

Vorwort	V
Einleitung	1
Kapitel 1 Anwendungsgebiete der Clusteranalyse	2
Einige Beispiele zu Anwendungen der Clusteranalyse	2
Kapitel 2 Hinweis zur Benutzung des Handbuchs	4
Kapitel 3 Grundkonzepte bei CLUSTAN	5
Classification Arrays	5
CLUSDECK	5
Kapitel 4 Wie benutzt man CLUSTAN?	7
CLUSTAN-Prozedur-Schlüsselwörter	7
Beispiel eines CLUSTAN-Jobs	8
Kapitel 5 Vorschläge für Analysen	9
(a) Große Populationen	9
(b) Populationen vom Umfang kleiner als 150	10
(c) Kleine Populationen (n kleiner als 40)	11
(d) Berechnung neuer Ähnlichkeitsmaße	11
(e) Verbindungen zu SPSS, MDSCAL und BMD	11
Kapitel 6 Veröffentlichte Methoden	12
(a) Hierarchische Fusion	12
(b) Monothetische Division	12
(c) Interative Relokation	12
(d) Andere Methoden	12
Kapitel 7 Prozedur FILE: liest kontinuierliche und binäre Daten	13
Datenaufbereitung	13
Populationsbeschränkungen	13
Kontinuierliche Daten	13
Normierung kontinuierlicher Daten	15
Hauptkomponentenanalyse	15
Binäre Daten	15
Gemischte Daten	17
Ausgabe der Ergebnisse	18
FILE-Spezifikationen	18
Deckaufbau	20
Vollständiges Beispiel	21
Kapitel 8 Prozedur CORREL: berechnet Abstands-/Ähnlichkeitsmatrix	22
Ähnlichkeitsmatrix	22
Abstandskoeffizienten	23
K-Linkage-Listen	23
Maskierungsmöglichkeit - Regeln	23
Beispiele zur Maskierungsmöglichkeit	25
CORREL-Spezifikationen	26
Deckaufbau	27
Kurzlauf	27

Kapitel 9	Prozedur DISTIN: liest Abstands-/Ähnlichkeitsmatrix	28
	Beobachtete Ähnlichkeiten	28
	Clustern einer Korrelationsmatrix	28
	Extern berechnete Ähnlichkeiten	29
	Inter-Cluster-Ähnlichkeiten	29
	Große Stichproben	29
	In Verwendung mit Prozedur HIERARCHY	29
	Hinzufügen oder Ersetzen einer Ähnlichkeitsmatrix	30
	DISTIN-Spezifikationen	30
	Deckaufbau	31
	Kurzlauf	31
Kapitel 10	Prozedur HIERARCHY: agglomerative Clustermethoden	33
	Methode	33
	Rechnenmethode	33
	HIERARCHY-Optionen und mögliche Ähnlichkeitskoeffizienten	34
	Ähnlichkeitsmatrix	35
	Linkage-Methoden. Optionen 1-3	35
	Centroid Sorting. Option 4	35
	Optionen 5-8	35
	Zeichnen von Dendrogrammen	36
	HIERARCHY-Spezifikationen	37
	Deckaufbau	37
	Analyse der Klassifikationen	38
	Kurzlauf	38
Kapitel 11	Prozedur TREE: druckt Dendrogramme und Division-Bäume	40
	Dendrogrammdateien	40
	Ausgabestilarten	40
	Fallkennungen	42
	TREE-Spezifikationen	42
	Deckaufbau	43
	Kurzlauf	43
Kapitel 12	Prozedur CENTROID: agglomerative Centroid-Methoden	44
	Methode	44
	Rechenmethode	44
	CENTROID-Spezifikationen	45
	Deckaufbau	45
	Kurzlauf	45
Kapitel 13	Prozedur RELOCATE: Methoden zum Partitionieren	46
	Prinzip des Prozesses	46
	Spezifizierung der ersten Klassifikation	46
	Die Removal-Option	47
	Suche nach der Lösung eines globalen Optimums	47
	Schwellenwert für minimale Clustergröße	48
	Erstellen eines Residuums mit der Ähnlichkeitsschwelle	48
	Maximale Anzahl der Relokationszyklen	48
	Hierarchische Fusion ohne Relokation	48
	Auswahl der Daten	49
	Ausgabe der Ähnlichkeiten	49
	RELOCATE-Spezifikationen	51
	Spezielle Größenbeschränkungen	53
	Deckaufbau	53
	Kurzlauf	54

Kapitel 14	Prozedur MODE: Mode Analysis	55
	Methode	55
	Die K-Linkage-Alternative	56
	Probabilistische Interpretation	56
	Steuerung der Ausgabe	57
	Endbedingungen	57
	Zusammenfassung der Eingabeparameter und der Default-Optionen	58
	MODE-Spezifikationen	59
	Deckaufbau	59
	Kurzlauf	59
Kapitel 15	Prozedur DENSITY: verbesserte Methode des Mode-Suchens	60
	Vergleich mit der Mode Analysis	60
	Dichteschätzer	61
	Modifiziertes Dendrogramm	62
	Testen auf natürliche Cluster	62
	Default-Werte für die Dichteparameter	63
	DENSITY-Spezifikationen	63
	Deckaufbau	63
	Kurzlauf	63
Kapitel 16	Prozedur DIVIDE: monothetisch divisive Methoden	65
	Methode	65
	Standard-Methoden	66
	Division-Baum	66
	Rechenmethode	67
	Trace-Funktion	67
	Group Analysis	68
	Association Analysis	68
	Beschreibung der Eingabe	68
	DIVIDE-Spezifikationen	69
	Deckaufbau	70
	Kurzlauf	70
Kapitel 17	Prozedur KNEND: Jardine-Sibson B(k) Methode	71
	Graphentheoretische Beschreibung	71
	Jardine-Sibson Algorithmus	71
	Cole-Wishart Algorithmus	71
	Cluster-Erkennung	72
	KNEND-Spezifikationen	72
	Deckaufbau	73
	Kurzlauf	73
Kapitel 18	Prozedur DNDRITE: Minimalbaum-Partition	74
	Optimierung der Fehlersumme	74
	DNDRITE-Statistiken	75
	Summe der Ähnlichkeiten innerhalb der Gruppen WS	75
	Verallgemeinerung auf alle Ähnlichkeiten	75
	Erzeugung und Analyse von Klassifikationen	76
	DNDRITE-Spezifikationen	76
	Deckaufbau	77
	Kurzlauf	77

		XI
Kapitel 19	Prozedur EUCLID: nicht-lineare Programmierungsmethode	78
	Methode	78
	Genauigkeit und Konvergenzgeschwindigkeit	78
	Spezifizierung der Clusteranzahl	79
	Trace der Iterationen und Klassifikationen	80
	EUCLID-Spezifikationen	81
	Deckaufbau	81
	Kurzlauf	81
Kapitel 20	Prozedur RESULT: druckt Daten und Clusterstatistiken	83
	Wie funktionieren die Print-Optionen?	83
	Classification Arrays und Clusterdiagnostik	83
	Definition der clusterdiagnostischen Statistiken	84
	RESULT-Spezifikationen	85
	Deckaufbau	87
	Kurzlauf	87
Kapitel 21	Prozedur DUMP: Ausdruck und Ausgabe von Klassifikationen	88
	DUMP-Parameterkarte	88
	Beschreibung	88
	DUMP-Spezifikationen	90
	Deckaufbau	90
Kapitel 22	Prozedur LABELS: liest alphanumerische Kennungen	91
	Anwendungsmethode	91
	Kennungen für numerische Variablen	91
	Kennungen für binäre Merkmale	91
	Kennungen für Fälle	91
	Kennungen für Faktoren	91
	LABELS-Spezifikationen	92
	Deckaufbau	92
	Kurzlauf	93
Kapitel 23	Graph-Plotting-Konstanten	94
	Herkömmliche Trommel-Plotter	94
	Wahl einer Mikrofilm-Kamera	94
	Wahl eines Mikrofilm-Maßstabs	94
	Default-Plot-Dimensionen	95
	Größe der Legenden	95
Kapitel 24	Prozedur PLINK: zeichnet Dendrogramme und Division-Bäume	96
	Normale Anwendungsmethode	96
	Dendrogrammdateien	96
	Von Hand erstellte Dendrogrammdecks	96
	Wahl der Legenden	97
	Zeichnen eines Teilbaums	97
	Dendrogramm-Transformationen	97
	PLINK-Spezifikationen	98
	Deckaufbau	99
	Kurzlauf	100

Kapitel 25	Prozedur SCATTER: zeichnet Streu- und Clusterdiagramme	101
	Wahl der Achsen	101
	Streudiagramm	101
	Zeichnen der Fallzahlen oder -kennungen	101
	Clusterdiagramm	102
	Clusterkreise	102
	Cluster-Begrenzungslinien	103
	Minimalbaum	103
	Achsenlängen	103
	Zeichnen mit geographischen Koordinaten	104
	Zeichnen signifikanter Bindungen	105
	Zeichengröße	105
	Koordinatenachsen	105
	Achsenbeschriftung	105
	Datenzusammenfassung	106
	SCATTER-Spezifikationen	107
	Deckaufbau	110
	Kurzlauf	110
Kapitel 26	SPSS: liest Daten von SPSS-Systemdateien	111
	Methode	111
	Beschränkungen	111
	Statistische Transformationen	111
	Variablen-Identifikation	112
	Code für fehlende Werte	112
	SPSS-Spezifikationen	112
	Deckaufbau	113
	Kurzlauf	113
Kapitel 27	Utility-Prozeduren	114
	RESTART-Prozedur	114
	SIZE-Prozedur	114
	TRACE-Schlüsselwort	115
	COMMENT-Schlüsselwort	115
	STOP-Schlüsselwort	115
Kapitel 28	Betrachtung der Ähnlichkeitskriterien	116
	Dreiteiliger Clusteralgorithmus	116
	Wahl des Ähnlichkeitskriteriums	117
	Ähnlichkeitskriterien-erzeugende Funktion	117
	USER-Funktion	118
Kapitel 29	Ähnlichkeitskoeffizienten	120
	Allgemeine Bezeichnungsweise	121
	Binäre Bezeichnungsweise	121
Kapitel 30	Spezifikationen der Job-Steuersprache	135
	Quelldaten auf Band oder Platte	135
	CLUSTAN-SAVE-Datei	135
	Plotting JCL	135
	Stanzen von Dendrogrammdecks	136
	Stanzen von Classification Arrays	136
	SPSS SAVE-Datei	136
	Beispiele	136
	Kopplung von CLUSTAN mit BMDX72	136
	Zeichnen kanonischer Variablen aus BMD07M	137
	Benutzung einer von MDSCAL erzeugten Konfiguration	137

Kapitel 31	Beispiel	139
	Diskussion	140
	Job für Plot-Beispiele	140
	Diskussion	142
	Graphen-Beispiele	143
Kapitel 32	Testdatenplan	148
	Job A	148
	Job B	149
	Job C	151
	Job D	152
	Job E	153
	Diskussion	154
	Job F	155
	Diskussion	157
Kapitel 33	SPSS-Beispiel	159
	Diskussion	159
Anhang:	CLUSTAN-Fehlermeldungen	160
Literaturverzeichnis		176
Sachverzeichnis		178