Thomas Cleff

Deskriptive Statistik und Explorative Datenanalyse

Eine computergestützte Einführung mit Excel, SPSS und STATA

3., überarbeitete und erweiterte Auflage

Springer

Inhaltsverzeichnis

1 Statistik und empirische Forschung 1
1.1 Statistik lügt? 1
1.2 Zwei Arten von Statistik 3
1.3 Statistik als Erkenntnisprozess 5
1.4 Phasen empirischer Forschung 7
1.4.1 Von der Erkundung zur Theorie 7
1.4.2 Von der Theorie zum Modell 8
1.4.3 Vom Modell zur „Business Intelligence“ 12
Literatur 14

2 Vom Zahlenwust zum Datensatz 15
2.1 Möglichkeiten der Datenbeschaffung 15
2.2 Die Entscheidung für ein Skalenniveau 18
2.3 Datenerfassung mit dem Computer: Skalierung und Kodierung 22
2.4 Fehlende Werte oder Missing Values 23
2.5 Ausreißer und offensichtlich falsche Werte 26
2.6 Übungsaufgaben zum Abschnitt 27
Literatur 28

3 Vom Datensatz zur Information 29
3.1 Erste Auswertungsschritte und grafische Darstellungen 29
3.2 Lageparameter als Informationsreduktion 36
3.2.1 Modus oder Modalwert 37
3.2.2 Der Mittelwert 37
3.2.3 Geometrisches Mittel 42
3.2.4 Harmonisches Mittel 44
3.2.5 Der Median 47
3.2.6 Quartile und Quantile 50
3.3 Boxplot – Erster Einblick in die Verteilung 51
3.4 Streuungsparameter 54
3.4.1 Die Standardabweichung und die Varianz 55

3.4.2 Der Variationskoeffizient 57
3.5 Schiefe und Kurtosis 59
3.6 Robustheit von Parametern 62
3.7 Konzentrationsmaße 63
3.8 Berechnung univariater Parameter mit dem Computer 66
3.8.1 Berechnung univariater Parameter mit SPSS 66
3.8.2 Berechnung univariater Parameter mit Stata 67
3.8.3 Berechnung univariater Parameter mit Excel 2010 68
3.9 Übungsaufgaben zum Abschnitt 69
Literatur 72

4 Bivariate Zusammenhänge 73
4.1 Bivariate Skalenniveau-Kombinationen 73
4.2 Zusammenhang zweier nominaler Variablen 74
4.2.1 Kontingenztabellen 74
4.2.2 Die Chi-Quadrat Berechnung 75
4.2.3 Der Phi-Koeffizient 80
4.2.4 Der Kontingenzkoeffizient 83
4.2.5 Cramers V 84
4.2.6 Nominale Zusammenhänge mit SPSS 85
4.2.7 Nominale Zusammenhänge mit Stata 89
4.2.8 Nominale Zusammenhänge mit Excel 90
4.2.9 Übungsaufgaben zum Abschnitt 91
4.3 Zusammenhang zweier metrischer Variablen 94
4.3.1 Das Streudiagramm 94
4.3.2 Der Korrelationskoeffizient nach Bravais-Pearson 98
4.4 Zusammenhang ordinalskalierter Variablen 101
4.4.1 Die Rangkorrelation nach Spearman (Rho) 102
4.4.2 Kendalls Tau (τ) 109
4.5 Zusammenhangsmaße zweier Variablen mit unterschiedlichem Skalenniveau 115
4.5.1 Zusammenhang nominaler und metrischer Variablen 115
4.5.2 Zusammenhang nominaler und ordinaler Variablen 117
4.5.3 Zusammenhang ordinaler und metrischer Variablen 118
4.6 Korrelationsrechnung mit dem Computer 119
4.6.1 Korrelationsrechnung mit SPSS 120
4.6.2 Korrelationsrechnung mit Stata 121
4.6.3 Korrelationsrechnung mit Excel 122
4.7 Scheinkorrelationen 123
4.7.1 Partielle Korrelation 126
4.7.2 Partielle Korrelation mit SPSS 128
4.7.3 Partielle Korrelation mit Stata 128

4.7.4 Partielle Korrelation mit Excel . . . 129
4.8 Übungsaufgaben zum Abschnitt . . . 130
Literatur . . . 133

5 Regressionsanalyse . . . 135
5.1 Erste Schritte einer Regressionsanalyse . . . 135
5.2 Koeffizienten der bivariaten Regression . . . 138
5.3 Multivariate Regressionskoeffizienten . . . 143
5.4 Die Anpassungsgüte der Regression . . . 144
5.5 Regressionsrechnung mit dem Computer . . . 146
5.5.1 Regressionsrechnung mit Excel . . . 146
5.5.2 Regressionsrechnung mit SPSS und Stata . . . 148
5.6 Anpassungsgüte multivariater Regressionen . . . 149
5.7 Regression mit unabhängiger Dummy-Variable . . . 150
5.8 Hebelwirkungen von Beobachtungen . . . 153
5.9 Nichtlineare Regressionen . . . 154
5.10 Ansätze einer Regressionsdiagnostik . . . 158
5.11 Übungsaufgaben zum Abschnitt . . . 164
Literatur . . . 170

6 Zeitreihen- und Indexrechnung . . . 171
6.1 Preisindizes . . . 172
6.2 Mengenindizes . . . 180
6.3 Wertindizes (Umsatzindizes) . . . 182
6.4 Deflationierung von Zeitreihen . . . 182
6.5 Umbasierung und Verkettung von Indizes . . . 184
6.6 Übungsaufgaben zum Abschnitt . . . 185
Literatur . . . 187

7 Clusteranalyse . . . 189
7.1 Hierarchische Clusteranalyse . . . 190
7.2 Die Clusterzentrenanalyse . . . 206
7.3 Clusteranalyse mit dem Computer . . . 208
7.3.1 Clusteranalyse mit SPSS . . . 209
7.3.2 Clusteranalyse mit Stata . . . 209
7.4 Übungsaufgaben zur Clusteranalyse . . . 211
Literatur . . . 214

8 Faktorenanalyse . . . 217
8.1 Faktorenanalyse: Grundlagen, Vorgehensweise und Interpretation . . . 217
8.2 Faktorenanalyse mit dem Computer . . . 229
8.2.1 Faktorenanalyse mit SPSS . . . 229
8.2.2 Faktorenanalyse mit Stata . . . 231

8.3 Übungsaufgaben zur Faktorenanalyse 232
Literatur 234

9 Lösungen der Übungsaufgaben 235

Formelsammlung 253

Sachverzeichnis 261