



Prof. Dr. Achim Bühl

SPSS 20

Einführung in die moderne Datenanalyse

13., aktualisierte Auflage

PEARSON

Higher Education
München • Harlow • Amsterdam • Madrid • Boston
San Francisco • Don Mills • Mexico City • Sydney
a part of Pearson pic worldwide

Inhaltsverzeichnis

Vorwort zur dreizehnten Auflage	15
1 Installation und Historie von SPSS	17
1.1 Systemvoraussetzungen von SPSS 20.	17
1.2 Installation von SPSS 20.	17
1.2.1 Deinstallation älterer Versionen.	18
1.2.2 Installation von SPSS 20 unter Windows 7.	18
1.3 Verknüpfung unter Windows 7 erstellen.	23
1.4 Arbeitsverzeichnis einrichten.	25
1.5 Übungsdateien des Buchs downloaden.	26
1.6 Programmeinstellungen von SPSS.	27
1.7 Die einzelnen Module von SPSS.	32
1.8 Historie des Programmsystems SPSS.	35
2 SPSS Statistics im Überblick	39
2.1 Auswahl einer Statistik-Prozedur.	40
2.1.1 Variablen auswählen.	43
2.1.2 Unterdialogboxen.	44
2.2 Einstellungen für den Daten-Editor.	46
2.3 Die Symbolleiste.	48
2.4 Erstellen und Editieren von Grafiken.	50
2.5 Der Viewer.	56
2.6 Editieren von Tabellen.	61
2.6.1 Der Pivot-Tabellen-Editor.	62
2.6.2 Weitere Bearbeitungsmöglichkeiten.	65
2.6.3 Symbole des Viewers.	71
2.7 Der Syntax-Editor.	72
2.8 Informationen zur Datendatei.	78
2.9 Das Hilfesystem.	83
2.10 Export der Ausgabe.	86
2.10.1 Statistische Ergebnisse nach Word übertragen.	86
2.10.2 Statistische Ergebnisse nach Word exportieren.	90
2.10.3 Diagramme nach Word übertragen.	93
2.10.4 Pivot-Tabellen und Diagramme in HTML-Dokumenten verwenden.	95
3 Datenaufbereitung	99
3.1 Kodierung und Kodeplan.	100
<i>i</i> 3.2 Datenmatrix	102
3.3 Start von SPSS.	103

Inhaltsverzeichnis

3.4	Daten-Editor	104
3.4.1	Definition der Variablen	104
3.4.2	Dateneingabe	118
3.5	Speichern einer Datendatei	121
3.6	Variablendeklarationen kopieren	121
3.7	Arbeiten mit mehreren Datendateien	125
3.7.1	Kopieren und Einfügen zwischen Datendateien	126
3.7.2	Umbenennen von Daten-Sets	128
3.8	Zusammenfügen von Datendateien	128
3.8.1	Fallweises Zusammenfügen	128
3.8.2	Variablenweises Zusammenfügen	134
3.9	Einlesen bereits vorhandener Daten	140
3.9.1	Einlesen von Daten mit Hilfe des SPSS-Assistenten	141
3.9.2	Einlesen von Daten mit Hilfe der Syntax	143
3.10	Arbeitssitzung beenden	145
Häufigkeitsauszählungen		147
4.1	Häufigkeitstabellen	147
4.2	Ausgabe statistischer Kennwerte	148
4.3	Mediän bei gehäuften Daten	153
4.4	Formate für Häufigkeitstabellen	157
4.5	Grafische Darstellung	158
Statistische Grundbegriffe und Kennwerte		165
5.1	Voraussetzungen für die Anwendung eines statistischen Tests	165
5.1.1	Skalenniveaus	165
5.1.2	Normalverteilung	168
5.1.3	Abhängigkeit und Unabhängigkeit von Stichproben	169
5.2	Übersicht über gängige Mittelwerttests	169
5.3	Die Irrtumswahrscheinlichkeit p	170
5.4	Statistischer Wegweiser	172
5.4.1	Strukturierung, Eingabe und Überprüfung der Daten	172
5.4.2	Deskriptive Auswertung	173
5.4.3	Analytische Statistik	173
5.5	Statistische Kennwerte	174
5.5.1	Deskriptive Statistiken	176
5.5.2	Fälle zusammenfassen	179
5.5.3	Verhältnisstatistiken	181
Datenselektion		185
6.1	Auswahl von Fällen	185
6.1.1	Einteilung der Operatoren	187
6.1.2	Relationale Operatoren	187
6.1.3	Logische Operatoren	188
6.1.4	Boolesche Algebra	188
6.1.5	Funktionen	191
6.1.6	Eingabe eines Konditional-Ausdrucks	193
6.1.7	Beispiele für Datenselektionen	196

Inhaltsverzeichnis

6.2	Ziehen einer Zufallsstichprobe	198
6.3	Fälle sortieren	200
6.4	Aufteilung der Fälle in Gruppen	201
Datenmodifikation.		207
7.1	Berechnung von neuen Variablen	207
7.1.1	Formulierung numerischer Ausdrücke	209
7.1.2	Funktionen	211
7.1.3	Einbindung der Syntax in den dialoggesteuerten Ablauf	216
7.2	Erstellen von Variablen mit Hilfe des Bereichseinteilers	218
7.3	Zählen des Auftretens bestimmter Werte	224
7.4	Umkodieren von Werten	227
7.4.1	Manuelles Umkodieren	227
7.4.2	Automatisches Umkodieren	232
7.5	Bedingte Berechnung von neuen Variablen	235
7.5.1	Formulierung von Bedingungen	235
7.5.2	Bildung eines Indexes	237
7.6	Aggregieren von Daten	242
7.7	Rangtransformationen	246
7.7.1	Beispiel einer Rangtransformation	246
7.7.2	Rangtypen	248
7.8	Gewichten von Fällen	251
7.8.1	Korrektur bei nicht gegebener Repräsentativität	251
7.8.2	Analyse von gehäuften Daten	257
7.9	Beispiele für die Berechnung neuer Variablen	261
7.9.1	Erstes Beispiel: Berechnung des Benzinverbrauchs	261
7.9.2	Zweites Beispiel: Berechnung des Datums des Ostersonntags	262
Datenexploration.		267
8.1	Aufdeckung von Eingabefehlern	267
8.2	Überprüfung der Verteilungsform	268
8.3	Berechnung von Kennwerten	268
8.4	Explorative Datenanalyse	268
8.4.1	Analysen ohne Gruppierungsvariablen	269
8.4.2	Analysen für Gruppen von Fällen	276
Kreuztabellen.		281
9.1	Erstellen von Kreuztabellen	281
9.2	Grafische Veranschaulichung von Kreuztabellen	295
9.3	Statistiken für Kreuztabellen	297
9.3.1	Chi-Quadrat-Test	298
9.3.2	Korrelationsmaße	301
9.3.3	Assoziationsmaße für nominalskalierte Variablen	304
9.3.4	Assoziationsmaße für ordinalskalierte Variablen	309
9.3.5	Weitere Assoziationsmaße	310

Inhaltsverzeichnis

10	Exakte Testmethoden.	317
10.1	Exakte p-Werte	319
10.2	Monte-Carlo-Methode	321
10.3	Integration in das SPSS-Basis-Modul	324
10.4	Nichtparametrische exakte Tests	328
10.4.1	Vergleich von zwei unabhängigen Stichproben	329
10.4.2	Vergleich von zwei abhängigen Stichproben	330
10.4.3	Vergleich von mehr als zwei unabhängigen Stichproben	333
10.4.4	Vergleich von mehr als zwei abhängigen Stichproben	334
10.4.5	Ein-Stichproben-Tests	335
10.4.6	Schnelle Berechnung	338
10.5	Statistiken für Kreuztabellen	338
11	Analyse von Mehrfachantworten.	345
11.1	Dichotome Methode	345
11.1.1	Definition von Sets	346
11.1.2	Häufigkeitstabellen für dichotome Setvariablen	347
11.1.3	Kreuztabellen mit dichotomen Setvariablen	349
11.2	Erstellen von Ranking-Listen	353
11.3	Kategoriale Methode	357
11.3.1	Definition von Sets	359
11.3.2	Häufigkeitstabellen für kategoriale Setvariablen	359
11.3.3	Kreuztabellen mit kategorialen Setvariablen	360
11.4	Dichotome und kategoriale Methode im Vergleich	362
12	Mittelwertvergleiche.	365
12.1	Vergleich von zwei unabhängigen Stichproben	366
12.2	Vergleich von zwei abhängigen Stichproben	368
12.3	Vergleich von mehr als zwei unabhängigen Stichproben	370
12.3.1	Zerlegen in Trendkomponenten	373
12.3.2	A-priori-Kontraste	373
12.3.3	A-posteriori-Tests	374
12.3.4	Weitere Optionen	375
12.4	Vergleich von mehr als zwei abhängigen Stichproben	375
12.5	Einstichproben-t-Test	377
12.6	Einbindung der Syntax in den dialoggesteuerten Ablauf	378
13	Nichtparametrische Tests.	381
13.1	Vergleich von zwei unabhängigen Stichproben	382
13.1.1	U-Test nach Mann und Whitney	382
13.1.2	Moses-Test	385
13.1.3	Kolmogorov-Smirnov-Test	386
13.1.4	Wald-Wolfowitz-Test	387
13.2	Vergleich von zwei abhängigen Stichproben	388
13.2.1	Wilcoxon-Test	388
"	13.2.2 Vorzeichen-Test	392
13.2.3	Chi-Quadrat-Test nach McNemar	394

Inhaltsverzeichnis

13.3	Vergleich von mehr als zwei unabhängigen Stichproben.	395
13.3.1	H-Test nach Kruskal und Wallis.	395
13.3.2	Median-Test	397
13.4	Vergleich von mehr als zwei abhängigen Stichproben.	398
13.4.1	Friedman-Test.	398
13.4.2	Kendalls W.	400
13.4.3	Cochrans Q.	401
13.5	Kolmogorov-Smirnov-Test zur Überprüfung der Verteilungsform . .	402
13.6	Chi-Quadrat-Einzelttest	403
13.7	Binomial-Test	407
13.8	Sequenzanalyse.	409
13.9	Nichtparametrische Tests mit Hilfe des Model Viewer.	410
13.9.1	U-Test nach Mann und Whitney.	410
13.9.2	H-Test nach Kruskal und Wallis.	415
14	Korrelationen.	419
14.1	Korrelationskoeffizient nach Pearson.	421
14.2	Rangkorrelationskoeffizienten nach Spearman und Kendali.	422
14.3	Partielle Korrelationen.	423
14.4	Distanz- und Ähnlichkeitsmaße.	426
14.5	Der Intraclass Correlation Coefficient (ICC).	429
14.6	Einbindung der Syntax in den dialoggesteuerten Ablauf.	430
15	Regressionsanalyse.	433
15.1	Einfache lineare Regression.	434
15.1.1	Berechnen der Regressionsgleichung.	435
15.1.2	Neue Variablen speichern.	437
15.1.3	Zeichnen einer Regressionsgeraden.	439
15.2	Multiple lineare Regression.	442
15.3	Nichtlineare Regression.	446
15.4	Binäre logistische Regression.	452
15.5	Multinomiale logistische Regression.	461
15.6	Ordinale Regression.	472
15.7	Probitanalyse.	480
15.8	Kurvenanpassung.	487
15.9	Gewichtsschätzung.	491
15.10	Partielle kleinste Quadrate.	495
15.11	Zweistufige kleinste Quadrate.	495
15.12	Kategoriale Regression.	496
15.12.1	Prinzip der Kategorienquantifikationen.	496
15.12.2	Zweites Beispiel: Der Untergang der Titanic.	504
15.12.3	Diskretisierung von Variablen.	510
16	Varianzanalysen.	517
16.1	Univariate Varianzanalyse.	519
16.1.1	Univariate Varianzanalyse (allgemeines lineares Modell)	520
16.1.2	Univariate Varianzanalyse nach Fisher.	527
16.1.3	Univariate Varianzanalyse mit Messwiederholung.	529

16.2	Kovarianzanalyse.	534
16.3	Multivariate Varianzanalyse.	536
16.4	Varianzkomponenten.	538
16.5	Lineare gemischte Modelle.	542
16.5.1	Varianzanalyse mit festen Effekten.	542
16.5.2	Kovarianzanalyse mit festen Effekten.	546
16.5.3	Analyse mit festen und zufälligen Effekten.	548
16.5.4	Analyse mit wiederholten Messungen.	550
17	Diskriminanzanalyse.	555
17.1	Beispiel aus der Medizin.	555
17.2	Beispiel aus der Soziologie.	564
17.3	Beispiel aus der Biologie.	572
17.4	Diskriminanzanalyse mit drei Gruppen.	574
18	Reliabilitätsanalyse.	579
18.1	Richtig-falsch-Aufgaben.	580
18.2	Stufen-Antwort-Aufgaben.	587
19	Faktorenanalyse.	589
19.1	Rechenschritte und Verfahrenstypen der Faktorenanalyse.	589
19.2	Explorative Faktorenanalyse.	590
19.2.1	Beispiel aus der Soziologie.	590
19.2.2	Beispiel aus der Psychologie.	598
19.3	Konfirmatorische Faktorenanalyse.	608
19.3.1	Beispiel aus der Freizeitforschung.	608
19.3.2	Grafische Darstellung des Rechnens mit Faktorwerten	616
19.3.3	Beispiel aus der Medienwissenschaft.	620
19.4	Das Rotationsproblem.	624
20	Clusteranalyse.	627
20.1	Das Prinzip der Clusteranalyse.	628
20.2	Hierarchische Clusteranalyse.	632
20.2.1	Hierarchische Clusteranalyse mit zwei Variablen.	632
20.2.2	Hierarchische Clusteranalyse mit mehr als zwei Variablen	637
20.2.3	Hierarchische Clusteranalyse mit vorgeschalteter Faktorenanalyse.	640
20.3	Ähnlichkeits- und Distanzmaße.	644
20.3.1	Intervallskalierte (metrische) Variablen.	644
20.3.2	Häufigkeiten.	647
20.3.3	Binäre Variablen.	648
20.4	Fusionierungsmethoden.	649
20.5	Clusteranalyse für hohe Fallzahlen (Clusterzentrenanalyse).	650
20.6	Die Two-Step-Clusteranalyse.	656
20.6.1	Die Two-Step-Clusteranalyse per Syntax ohne Model Viewer.	658
20.6.2	Die Two-Step-Clusteranalyse per Model Viewer.	676

21	Klassifikationsanalyse.	689
21.1	Einführendes Beispiel aus der Geschichtswissenschaft.	691
21.1.1	Erstellen einer Analysedatei.	691
21.1.2	Erzeugung und Interpretation eines Baumdiagramms	693
21.1.3	Interpretation der Vorhersagewerte.	699
21.1.4	Arbeiten mit dem Baumeditor.	702
21.2	Vertiefungsbeispiel aus dem Bereich der Wirtschaftswissenschaften	706
21.2.1	Erstellen einer Analysedatei.	707
21.2.2	Erzeugung und Interpretation eines Baumdiagramms	708
21.2.3	Erleichterung beim Erstellen der finalen Tabelle.	714
21.3	Der CHAID-Algorithmus als Klassifikationsmethode.	718
21.3.1	Erstellen einer Analysedatei.	719
21.3.2	Erzeugung und Interpretation eines Baumdiagramms	720
21.3.3	Ansichten und Navigation durch den Baum.	724
21.3.4	Analyse der finalen Segmente.	729
21.4	Der Exhaustive-CHAID-Algorithmus als Klassifikationsmethode	733
21.4.1	Erstellen einer Analysedatei.	734
21.4.2	Erzeugung eines Baumdiagramms.	735
21.4.3	Betrachtung des Baummodells und der Baumtabelle	737
21.4.4	Gewinnzusammenfassung, Risiko und Klassifikation	740
21.4.5	Vorhergesagte Werte.	742
21.4.6	Analyse der finalen Segmente.	743
21.5	Der CRT-Algorithmus als Klassifikationsmethode.	745
21.5.1	Der binäre Algorithmus in vergleichender Betrachtung	746
21.5.2	Vertiefungsbeispiel zum CRT-Algorithmus.	752
21.6	Der QUEST-Algorithmus als Klassifikationsmethode.	760
21.6.1	Erstellen einer Analysedatei.	761
21.6.2	Erzeugung und Interpretation eines Baumdiagramms	764
21.6.3	Analyse der Vorhersagewerte.	767
21.6.4	Analyse der Endknoten.	768
21.6.5	Dichotomisierung der Zielvariablen.	770
21.6.6	Analyse einzelner Parteien.	774
21.7	Die Hilfeoption des Baumeditors.	779
22	Loglineare Modelle.	781
22.1	Eine typische Anwendungssituation.	781
22.2	Das Prinzip der loglinearen Modelle.	784
22.3	Überblick über die loglinearen Modelle.	786
22.4	Hierarchisches loglineares Modell	787
22.5	Allgemeines loglineares Modell.	796
22.6	Logit-loglineares Modell.	802
23	Überlebens- und Ereignisdatenanalyse.	809
23.1	Sterbetafeln.	810
23.1.1	Einführende Beispiele aus der Medizin.	810
23.1.2	Vertiefende Beispiele aus der Soziologie.	819
23.2	Kaplan-Meier-Methode.	826

23.3	Regressionsanalyse nach Cox	831
23.3.1	Beispiel aus der Medizin	831
23.3.2	Beispiel aus der Ökonomie.	837
23.4	Cox-Regression mit zeitabhängigen Kovariaten.	838
24	Multidimensionale Skalierung.	843
24.1	Das Prinzip der MDS.	844
24.2	Beispiel aus dem Marketing-Bereich.	849
24.3	Ähnlichkeiten aus Daten erstellen.	851
24.4	Multidimensionale Skalierung und Faktorenanalyse.	859
25	Korrespondenzanalyse.	863
25.1	Einfache Korrespondenzanalyse.	864
25.1.1	Das Prinzip der einfachen Korrespondenzanalyse.	865
25.1.2	Beispiel einer Produktpositionierung	875
25.1.3	Das Seriationsproblem in der Archäologie.	881
25.2	Multiple Korrespondenzanalyse mit Nominalvariablen.	884
25.2.1	Erstes Beispiel: Produktpositionierung im Marketingbereich	884
25.2.2	Zweites Beispiel: Visualisierung der Variablenzusammenhänge.	890
25.2.3	Drittes Beispiel: Darstellung der Kategorienquantifikationen.	896
25.3	Multiple Korrespondenzanalyse mit beliebigen Variablen.	901
25.3.1	Erstes Beispiel: Alle Variablen numerisch.	901
25.3.2	Zweites Beispiel: Numerische und nominalskalierte Variablen.	907
25.4	Kanonische Korrespondenzanalyse.	913
26	Conjoint-Analyse.	929
26.1	Zielsetzung	929
26.2	Vorstellung eines Beispiels.	931
26.3	Erstellung eines orthogonalen Designs.	931
26.4	Die Ausgabe des orthogonalen Designs.	936
26.5	Die Prozedur CONJOINT.	939
27	Berichte und Gruppenwechsel.	947
27.1	Zeilenweise Berichte.	947
27.1.1	Erstellen eines einfachen Berichts.	947
27.1.2	Zweistufiger Gruppenwechsel.	952
27.1.3	Dreistufiger Gruppenwechsel.	953
27.1.4	Berichts-Layout	957
27.2	Spaltenweise Berichte.	962
27.3	Komprimierte Berichtsausgabe.	968
27.4	Übungsaufgaben.	971
28	Diagramme.	973

28.1	Balkendiagramm	973
28.1.1	Einfaches Balkendiagramm: Darstellung von Häufigkeiten	973
28.1.2	Einfaches Balkendiagramm: Kennwerte einer metrischen Variablen	977
28.1.3	Gruppiertes Balkendiagramm	983
28.1.4	Gestapeltes Balkendiagramm	985
28.2	Liniendiagramm	988
28.2.1	Einfaches Liniendiagramm	989
28.2.2	Mehrfaches Liniendiagramm	991
28.3	Flächendiagramm	993
28.3.1	Einfaches Flächendiagramm	993
28.3.2	Gestapeltes Flächendiagramm	995
28.4	Kreisdiagramm	996
28.5	Streudiagramm	998
28.6	Histogramm	1003
28.6.1	Einfaches Histogramm	1003
28.6.2	Gestapeltes Histogramm	1005
28.7	Hoch-Tief-Diagramme	1006
28.7.1	Einfache Hoch-Tief-Schluss-Diagramme	1006
28.7.2	Einfache Bereichsbalken	1009
28.7.3	Gruppierte Hoch-Tief-Schluss-Diagramme	1010
28.7.4	Differenzliniendiagramme	1010
28.8	Boxplot	1011
28.8.1	Einfacher Boxplot	1011
28.8.2	Gruppiertes Boxplot	1013
28.9	Doppelachsen	1014
28.9.1	Zwei Y-Achsen mit kategorialer X-Achse	1014
28.9.2	Zwei Y-Achsen mit metrischer X-Achse	1017
28.10	Erstellen eines Diagramms aus einer Pivot-Tabelle	1018
Anhang A Verzeichnis der verwendeten Dateien		1021
Anhang B Weiterführende Literatur		1033
Index		1039