &<3,5%)		wachsend (>3,5%)		schrumpfend (<-3,5%)		stabil (~-3,5% &<3,5%)		wachsend (>3,5%)	
		Anzahl	Anzahl als Zeilen%	Anzahl	Anzahl als Zeilen%	Anzahl	Anzahl als Zeilen%	Anzahl	Anzahl als Zeilen%	Anzahl	Anzahl als Zeilen%	Anzahl	Anzahl als Zeilen%
Ost_West	West	108	33,3%	190	58,6%	26	8,0%	26	8,0%	200	61,7%	98	30,2%
	Berlin	0	0,0%	1	100,0%	0	0,0%	0	0,0%	1	100,0%	0	0,0%
	Ost	53	75,7%	13	18,6%	4	5,7%	43	71,7%	8	13,3%	9	15,0%

8. Beispiel 4: Hinzufügen einer metrischen Variablen

Wird statt einer nominalen oder ordinalen Variable eine metrische Variable hinzugefügt (hier: Arbeitslosenquote), wird standardmäßig deren Mittelwert ausgegeben, unterteilt nach den ausgewählten nominalen oder ordinalen Variablen.



Ergebnis im Dialogfeld:

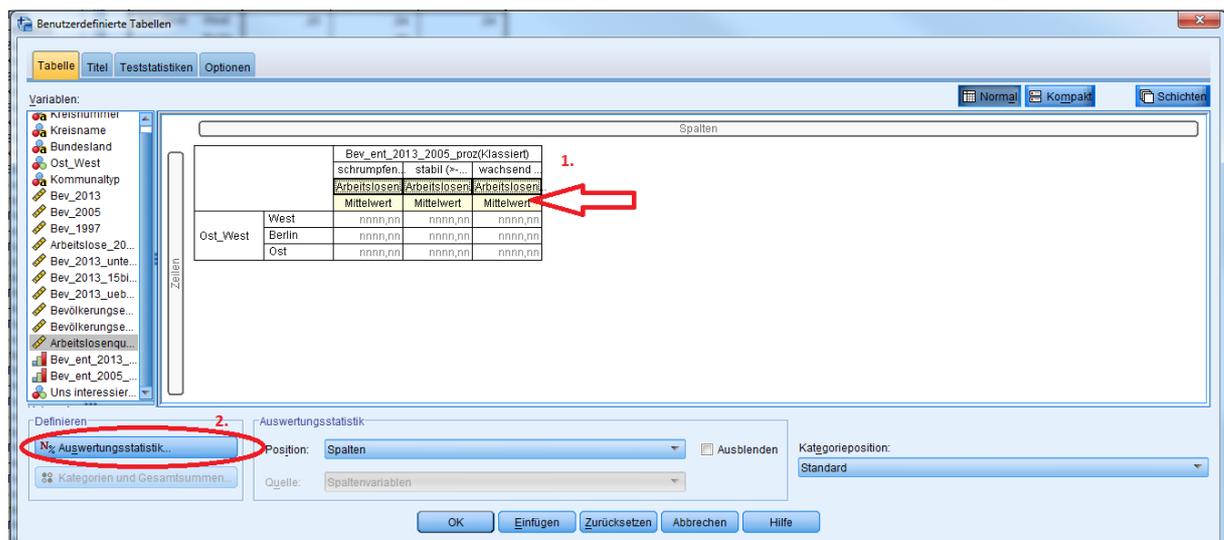


Ergebnis nach Klick auf OK (auch diese Tabelle kann und sollte noch per Doppelklick bearbeitet werden und es sollten beispielsweise bessere Spaltenüberschriften vergeben werden):

➔ **Benutzerdefinierte Tabellen**

		Bev_ent_2013_2005_proz(klassiert)		
		schrumpfend (<-3,5%)	stabil (>-3,5% &<3,5%)	wachsend (>3,5%)
		Arbeitslosenq uote 2013 (Anteil Arbeitslose an Bevölkerung zwischen 15 und 65 in %)	Arbeitslosenq uote 2013 (Anteil Arbeitslose an Bevölkerung zwischen 15 und 65 in %)	Arbeitslosenq uote 2013 (Anteil Arbeitslose an Bevölkerung zwischen 15 und 65 in %)
		Mittelwert	Mittelwert	Mittelwert
Ost_West	West	,05	,04	,04
	Berlin	.	,09	.
	Ost	,08	,08	,07

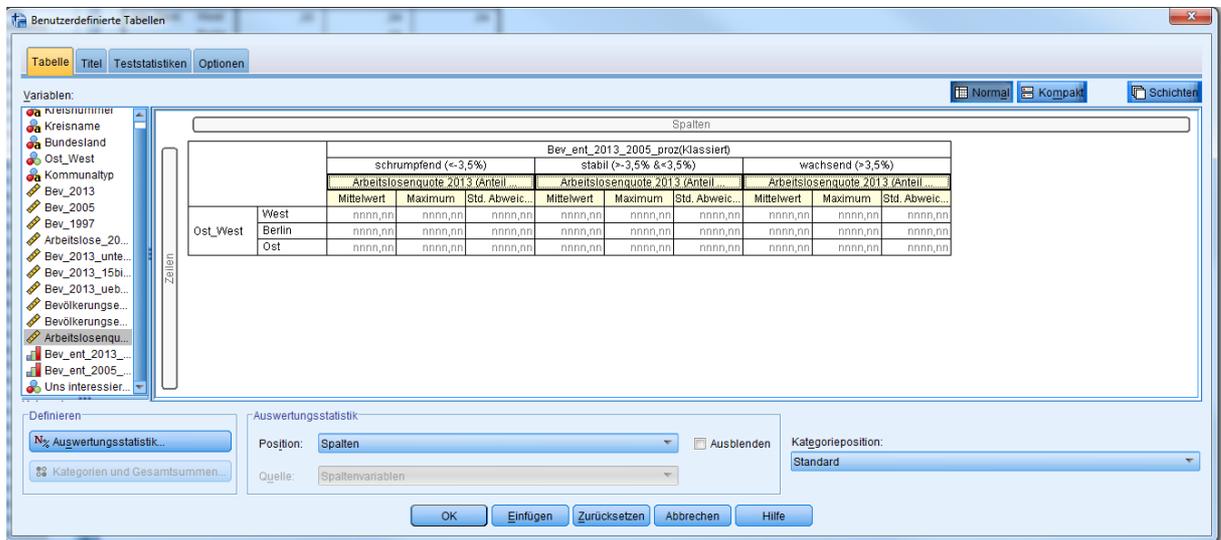
9. **Beispiel 4a: Anpassen der Statistik für eine Durchschnittstabelle (Kombination einer metrischen mit ordinalen oder nominalen Variablen).** Auch bei einer metrischen Variablen können die Auswertungsstatistiken entsprechend angepasst werden (ausgehend von Beispiel 4). Hier muss erst die metrische Variable markiert werden, indem in den entsprechenden Spaltenkopf geklickt wird (1). Anschließend kann das Dialogfeld „Auswertungsstatistiken“ aufgerufen werden (2):



Auch hier müssen die entsprechenden Statistiken ausgewählt werden (hier: Maximum und Standardabweichung) und durch Klick auf den Pfeil in die Auswahl übernommen werden (1.). Anschließend kann das Dialogfeld per Klick auf „Der Auswahl zuweisen“ geschlossen werden (2.)



Ergebnis im Dialogfeld:



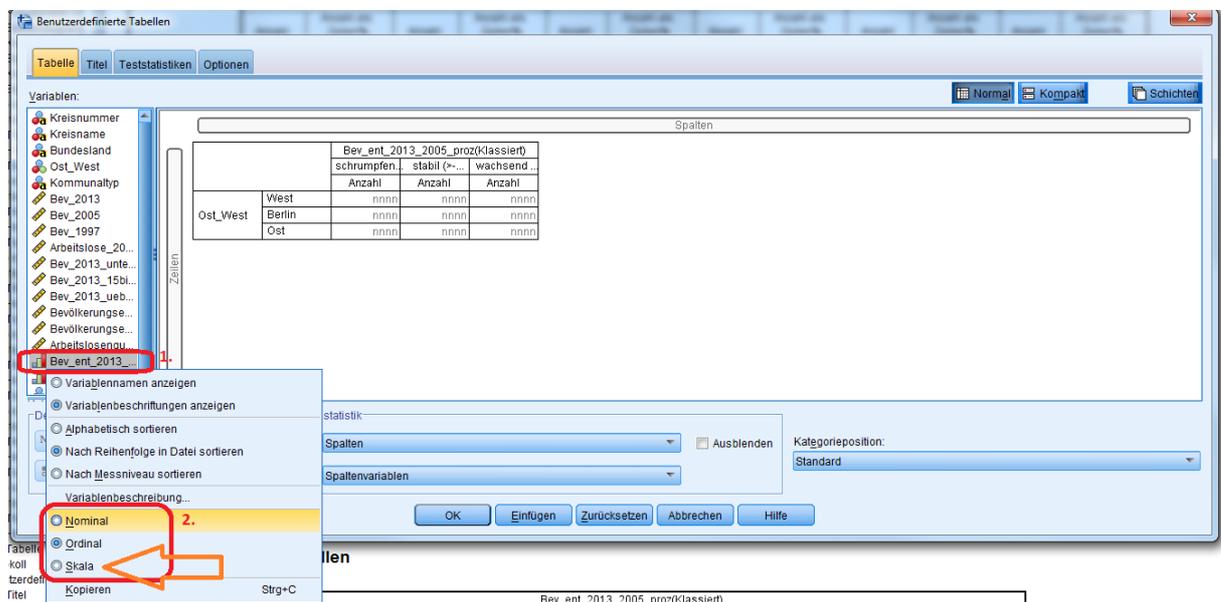
Ergebnis nach Klick auf OK:

➔ **Benutzerdefinierte Tabellen**

		Bev_ent_2013_2005_proz(Klassiert)								
		schrumpfend (<-3,5%)			stabil (>-3,5% &<3,5%)			wachsend (>3,5%)		
		Arbeitslosenquote 2013 (Anteil Arbeitslose an Bevölkerung zwischen 15 und 65 in %)			Arbeitslosenquote 2013 (Anteil Arbeitslose an Bevölkerung zwischen 15 und 65 in %)			Arbeitslosenquote 2013 (Anteil Arbeitslose an Bevölkerung zwischen 15 und 65 in %)		
		Mittelwert	Maximum	Standardabweichung	Mittelwert	Maximum	Standardabweichung	Mittelwert	Maximum	Standardabweichung
Ost_West	West	,05	,12	,02	,04	,10	,02	,04	,07	,01
	Berlin	.	.	.	,09	,09
	Ost	,08	,13	,02	,08	,09	,01	,07	,08	,01

10. Beispiel 5: Hinzufügen des Medians einer ordinalen Variablen als Auswertestatistik. Soll für eine ordinale Variable (wie z.B. die Klassierung der Bevölkerungsentwicklung nach Schrumpfung, Stabil, Wachstum) der Median hinzugefügt werden, so geht dies nicht ohne einen kleinen Umweg, da in der Standardeinstellung diese Statistik nicht für ordinale Variablen angezeigt wird.

Als Lösung kann das Skalenniveau der Variable vorübergehend auf ‚metrisch‘ geändert werden. Dies geschieht durch Rechtsklick auf die Variable (hier: Bev_ent_2013_2005_klass), bevor sie dem Diagramm hinzugefügt wird (1.). Hier kann das Messniveau auf ‚Skala‘ (=metrisch) geändert werden (2.). Diese Änderung gilt nur, solange das Dialogfeld „Benutzerdefinierte Tabelle“ geöffnet ist.



Anschließend kann die jetzt als metrisch klassifizierte Variable - wie hier gezeigt ggf. erneut - der Tabelle hinzugefügt werden:

Benutzerdefinierte Tabellen

Tabelle Titel Teststatistiken Optionen

Variablen: Normal Kompakt Schichten

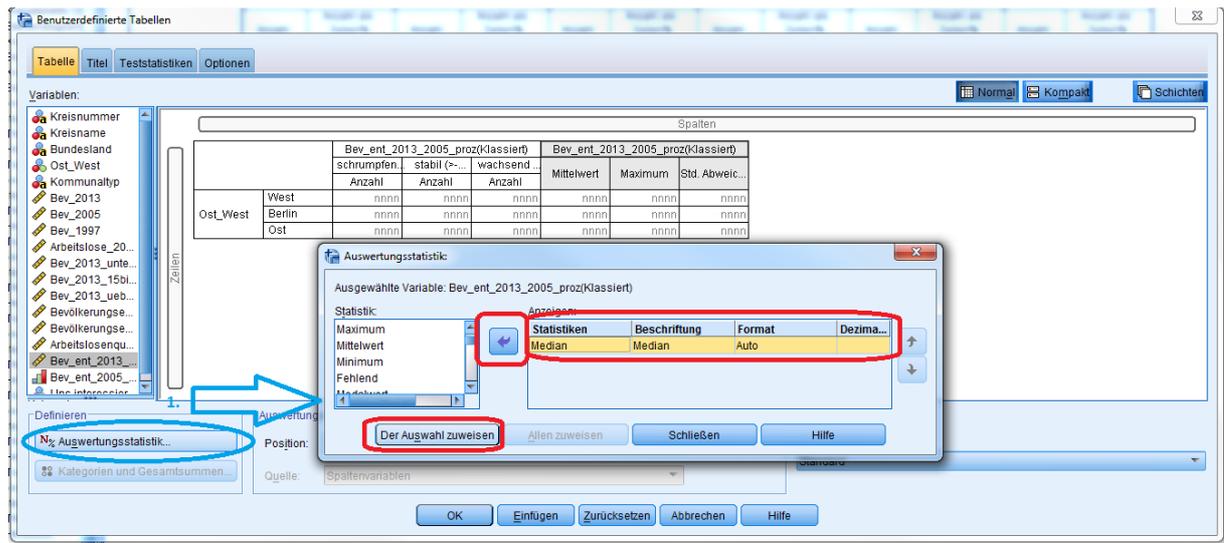
		Bev_ent_2013_2005_proz...	klassiert	
		schrumpfen	stabil (*-...)	wachsend...
		Anzahl	Anzahl	Anzahl
Ost_West	West	nnnn	nnnn	nnnn
	Berlin	nnnn	nnnn	nnnn
	Ost	nnnn	nnnn	nnnn

Definieren: % Auswertungsstatistik... Kategorien und Gesamtsummen...

Auswertungsstatistik: Position: Spalten Quelle: Spaltenvariablen Ausblenden Kategorieposition: Standard

OK Einfügen Zurücksetzen Abbrechen Hilfe

Jetzt können die Auswertungsstatistiken wie im letzten Beispiel beschrieben angepasst werden. Ggf. müssen zuerst alle momentan angezeigten Statistiken entfernt werden, indem sie in der rechten Box markiert und per Klick auf den Pfeil nach links entfernt werden. Anschließend kann als Statistik „Median“ ausgewählt werden. Dieser muss durch Klick auf den Pfeil in die Auswahl übernommen werden, anschließend kann das Dialogfeld per Klick auf „Der Auswahl zuweisen“ geschlossen werden:



Ergebnis nach Klick auf OK - für West und Berlin ist die 2. Kategorie der Median („stabil“), für Ost ist die 1. Kategorie der Median („schrumpfend“):

➔ **Benutzerdefinierte Tabellen**

		Bev_ent_2013_2005_proz(Klassiert)			Bev_ent_2013_2005_proz (Klassiert)
		schrumpfend (<-3,5%)	stabil (>-3,5% &<3,5%)	wachsend (>3,5%)	Median
		Anzahl	Anzahl	Anzahl	
Ost_West	West	108	190	26	2
	Berlin	0	1	0	2
	Ost	53	13	4	1

11. Export einer Tabelle und Import in Word.

Durch Rechtsklick auf die Tabelle im Eingabefenster und dann Klick auf ‚Kopieren‘ oder ‚Kopieren Spezial‘ kann die Tabelle in die Zwischenablage kopiert und anschließend z.B. in Powerpoint oder Word weiterverwendet werden.

		Bev_ent_2013_2005_proz(Klassiert)						Bev_ent_2005_1997_proz(Klassiert)					
		schrumpfend (<-3,5%)		stabil (>-3,5% &<3,5%)		wachsend (>3,5%)		schrumpfend (<-3,5%)		stabil (>-3,5% &<3,5%)		wachsend (>3,5%)	
		Anzahl	Anzahl als Zeilen%	Anzahl	Anzahl als Zeilen%	Anzahl	Anzahl als Zeilen%	Anzahl	Anzahl als Zeilen%	Anzahl	Anzahl als Zeilen%	Anzahl	Anzahl als Zeilen%
Ost_West	West	108	33,3%	190	58,6%	26	8,0%	26	8,0%	26	8,0%	26	30,2%
	Berlin	0	0,0%	1	100,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
	Ost	53	75,7%	13	18,6%	4	5,7%	43	71,7%	9	15,0%	9	15,0%

Nach Einfügen in Word (Strg+C, ‚Ursprüngliche Formatierung beibehalten‘) passiert es häufig, dass die Tabellen breiter sind als das Word-Dokument:

In diesem Fall kann durch Rechtsklick | ‚AutoAnpassen‘ | ‚Größe an Fenster anpassen‘ die Tabellengröße an die Seitenbreite angepasst werden.

Ergebnis:

		Bev_ent_2013_2005_proz(Klassiert)						Bev_ent_2005_1997_proz(Klassiert)					
		schrumpfend (<-3,5%)		stabil (>-3,5% &<3,5%)		wachsend (>3,5%)		schrumpfend (<-3,5%)		stabil (>-3,5% &<3,5%)		wachsend (>3,5%)	
		Anzahl	Anzahl als Zeilen%	Anzahl	Anzahl als Zeilen%	Anzahl	Anzahl als Zeilen%	Anzahl	Anzahl als Zeilen%	Anzahl	Anzahl als Zeilen%	Anzahl	Anzahl als Zeilen%
<u>Ost_West</u>	West	108	33,3%	190	58,6%	26	8,0%	26	8,0%	200	61,7%	98	30,2%
	Berlin	0	0,0%	1	100,0%	0	0,0%	0	0,0%	1	100,0%	0	0,0%
	Ost	53	75,7%	13	18,6%	4	5,7%	43	71,7%	8	13,3%	9	15,0%