

Sommersemester 2009 Statistik mit SPSS



Methodenzentrum Sozialwissenschaften

Göttingen, 08.05.2009

Dozentin: Esther Ochoa Fernández



Überblick

- 1. Arbeitsschritte bei der Datenanalyse
- 2. Allgemeine Regeln bei den SPSS-Syntax-Anweisungen, SPSS-Hilfe
- 3. Datenbeschreibungs- und Datentransformationsbefehle



1. Arbeitsschritte bei der Datenanalyse



1. Arbeitsschritte bei der Datenanalyse

Datenmanagement

(Einlesen von Daten, Teilen von Datensätzen, Hinzufügen von Variablen oder Fällen etc.)

Datentransformation

(Bereinigung der Daten, umkodieren oder Kategorisieren von Variablen etc.)

Datenanalyse

(Tabellen, Grafiken, deskriptive Statistiken Regressions-, oder Faktorenanalyse etc.)

Göttingen, 08.05.2009

Statistik mit SPSS



2. Allgemeine Regeln bei SPSS-Syntax-Anweisungen



2. Regeln bei SPSS-Syntax-Anweisungen

- ➔ Jeder SPSS-Befehl beginnt mit einem Schlüsselwort (z.B. FREQUENCIES oder GRAPH).
- ➔ Jeder Befehl beginnt mit einer neuen Zeile und endet mit einem Punkt.
- Befehle und Unterbefehle dürfen (meistens) mit den ersten drei Buchstaben abgekürzt werden (z.B. fre, var lab, val lab, cor, reg).
- → Unterbefehle werden mit einem / abgetrennt.



2. Regeln bei SPSS-Syntax-Anweisungen

- → Keine Unterscheidung zwischen Groß- und Kleinschreibung.
- Variablennamen (z.B. v15 oder vsex) müssen ausgeschrieben werden (keine Sonderzeichen, Leerzeichen).
- Zum Nachlesen: Wittenberg/Cramer-Skript, S. 31 ff; Brosius, Kapitel 08.
- → Syntax-Hilfe:

- [/=]
- → SPSS-Hilfe: Syntax-Guide



Syntax-Hilfe

Beispiele 3. Sitzung - SPSS Syntax Editor	
File Edit View Data Transform Analyze Graphs Utilities Run Window Help	
≝∎⊜ ¤ ∽ ■ ‱ № № № № № !	
****Beispiele 3. Sitzung.	
fre v3 to v6/stat.	
SPSS-Syntax-Hi	lfe
	▼
PSPSS Processor is ready	



Base-System		×
Ausblenden Zurück	Vorwärts Abbrechen Aktualisieren Schriftart <u>O</u> ptionen	
nhalt I <u>n</u> dex <u>S</u> u	FREQUENCIES	ter 🗧
🗉 🔶 FILE 🔼		
🗉 🔷 FILE	PREQUENCIES VARIABLES=Variist [Variist]	
🗉 🔶 FILE	[/FORMAT= [{NOTABLE }] [{ AVALUE** }]	
🗉 🔷 FILTI	$\{LIMIT(n)\}$ $\{DVALUE\}$	
🗉 🗣 FINIS	{DFREQ }	
■ ♥ FIT		
■ ♥ FLIP	[/MISSING=INCLUDE]	
	[/BARCHART=[MINIMUM(n)] [MAXIMUM(n)] [{ FREQ (n) }]]	
	$\{PERCENT(n)\}$	
	{PERCENT} {NOMISSING}	
B ♥ GET		
∎ ♥GET	[/HISTOGRAM=[MINIMUM(n)] [MAXIMUM(n)] [{FREQ(n) }] [{NONORMAL}]]	
■ ♥ GET	(NORMAL)	
	[/GROUPED=varlist [{(width) }]]	
	{(boundary list)}	
	[/NTILES=n]	
B SIGRA	[/PERCENTILES=value list]	
B SIMPC	[/STATISTICS=[DEFAULT] [MEAN] [STDDEV] [MINIMUM] [MAXIMUM]	
■ ♦ INCL	[SEMEAN] [VARIANCE] [SKEWNESS] [SESKEW] [RANGE]	
	[MODE] [KURTOSIS] [SEKURT] [MEDIAN] [SUM] [ALL]	
🗉 🔷 INPU		
🗉 🐟 INSE	[/ORDER=[{ANALYSIS}] [{VARIABLE}]	
🗉 🐟 KEYE		
🗉 🚸 LEAV 🚩	** Default if subcommand is omitted or specified without keyword.	
	Dieser Befehl liest den aktiven Datensatz und löst die Ausführung aller offenen Befehle aus. Unter Command	~
ttingen,		
05.2009	Statistik mit SPSS	



SPSS-Hilfe: Syntax-Guide

*ZA3762 allbus2004.SAV [DatenSet2] - SPSS Daten-Editor						
Datei	Bearbeite	n Ansicht	Daten Transform	nieren Analysieren	Grafiken Extras Fenster	Hilfe
🗁 🔒 🚔	📴 🔶 🕈	🖢 🕼 🦛 🖷 🖬	i 🗄 🤹 🖪 🐨 🥥 🌢			Themen
	Name	Тур	Spaltenformat	Dezimalstellen	Variablenlabel Wertelabels Fe	Lernprogramm
1	v1	Numerisch	4	0	STUDIENNUM (3762, ALLBUS K	Fallstudien
2	v2	Numerisch	6	0	IDENTIFIKATIO Kein Ko	Statistics Coach
3	v3	Numerisch	1	0	ERHEBUNGSG {1, ALTE BUND K	Command Syntax Reference
4	v4	Numerisch	1	0	DEUTSCHE ST {1, JA} 7	SPSS Developer Control
5	v5	Numerisch	1	0	WIRTSCHAFTS {1, SEHR GUT}. 7	All stilles
6	v6	Numerisch	1	0	WIRTSCHAFTS {1, SEHR GUT}. 7	Algorithms
7	v7	Numerisch	1	0	WIRTSCHAFTS {1, WESENTLI 7	SPSS-Homepage
8	v8	Numerisch	1	0	WIRTSCHAFTS {1, WESENTLI 7	Info
9	v9	Numerisch	1	0	FREIZEIT: BUE {1, TAEGLICH}. 7	Lizenzautorisierung
10	v10	Numerisch	1	0	FREIZEIT: ZEIT {1, TAEGLICH}. 7	Nach Updates überprüfen
11	v11	Numerisch	1	0	FREIZEIT: LP, {1, TAEGLICH}. 7	Produkt registrieren
12	v12	Numerisch	1	0	FREIZEIT: VID {1, TAEGLICH}. 7	Flodukt registrieren
13	v13	Numerisch	1	0	FREIZEIT: MIT {1, TAEGLICH}. 7	- HI 8 Rechts
14	v14	Numerisch	1	0	FREIZEIT: INTE {1, TAEGLICH}. 7	- HI 8 Rechts
15	v15	Numerisch	1	0	FREIZEIT: SIC {1, TAEGLICH}. 7	- HI 8 Rechts
16	v16	Numerisch	1	0	FREIZEIT: EIN {1, TAEGLICH}. 7	- HI 8 Rechts
17	v17	Numerisch	1	0	FREIZEIT: SPA {1, TAEGLICH}. 7	- HI 8 Rechts
18	v18	Numerisch	1	0	FREIZEIT: YOG {1, TAEGLICH}. 7	- HI 8 Rechts
19	v19	Numerisch	1	0	FREIZEIT: ESS {1, TAEGLICH}. 7	- HI 8 Rechts
20	v20	Numerisch	1	0	FREIZEIT: BES {1, TAEGLICH}. 7	- HI 8 Rechts
21	v21	Numerisch	1	0	FREIZEIT: BES {1, TAEGLICH}. 7	- HI 8 Rechts
22	v22	Numerisch	1	0	FREIZEIT: GES {1, TAEGLICH}. 7	- HI 8 Rechts
23	v23	Numerisch	1	0	FREIZEIT: AUS {1, TAEGLICH}. 7	- HI 8 Rechts
24	v24	Numerisch	1	0	FREIZEIT: BET {1, TAEGLICH}. 7	- HI 8 Rechts
25	v25	Numerisch	1	0	FREIZEIT: EHR {1, TAEGLICH}. 7	- HI 8 Rechts
26	v26	Numerisch	1	0	FREIZEIT: KIR {1, TAEGLICH}. 7	- HI 8 Rechts
27	v27	Numerisch	1	0	FREIZEIT: KUN {1, TAEGLICH}. 7	- HI 8 Rechts
28	v28	Numerisch	1	0	FREIZEIT: BAS {1, TAEGLICH}. 7	- HI 8 Rechts
29	v29	Numerisch	1	0	FREIZEIT: AKTI {1. TAEGLICH}. 7	- HI 8 Rechts







3. Datenbeschreibungsund Datentransformationsbefehle



3. Datenbeschreibungs-/ Datentransformationsbefehle

- a) MISSING VALUES (Datenbereinigung, Definition fehlender Werte)
- b) RECODE (Werte einer bekannten Variablen werden verändert bzw. alte Werte durch neue ersetzt)
- c) VARIABLE LABELS (Variablen mit Etikett versehen)
- d) VALUE LABELS (Werte mit Etikett versehen)

Hilfsbefehl:

a) COMMENT (Kommentar im Syntaxfenster:***)



a) Datenbereinigung: MISSING VALUES Anwendungsbeispiel:

- Häufig sollen "weiß nicht" oder "Antwort verweigert" aus der Analyse ausgeschlossen werden
 - Automatisch fehlende Werte (System missing values)
 - Benutzerdefinierte fehlende Werte (User missing values)

Beispiel: Fehlende Werte bei der Variable "vsex" definieren (Wahlstudie 2005)

Göttingen, 08.05.2009

Statistik mit SPSS



vsex Geschlecht								
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente			
Gültig	1 Mann	856	51,4	51,4	51,4			
	2 Frau	806	48,4	48,4	99,8			
	9 k.A.	3	,2	,2	100,0			
1	Gesamt	1665	100,0	100,0				

Die Fälle, die den Wert "9" aufweisen, sollen als "fehlend" definiert und aus den folgenden Analysen ausgeschlossen werden







Regeln zum MISSING VALUES Befehl

Es dürfen max. drei Werte/ Wertebereiche als fehlende Werte definiert werden, Beispiel:

miss val v12 (7 thru 9, 99).

Schlüsselwörter für die Werteliste: Lowest, highest, thru miss val v12 (7 thru hi). miss val v14 (lo thru 3,99).



Regeln zum MISSING VALUES Befehl

Es können in einem Missing Values Befehl die Werte mehrerer Variablen als fehlend definiert werden:

> miss val v130a to v130k (8, 9) /v69 v70 (96 thru 99).

- Die Variablen in der TO-Gruppe müssen der Reihenfolge in der Datenmatrix entsprechen
- Der Missing Values Befehl kann mit folgendem Befehl rückgängig gemacht werden:

miss val v12 ().

b) RECODE und RECODE (INTO)

 RECODE: Variable wird umkodiert und überschrieben
RECODE (INTO): Variable wird umkodiert und bekommt neuen Namen

Anwendungsbeispiele:

- Umpolen von Kodierungen
- Zusammenfassen von Werten einer kontinuierlichen Variable, etwa Altersgruppen

Beispiel 1: Rekodierung der Variable v318 "Konsumhäufigkeit Bier und Wein" (Allbus 2004) in eine vierstufige Variable



b) RECODE und RECODE (INTO)

Häufigkeitstabelle ursprüngliche Variable

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig 3	1 MEHRMALS TAEGLICH	67	2,3	2,3	2,3
	2 <fast> TAEGLICH</fast>	331	11,2	11,3	13,5
2	3 MEHRMALS DIE WOCHE	589	20,0	20,0	33,6
	4 EINMAL DIE WOCHE	599	20,3	20,4	53,9
	5 2X - 3X IM MONAT	339	11,5	11,5	65,5
1	6 MONATLICH, SELTENER	448	15,2	15,2	80,7
0	7 NIE	568	19,3	19,3	100,0
	Gesamt	2941	99,8	100,0	
Fehlend	99 KEINE ANGABE	5	,2		
Gesam	t	2946	100,0		

v318 KONSUMHAUEFIGKEIT: BIER ODER WEIN



b) RECODE und RECODE (INTO) Variablen rekodieren





b) RECODE und RECODE (INTO)

Häufigkeitstabelle neue Variable "fre_alk"

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	,00	568	19,3	19,3	19,3
	1,00	787	26,7	26,8	46,1
	2,00	1188	40,3	40,4	86,5
	3,00	398	13,5	13,5	100,0
	Gesamt	2941	99,8	100,0	
Fehlend	System	5	,2		
Gesamt		2946	100,0		

fre_alk



b) RECODE und RECODE (INTO)

Häufigkeitstabelle neue Variable "fre_alk"





c und d) Variable und Werte Labels

Beispiele 3. Sitzung - SPSS Syntax Editor	
File Edit View Data Transform Analyze Graphs Utilities Run Window Help	
var lab fre_alk 'Konsumhäufigkeit Bier und Wein'.	
val lab fre_alk 0 'nie' 1 'seltener als 1x Woche' 2 'mind. 1x Woche' 3 'täglich	'. 📕
	-
📍 SPSS Processor is ready 🛛 👘 🖉 Ln 22 Col 1	



c und d) Variable und Werte Labels

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	,00 nie	568	19,3	19,3	19,3
	1,00 seltener als 1x Woche	787	26,7	26,8	46,1
	2,00 mind. 1x Woche	1188	40,3	40,4	86,5
	3,00 täglich	398	13,5	13,5	100,0
	Gesamt	2941	99,8	100,0	
Fehlend	System	5	,2		
Gesamt		2946	100,0		

fre_alk Konsumhäufigkeit Bier und Wein



Hilfsbefehl: COMMENT (Kommentar im Syntaxfenster:****)





Alle Befehle zusammen

