

SPSS-Kurzanleitung (SPSS-Version 10.0)

Vorabinformationen

Übungsdateien (!!wichtig!!)	alle Übungsdateien, mit denen die Beispiele gerechnet werden können, befinden sich im Internet unter ftp://ftp.gwdg.de/pub/agrار-marketing/KonsVerh-Uebung/
-----------------------------	--

- | | |
|---|--|
| Verschiedene SPSS-Bausteine (Module) | <ul style="list-style-type: none"> • SPSS-Base (Datenmanagement, deskriptive Statistik, Regressions- Varianz-, Diskriminanz-, Cluster- und Faktorenanalyse) • SPSS-Professional Statistics (speziellere Regressionsverfahren, multidimensionale Skalierung) • SPSS-Advanced Statistics • SPSS-Tables (erweiterte Layoutmöglichkeiten bei der Tabellenerstellung) • SPSS-Categories (Verfahren zur Analyse kategorialer Daten, z.B. Korrespondenzanalyse) • SPSS-Conjoint (nur Conjoint-Analyse) • SPSS-CHAID (<u>C</u>hi-squared <u>A</u>utomatic <u>I</u>nteraction <u>D</u>etector zur Segmentierung) • SPSS-Data Entry (Fragebogengenerator für computergestützte Interviews) • SPSS-Exact Tests (Berechnung exakter Irrtumswahrscheinlichkeiten beim Vorliegen kleiner Fallzahlen bei nicht-parametrischen Tests, z.B. Chi-Quadrat-Test) • SPSS-Trends (Zeitreihenanalyse) • SPSS-AnswerTree (Weiterentwicklung von CHAID) • SPSS-Missing Value Analysis • SPSS-AMOS (Strukturgleichungsmodelle, pendant zu LISREL) |
|---|--|

SPSS-Dateneditor	<ul style="list-style-type: none"> • Dateneditor wird beim Aufrufen von SPSS gestartet • Die hier erzeugten und verwendeten Dateien enthalten die Datensätze in Tabellen(Spread-Sheet)-Form und werden mit der Endung *.sav abgespeichert • Funktionen im Dateneditor: Dateneingabe/Dateneinlesen, Datenbearbeitung, Statistische Berechnungen, Erzeugung von Tabellen und Graphiken
SPSS-Outputeditor	<ul style="list-style-type: none"> • Gibt die Ergebnisse der im Dateneditor in Auftrag gegebenen Aktionen automatisch aus (statistische Kennzahlen, Tabellen, Graphiken) • Ermöglicht Änderungen am Layout der Ergebnisse (Tabellen, Graphiken) aber keine Neuberechnung • vom Outputeditor aus können auch Statistiken und Graphiken aber keine Datentransformationen durchgeführt werden
SPSS-Syntaxeditor	<ul style="list-style-type: none"> • Ermöglicht manuelle (d.h. nicht menugesteuerte) Eingabe von SPSS-Befehlen • Spezielle statistische Verfahren (bzw. Optionen) können nur über den Syntaxeditor aktiviert werden • Bei häufiger Anwendung und guter Kenntnis der Syntax - effizienteres Arbeiten • Aufruf nicht automatisch sondern unter Datei/Neu/Syntax
SPSS-Skripteditor	<ul style="list-style-type: none"> • Der Skripteditor enthält die Programmiersprache Visual Basic, um vor allem Vorgänge der Outputgestaltung zu automatisieren • Möglichkeit des Einsatzes bzw. Modifikation mitgelieferter SPSS-Skripts ohne größere Programmierkenntnisse

Inhaltliche Aufgaben	SPSS-Befehle	Erläuterungen und Beispiele
1) Eingabe neuer Daten	Datei/Neu/Daten	
	Daten einfach in die Tabelle eintragen	<i>Bei diesem Verfahren vergibt SPSS automatisch Variablenamen und definiert diese (Skalenniveau, Spaltenbreite, etc). Besser ist es jedoch die Variablen vor der Dateneingabe manuell zu definieren</i>
2) Definieren von Variablen (<i>Format, Fehlende Werte, Label etc.</i>)	auf Spaltenkopf doppelklicken	<i>Zum Wechseln von der Variablenansicht zur Datenansicht: auf Spaltenkopf doppelklicken (bzw. Registrierkarte unten links auf der SPSS-Seite wählen); Zum Wechseln von der zur Datensatzansicht: auf die Variablenansicht Zeilenkopf doppelklicken</i>
	Name eintragen	<i>maximal 8 Zeichen, das erste Zeichen muss ein Buchstabe oder das @-Zeichen sein, einige SPSS-Schlüsselwörter dürfen nicht verwendet werden</i>
	Messniveau (=Skalenniveau) wählen	<i>im Zweifelsfall 'metrisch' wählen, da hiermit später mehr Berechnungen und statistische Analysen möglich sind</i>
	Variablentyp wählen	<i>in der Regel 'numerisch' oder 'string' (=Buchstaben oder Zahlen-Buchstaben-Kombinationen) wählen.</i>

Spaltenformat wählen	<i>Insbesondere. bei 'string'-Variablen die Variablenbreite (=Spaltenformat) erhöhen, da standardmäßig nur 8 Zeichen eingegeben werden können; wenn man nachträglich die Variablenbreite/das Spaltenformat ändert, werden die Zeichen, die nicht mehr passen für immer und ewig abgeschnitten</i>
Dezimalstellen wählen	<i>Es können mehr Dezimalstellen eingegeben werden als hier angegeben, angezeigt werden aber nicht mehr. Wenn die optische Spaltenbreite (siehe unten) zu klein gewählt ist, werden nicht alle Dezimalstellen angezeigt</i>
Label eingeben	<i>Label sind Beschriftungen, die die Variable näher beschreiben; wahlweise erscheinen Labels anstelle von Variablennamen in Graphiken, Tabellen oder Dialogfeldern zur Variablenauswahl</i>
Wertelabel eingeben	<i>auch für Werte können Labels ausgegeben werden (z.B. bei der Variablen 'Geschlecht: Wert =1, Wertelabel, weiblich - nach anklicken von 'Hinzufügen' wird nicht mehr 1 sondern 'weiblich' in Graphiken und Tabellen angezeigt); Auswählen ob Label oder Variablennamen angezeigt werden sollen: Bearbeiten/Optionen/Allgemein</i>
Fehlende Werte definieren	<i>es besteht die Möglichkeit z.B. verschiedene Fehlende Werte zu definieren, die von SPSS dann bei Berechnungen unberücksichtigt bleiben. Die Definition verschiedener fehlender Werte ist sinnvoll wenn man eine Differenzierung z.B. zwischen 'weiß nicht' 'Frage nicht beantwortet' 'Fragebogen nicht beantwortet' im Datensatz beibehalten will, SPSS aber alle 3 Varianten alle fehlenden Wert interpretieren soll.</i>
(optische) Spaltenbreite definieren	<i>Um die Spaltenbreite auf den Bildschirm zu definieren. Am Variableninhalt ändert sich dadurch nichts</i>

	Ausrichtung	<i>Links, rechts oder in der Mitte</i>
	Meßniveau	<i>Im Zweifelsfall ‚metrisch‘ wählen, da hiermit später mehr Berechnungen und statistische Analysen möglich sind</i>
3) Import von Excel (und anderen Dateiformaten) in den Dateneditor	Datei/Öffnen/Daten	<i>Der direkte Import von Access-Daten ist etwas komplizierter, Hier empfiehlt sich, die Access-Daten als Excel-Datei abzuspeichern und dann einzulesen Zum Üben kann die Datei 'Datensatz_Beckers.xls' verwendet werden. Diese Datei ist Kennwortgeschützt - das Kennwort wird für Studierende auf Wunsch bekannt gegeben. Bevor die Datei in SPSS eingelesen werden kann, ist sie mit Excel zu öffnen und neu - ohne Schreib/Leseschutz - abzuspeichern. Viele der folgenden Beispiele beziehen sich auf diese Datei.</i>
	Im Dialogfeld ‚Dateityp‘ das Dateiformat wählen und ‚Öffnen‘ klicken	
	Das Feld ‚Variablennamen aus der ersten Datenzeile lesen‘ aktivieren	<i>Wenn das Feld nicht aktiviert ist, werden die Variablennamen (d.h. die Spaltenüberschriften aus Excel) nicht importiert</i>
	Gegebenenfalls Variablendefinition überprüfen und ergänzen (siehe Punkt 2)	<i>Beim Import weißt SPSS bereits Variablentyp und Messniveau zu.</i>

4) Berechnung einer neuen Variablen	Transformieren/Berechnen	
	Zielvariable ins Dialogfeld eintragen	<i>Hier kann entweder a) ein neuer Name oder b) ein bereits bestehender Name eingegeben werden. Bei b) wird die ursprüngliche Variable überschrieben.</i>
	Ins Feld 'Numerischer Ausdruck' die Berechnungsformel eingeben	<i>Hier können z.B. zwei Variablen eingefügt und addiert werden (becker + albi); oder umfangreichere Mittelwertberechnungen durchgeführt werden: MEAN(beck23a,beck23b,...beck23n); Informationen über die Funktionen erhält man mit der rechten Maustaste; interessant ist noch die Funktion Range: RANGE(income27,0,3) - versieht alle Datensätze mit einer 1 bei denen als Einkommen die Kategorien 0 bis 3 angekreuzt worden sind.</i>
	gegebenenfalls eine 'Falls-Anweisung' angeben.	<i>Durch eine Falls-Anweisung wird eine Bedingung formuliert unter der die Berechnung durchgeführt wird. Mehrere Bedingungen können durch die Operatoren 'or' bzw. 'and' verknüpft werden. Führe die Berechnung durch falls 'income27 = 3 or income27 = 4)</i>
5) Umkodierung von Werten	Transformieren/Umkodieren/in andere Variablen	<i>Wenn die Einkommenseinteilung in 7 Klassen zu fein ist und man nur zwischen arm und reich differenzieren möchte kann man die Variable 'income27' umkodieren in eine andere Variable (aber auch in dieselbe und damit die Originalvariable ersetzen).</i>
	umzukodierende Variable mit dem Pfeil in das Dialogfeld schieben	

	Name und gegebenenfalls Label der neuen Variablen angeben und den Button 'ändern' wählen	
	alte und neue Werte eintragen und 'Hinzufügen'	<i>z.B. alter Wert/Bereich: 1 bis 2; neuer Wert: 1 - 'Hinzufügen' anschließend alter Wert/Bereich 3 bis 4 ; neuer Wert: 2 - 'Hinzufügen'</i>
	wenn alle Werte umkodiert sind mit 'weiter' abschicken	
6) Selektion von ausgewählten Datensätzen (auch zum Entdecken unplausibler Antworten)	Daten/Fälle auswählen	<i>Z.B. Auswahl nur der weiblichen Befragten, um damit getrennt weiterzurechnen</i>
	'Falls Bedingung zutrifft' ankreuzen	<i>optional kann auch aus dem vorhandenen Datensatz eine Zufallsstichprobe gezogen werden</i>
	unter 'Falls...' die Bedingung ins das Dialogfeld eingeben	<i>z.B. sex = 1, im folgenden werden nur Datensätze (=Zeilen) bei denen die Variable 'sex' gleich 1 ist in weitere Berechnungen einbezogen; nicht berücksichtigte Datensätze erkennt man im Dateneditor an einer durchgestrichenen Nummer</i>

7) Mehrfachantworten kodieren	Zwischen der 'Methode multipler Dichotomien' und der 'Methode multipler Kategorien' auswählen	<i>empfohlen wird die Methode der multiple Dichotomien: aus allen ankreuzbaren Items wird je eine dichotome Variable erstellt (wurden die Befragten gebeten, eine oder mehrere der Sportarten Fussball, Handball und Volleyball anzukreuzen, werden daraus im Datenblatt 3 ja-nein-Variablen: Fussball ja/nein, Handball ja/nein, Volleyball ja/nein. Mit dieser Methode lassen sich später mehr statistische Berechnungen durchführen als mit den multiplen Kategorien</i>
8) Mehrfachnennungen auswerten (bei multiplen Dichotomien)	Übungsdatei 'Mehrfachantworten.sav' öffnen	<i>Es handelt sich hierbei um eine Frage nach den durchgeführten Freizeitaktivitäten an einem Erholungsort. Die Antworten (z.B. Spazieren gehen, Wandern, Radtouren, ...) sind als dichotome Variablen kodiert worden.</i>
	Analysieren/Mehrfachantworten/Sets definieren	<i>Zwischenschritt: hier werden die dichotomisierten Variablen wieder zu einer Gruppe zusammengefasst, so dass eine übersichtlichere Darstellung in Tabellen und Kreuztabellen möglich ist.</i>
	Alle gewünschten Variablen ins Set kopieren und die Option 'Dichotomien' wählen	
	Einen neuen Namen für das Set definieren und gegebenenfalls einen Label festlegen	

gezählten Wert festlegen und das neue Variablenset 'Hinzufügen'	<i>Hier handelt es sich i.d.R. um eine 1, wenn das Ankreuzen einer Aktivität mit ja und das Nicht-Ankreuzen einer Aktivität mit nein kodiert ist.</i>
gegebenenfalls weitere Sets definieren oder das Dialogfeld schließen	
Analysieren/Mehrfachantworten/Häufigkeiten	<i>hiermit kann eine Häufigkeitstabelle erstellt werden, in der alle dichotomen Variablen in einer Tabelle zusammengefasst werden</i>
Testvariable auswählen und ok. klicken	<i>Die Häufigkeitstabelle wird im Output-Editor angezeigt</i>
Analysieren/Mehrfachantworten/Kreuztabellen	
Zeilen und Spaltenvariable für die Kreuztabelle auswählen	
Bereich definieren	<i>ist obligatorisch: wenn z.B. eine Kreuztabelle über 9 Einkommensgruppen gebildet werden soll, die von 1 bis 9 kodiert sind, ist hier der Bereich 1-9 einzugeben. Es können aber auch weniger Einkommensgruppen in der Tabelle erscheinen.</i>
9) Tabellen und Graphiken	<i>hier sei das Buch von Brosius (1999) 'SPSS 8 - professionelle Statistik unter Windows' empfohlen (insbesondere das Kapitel zu Pivottabellen).</i>

Marktsegmentierung mit SPSS

<p>10) Marktsegmentierung mithilfe von Kreuztabellen (2 Variablen)</p>	<p>Analysieren/Deskriptive Statistiken/Kreuztabellen</p>	
	<p>die 'zu kreuzenden' Variablen in das Zeilen und Spaltenfeld schieben</p>	<p><i>hierbei spielt es keine Rolle was Spalte und was Zeile ist; die folgenden Beispiele beziehen sich wieder auf den 'Datensatz_Beckers.xls'; z.B. die Variablen 'kauf16' und 'alter' womit die Kaufwahrscheinlichkeit für Aktivfrühstück und die Einordnung der Befragten in Altersklassen gemeint ist. (Tipp: Werden mehrere Variablen gleichzeitig in ein Feld geschoben, werden mehrere Kreuztabellen untereinander ausgewiesen)</i></p>
	<p>unter 'Statistik' die Option 'Chi-Quadrat' wählen anschließend 'weiter' klicken</p>	
	<p>unter 'Zellen' die Optionen 'beobachtet' und 'erwartet' anklicken - anschließend 'weiter' klicken</p>	

	nach dem Klicken von 'ok' wird die Kreuztabelle im Outputeditor angezeigt.	<i>Problem: 10 Zellen (50%) haben eine erwartete Häufigkeit kleiner 5. Der Chi-Quadrat-Test ist daher nicht gültig. Um das zu verhindern wird im nächsten Schritt die Variable 'kauf16' dichotomisiert.</i>
	Transformieren/Umkodieren/in andere Variablen	<i>siehe Punkt 5) 'Umkodierung von Variablen; z.B. Umwandeln von 'kauf16' in 'kaufdi' mit dem Label 'Kaufwahrscheinlichkeit dichotomisiert'. Hierzu den Bereich -2 bis 0 in 1 umkodieren und den Bereich 1 bis 2 in 2 umkodieren.</i>
11) Marktsegmentierung mithilfe von Kreuztabellen (3 Variablen)	Analysieren/Deskriptive Statistiken/Kreuztabellen	
	die ersten 'zu kreuzenden' Variablen in das Zeilen und Spaltenfeld schieben	<i>z.B. 'kauf16' und 'alter'</i>
	eine dritte Variable in das Dialogfeld 'Schicht 1 von 1' schieben	<i>z.B. sex</i>

nach dem Klicken von 'ok' wird die Kreuztabelle im Outputeditor angezeigt.	<p><i>Die Kreuztabelle kann nach dem Doppelklicken bearbeitet werden. Sollen z.B. Zeilen und Spalten vertauscht werden, muss mit dem Menübefehl 'Pivot/Pivotleisten' aktiviert werden (näheres in SPSS-Büchern).</i></p> <p><i>Anmerkung: Der Chi-Quadrat-test kann aber nicht für alle drei Variablen gleichzeitig durchgeführt werden, so dass ein Wert für den weiblichen und ein Teil für den männlichen Teil der Befragten ausgewiesen wird. Um dies zu umgehen, werden im folgenden zwei Variablen zu einer zusammengefasst (vgl. Lebenszykluskonstrukt oder Preisindex)</i></p>
Transformieren/Berechnen	<i>Erstellen eines Index aus den Variablen 'alter' und 'sex'</i>
Zielvariable benennen	<i>z.B. 'Index'</i>
erste Falls-Anweisung eintragen	<i>z.B. (alter = 1 alter = 2) & sex = 1</i>
auf 'weiter' klicken und im Feld 'Numerischer Ausdruck' die gewünschte Kodierung eintragen - anschließend 'OK' klicken.	<i>hier darf nur eine einfache Zahl (z.B. die 1) stehen. Verbal ausgedrückt passiert hier folgendes: Weise allen Befragten, die zwischen 20 und 29 Jahre oder die zwischen 30 und 39 Jahre alt und außerdem weiblich sind eine 1 zu.</i>
zweite Falls-Anweisung eintragen und wieder 'Numerischen Ausdruck' eintragen - usw.	<i>z.B. (alter = 3 alter = 4) & sex = 1</i> <i>Numerischer Ausdruck: 2</i>

	Analysieren/Deskriptive Statistiken/Kreuztabellen	<p>nach der Indexbildung kann mit der neuen Variablen 'Index' eine Kreuztabellierung (z.B. mit Kauf16) durchgeführt werden.</p> <p>Vorteil der tabellarischen Segmentierung: Rechenprozess kann leicht erläutert werden</p> <p>Problem1: Es lassen sich auf diese Weise keine intervallskalierten Variablen klassifizieren</p> <p>- Ausweg bei intervallskalierten Variablen: 'Transformieren/Variablen kategorisieren'</p> <p>Problem2: Die Segmentierung mit mehreren Variablen ist sehr mühsam</p> <p>Problem3: Die Gruppen/Cluster werden a priori festgelegt (Clustergröße ist nicht flexibel)</p> <p>Alle Probleme lassen sich mit der unter 13) behandelten Clusteranalyse lösen.</p>
12)	Analysieren/Mittelwerte vergleichen/Mittelwerte	<p>Angenommen es liegt eine Indexvariable vor (vgl. Punkt 11; z.B. die Variable 'index') und es soll festgestellt werden, ob sich die durch den Index gebildeten Cluster in der Einstellung zu einer Marke (z.B. Variable 'becker') unterscheiden</p>
	die metrische Variable in das Feld 'Abhängige Variablen' schieben	hier 'becker'
	die klassifizierende Variable in das Feld 'Unabhängige Variablen' schieben	hier 'index'
	unter Optionen im Feld Zellenstatistik nur 'Mittelwert' eintragen	das erhöht die Übersicht bei der Dateninterpretation

mit 'weiter' und 'OK' die Berechnung durchführen	<i>das Ergebnis wird im Outputeditor angezeigt</i>
Transformieren/Berechnen	<i>Die folgenden Schritte sind notwendig, um zu Überprüfen, ob die Mittelwerte der Gruppen von Gesamtmittelwert statistisch signifikant abweichen. Hierzu wird für jede der Gruppen eine neue Variablen gebildet. Anschließend wird für jede der neuen Variablen ein t-Test durchgeführt.</i>
a) Zielvariable festlegen	<i>z.B. 'becker_1'</i>
b) Falls-Anweisung eingeben	<i>z.B. 'index = 1'</i>
c) In des Feld die metrische Ausgangsvariable eintragen	<i>hier 'becker' Diese Schritte bewirken folgendes. In die Variable becker_1 wird nur die gemittelte Einstellung zu beckers bester eingetragen, falls die befragte Person weiblich und zwischen 20 und 39 Jahre ist.</i>
Schritte a) bis c) für alle Gruppen wiederholen	<i>Der Index enthält in diesem Fall vier Gruppen (weiblich/jung, weiblich/alt, männlich/jung, männlich/alt), so dass vier neue Variablen entstehen</i>
Analysieren/Mittelwerte vergleichen/T-Test bei einer Stichprobe	
Die neu gebildeten Variablen in das Feld 'Testvariablen' schieben	

	<p>Unter 'Testwert' den Gesamtmittelwert der metrischen Variablen eintragen und 'OK' klicken</p>	<p><i>Der Gesamtmittelwert ist bereits oben berechnet worden (5. Tabellenzeile unter diesem Punkt)</i></p> <p><i>Das Ergebnis des t-Tests wird im Outputeditor angezeigt.</i></p> <p><i>TIPP: die Ergebnisse lassen sich anschaulich in Form eines Polaritätenprofils darstellen. Vergleiche hierzu Herrmann und Homburg (1999): 'Marktforschung', S. 752.</i></p>
<p>13) Marktsegmentierung mithilfe der Clusteranalyse</p>	<p>Analysieren/Klassifizieren/hierarchische Cluster</p>	<p><i>Die theoretischen Grundlagen der Clusteranalyse werden bei Brosius (1999) 'SPSS 8 - professionelle Statistik unter Windows' und bei Backhaus (1996) 'Multivariate Analysemethoden - Eine anwendungsorientierte Einführung' erläutert.</i></p> <p><i>Das hierarchische Clusterverfahren kann bei einer größeren Anzahl an Beobachtungen nicht mehr durchgeführt werden. Ab ca. 200 Beobachtungen ist daher eine Clusterzentrenanalyse durchzuführen (vgl. Brosius 1999:721ff).</i></p>
	<p>diejenigen Variablen in das Feld 'Variablen' schieben, anhand derer eine Clusteranalyse durchgeführt werden soll</p>	<p><i>Die Variablenauswahl muss anhand von inhaltlichen Überlegungen getroffen werden. Nominalskalierte Variablen sind für eine Clusteranalyse ungeeignet (Ausnahme: dichotome Variablen)</i></p>
	<p>gegebenenfalls eine Variable ins Feld 'Fallbeschriftung' schieben</p>	<p><i>Wenn man z.B. die 20 wichtigsten Kunden gruppieren will, dann bietet sich als Fallbeschriftung der Kundename an. Im Output wird dann beispielsweise nicht der Fall 12 sondern Kunde 'Meyer' ausgewiesen, was die Interpretation der Ergebnisse erleichtert.</i></p>
	<p>Fälle oder Variablen auswählen</p>	<p><i>üblicherweise werden Datensätze (Fälle) gruppiert, möglich ist aber auch eine Clusterbildung über Variablen.</i></p>

<p>Unter 'Anzeigen' die Option 'Statistik' (gegebenenfalls auch 'Diagramme') ankreuzen</p>	<p><i>Nur wenn diese Felder angekreuzt sind, können unten weitere Statistik./Diagramm...-Optionen gewählt werden.</i></p> <p><i>Die Diagrammfunktion sollte deaktiviert werden, weil hier gegenüber 'Statistik' keine weiteren Erkenntnisse ausgegeben werden und der Berechnungsprozess etwas verlangsamt wird. Sie bietet sich allenfalls für didaktische Zwecke an, um den Clusterbildungsprozess graphisch zu veranschaulichen.</i></p>
<p>unter 'Statistik...' gewünschte Optionen wählen und anschließend 'weiter' klicken</p>	<p>Zuordnungsübersicht: <i>die Clusterbildungsschritte werden hier in Form einer Agglomerationstabelle aufgelistet. An dieser Zuordnungsübersicht (siehe Output der Clusteranalyse) ist vor allem die Spalte 'Koeffizienten' wichtig. Der Koeffizient drückt aus, welcher Abstand zwischen den beiden Clustern besteht, die auf dieser Stufe zusammengefasst werden. I.d.R. werden bei den letzten Fusionierungsschritten immer unähnlichere Gruppen zusammengefasst. Trägt man die Koeffizienten an der Y-Achse und die Anzahl der Cluster auf der X-Achse in ein Diagramm ein, so erkennt man bei vielen Clusteranalysen einen Knick im Graphen, der auf die optimale Anzahl der Cluster hindeutet (Elbow-Kriterium - vgl. Backhaus, 1996, S. 308).</i></p> <p>Distanz-Matrix: <i>im allgemeinen nicht notwendig, verlangsamt die Outputdarstellung</i></p> <p>Cluster-Zugehörigkeit: <i>Im Output wird jeder Datensatz einem der gebildeten Cluster zugeordnet. Daher muss entweder eine genaue Anzahl der Cluster oder zumindest ein Bereich angegeben werden. In der Praxis wird hier allerdings häufig 'keine' gewählt, weil man sich die Zuordnung der Datensätze unter 'speichern...' (siehe unten) auch direkt in die SPSS-Datendatei (z.B. Datensatz_Beckers.sav) eintragen lassen kann.</i></p>

<p>unter 'Methode...' gewünschte Optionen wählen und anschließend 'weiter' klicken</p>	<p>Clustermethode: <i>sehr verbreitet ist die Ward-Methode, die dahin tendiert, etwa gleich große Cluster zu bilden. Die Vor- und Nachteile verschiedener Methoden finden sich bei Backhaus (1996:297ff).</i></p> <p>Maß: <i>das gewünschte Distanzmaß kann hier in Abhängigkeit von der Variablenskalierung gewählt werden. In der Praxis werden häufig ordinalskalierte oder sogar binäre Variablen gleichbehandelt (vgl. hierzu auch die Beispielkalkulation in Brosius, 1999:695f). Am einfachsten kann daher im Bereich Intervall 'Quadrierter Euklidischer Abstand' gewählt werden. Eine ausführlichere Darstellung der Distanzmaße findet sich bei Backhaus (1996) und Brosius (1999).</i></p> <p>Standardisierung: <i>Wenn sich die Dimensionen der verwendeten Variablen stark unterscheiden, sollten die Werte standardisiert werden. Weit verbreitet ist die Transformation in Z-Werte (Zu Vor- und Nachteilen der Z-Transformation und weiteren Standardisierungen siehe Brosius, 1999:715f).</i></p>
--	--

<p>unter 'Speichern...' gewünschte Optionen wählen und anschließend 'weiter' klicken</p>	<p><i>Unter 'einzelne Lösung' wird die Anzahl der Cluster festgelegt. Nach dem Ausführen der Clusteranalyse wird jeder Datensatz einem Cluster zugeordnet. Diese Information wird in die SPSS-Datendatei (z.B. Datensatz_Beckers.sav) eingetragen. Mit der so entstandenen Klassifizierungsvariable können weitere Analyse (z.B. Kreuztabellierung) durchgeführt werden.</i></p> <p><i>Es ist zu empfehlen die gesamte Clusteranalyse zunächst ohne die Festlegung der Cluster-Zugehörigkeit zu rechnen und mithilfe der Zuordnungsübersicht (Elbow-Kriterium) eine sinnvolle Anzahl der Cluster zu bestimmen. Bei einer zweiten Rechnung sollte dann die so ermittelte Clusterzahl eingetragen werden.</i></p>
--	---