

# Inhaltsverzeichnis

<b>Vorwort</b>	13
<b>Vorwort zur zweiten Auflage</b>	15
<b>Fragestellungen und Methoden</b>	17
<b>Kapitel 1 Einführung</b>	21
1.1 Konzeption des Buchs .....	22
1.2 Aufbau des Buchs .....	23
1.3 Programmversionen von R .....	24
1.3.1 Das REdaS Package .....	24
1.4 Wie kann dieses Buch verwendet werden? .....	24
1.5 Typografische und andere Konventionen in diesem Buch .....	25
<b>Kapitel 2 Statistische Grundbegriffe</b>	29
2.1 Einige Beispiele .....	30
2.1.1 Hochrechnung (statistisches Schätzen) .....	31
2.1.2 Prüfen von Fragestellungen (Testen von Hypothesen) .....	32
2.1.3 Erstellen von Modellen (statistisches Modellieren) .....	34
2.2 Grundlegende Konzepte .....	35
2.3 Messung und Typen von Daten .....	38
2.4 Arten von Fragestellungen und Variablen .....	43
2.5 Zusammenfassung der Konzepte .....	45
2.6 Übungen .....	46
<b>Teil I Einführung in R</b>	49
<b>Kapitel 3 Erste Schritte</b>	51
3.1 Download und Installation von R .....	53
3.1.1 Download .....	53
3.1.2 Installation .....	54
3.1.3 Starten und Beenden von R .....	55
3.1.4 Anpassen von R .....	56
3.1.5 Installation von Erweiterungen (Contributed Packages) .....	58
3.2 Aller Anfang ist leicht .....	60
3.3 Übungen .....	80
3.4 R-Befehle im Überblick .....	81

<b>Kapitel 4</b>	<b>Daten in R – vom Fragebogen zum fertigen Datensatz</b>	<b>85</b>
4.1	Fragebogen und Kodierung .....	86
4.2	Erfassen der kodierten Daten .....	90
4.2.1	Eingabe der Daten .....	90
4.2.2	Abspeichern und Einlesen der Daten .....	94
4.3	Organisation eines Datensatzes – Data Frames .....	94
4.3.1	Manipulation einzelner Variablen eines Data Frame .....	94
4.3.2	Listen .....	97
4.3.3	Faktoren .....	101
4.3.4	Auswählen von Beobachtungseinheiten (Fällen) mit logischen Operationen .....	105
4.4	Transformieren der Daten bzw. Erzeugen von neuen Variablen .....	111
4.4.1	Berechnen neuer Variablen .....	111
4.4.2	Unendliche und undefinierte Werte .....	116
4.4.3	Umkodieren von Variablen .....	116
4.4.4	Modifikation eines Data Frame .....	122
4.4.5	Datenkontrolle .....	125
4.5	Übungen .....	128
4.6	R-Befehle im Überblick .....	128
<b>Kapitel 5</b>	<b>Mehr R</b>	<b>133</b>
5.1	Die R-Arbeitsumgebung .....	135
5.1.1	Der Workspace .....	136
5.1.2	Working Directory – das Arbeitsverzeichnis .....	138
5.2	R-Grafik .....	139
5.2.1	High-level Plotting Functions .....	139
5.2.2	Hinzufügen von Grafikelementen (low-level plotting functions) .....	144
5.2.3	Spezielle Einstellungen (Graphical Parameters) .....	151
5.2.4	Weitere Grafikfunktionen und Packages .....	156
5.3	Weiterverwenden des R-Outputs .....	159
5.4	Einlesen von R-Befehlen .....	161
5.4.1	Der R Editor .....	161
5.4.2	Einlesen von R-Skripten (Quellcode) .....	163
5.4.3	Direktes Kopieren von R-Code – Einfügen über die Zwischenablage .....	164
5.5	Einlesen und Schreiben externer Dateien .....	165
5.5.1	Daten aus Microsoft Excel bzw. LibreOffice Calc .....	165
5.5.2	Dateien aus anderen Statistikpaketen (z. B. SPSS) .....	168
5.5.3	Direktes Kopieren – Einfügen über die Zwischenablage .....	169
5.5.4	Schreiben von Dateien .....	170
5.6	Statistische Funktionen .....	171
5.6.1	Die drei Funktionen statistischer Verteilungen .....	171
5.6.2	Erzeugen von (Pseudo-)Zufallszahlen .....	173
5.6.3	Verfügbare Verteilungsfunktionen .....	174



6.6	Zusammenfassung der Konzepte .....	262
6.7	Übungen .....	263
6.8	Vertiefung: Die Chi-Quadrat-Verteilung oder wie entsteht ein $p$ -Wert? ...	264
6.9	R-Befehle im Überblick .....	267
<b>Kapitel 7</b>	<b>Mehrere kategoriale Variablen</b> .....	<b>269</b>
7.1	Beschreibung mehrerer kategorialer Variablen .....	270
7.1.1	Numerische Beschreibung .....	271
7.1.2	Grafische Beschreibung .....	274
7.2	Ist die Verteilung von Häufigkeiten in verschiedenen Gruppen gleich?... ..	277
7.3	Unterscheiden sich Anteile in zwei oder mehreren Gruppen? .....	282
7.3.1	Eingabe einer Tabelle .....	283
7.3.2	Vergleich der Anteile .....	283
7.3.3	Exakter Test nach Fisher .....	284
7.4	Sind zwei kategoriale Variablen unabhängig? .....	286
7.4.1	Datenaufbereitung .....	286
7.4.2	Unabhängigkeitstest .....	287
7.5	Unterscheidet sich das Risiko in zwei Gruppen? .....	290
7.5.1	Odds-Ratio .....	291
7.5.2	Odds-Ratio-Test .....	293
7.6	Wie kann man Veränderungen von Anteilen testen? .....	294
7.6.1	Unabhängige und abhängige Stichproben .....	294
7.6.2	McNemar-Test .....	295
7.7	Zusammenfassung der Konzepte .....	296
7.8	Übungen .....	297
7.9	R-Befehle im Überblick .....	299
<b>Teil III</b>	<b>Metrische Daten</b> .....	<b>301</b>
<b>Kapitel 8</b>	<b>Eine metrische Variable</b> .....	<b>303</b>
8.1	Wie kann man die Verteilung einer metrischen Variablen beschreiben? ..	304
8.1.1	Klassifizieren, Tabellen und Histogramme .....	305
8.1.2	Maßzahlen zur Beschreibung der Verteilung .....	309
8.1.3	Boxplot .....	317
8.1.4	Ausreißer .....	319
8.1.5	Weitere grafische Beschreibungsmethoden .....	320
8.2	Ist der Mittelwert der Grundgesamtheit anders als eine bestimmte Vorgabe? .....	326
8.3	Bereich für den Mittelwert in einer Grundgesamtheit .....	330
8.4	Folgt eine metrische Variable einer bestimmten Verteilung? .....	332
8.4.1	Q-Q-Plot .....	333
8.4.2	Kolmogorov-Smirnov-Test und Shapiro-Wilk-Test .....	336
8.4.3	Anpassungstest mit der $\chi^2$ -Verteilung .....	338

8.5	Zusammenfassung der Konzepte .....	341
8.6	Übungen .....	341
8.7	R-Befehle im Überblick .....	342
<b>Kapitel 9</b>	<b>Mehrere metrische Variablen</b> .....	<b>345</b>
9.1	Wie stark ist der Zusammenhang zwischen zwei metrischen Variablen? ..	347
9.1.1	Grafische Beschreibung .....	348
9.1.2	Korrelationskoeffizient nach Pearson .....	349
9.1.3	Korrelationskoeffizient nach Spearman .....	352
9.2	Welche Form hat der Zusammenhang zwischen zwei Variablen? .....	352
9.2.1	Lineares Regressionsmodell .....	353
9.2.2	Rechenergebnisse .....	355
9.3	Vorhersage von Werten .....	358
9.3.1	Punktprognosen .....	358
9.3.2	Intervallprognosen .....	360
9.4	Zusammenhang einer mit mehreren Variablen .....	362
9.4.1	Multiple lineare Regression .....	363
9.4.2	Kategoriale als erklärende Variablen .....	365
9.4.3	Modellselektion .....	368
9.4.4	Modelldiagnostik .....	370
9.4.5	Prognose .....	374
9.5	Unterscheiden sich Mittelwerte zu zwei oder mehreren Zeitpunkten? ...	375
9.5.1	Grafische Beschreibung .....	376
9.5.2	Analyse der Fragestellung .....	377
9.6	Wie kann man den zeitlichen Verlauf einer Variablen beschreiben? .....	379
9.6.1	Zeitreihen .....	379
9.6.2	Zeitreihenzerlegung .....	380
9.6.3	Trend .....	381
9.6.4	Saison .....	387
9.6.5	Zusammenfassung der Zeitreihenzerlegung .....	389
9.6.6	Prognose .....	390
9.6.7	Autokorrelation .....	392
9.7	Zusammenfassung der Konzepte .....	394
9.8	Übungen .....	395
9.9	R-Befehle im Überblick .....	397
<b>Teil IV</b>	<b>Metrische und kategoriale Daten</b> .....	<b>399</b>
<b>Kapitel 10</b>	<b>Metrische und kategoriale Variablen</b> .....	<b>401</b>
10.1	Unterscheiden sich die Mittelwerte in zwei Gruppen? .....	403
10.1.1	Grafische und numerische Beschreibung .....	403
10.1.2	Analyse der Fragestellung .....	404

10.2	Unterscheidet sich die Lage einer Variablen zwischen zwei Gruppen? ...	406
10.2.1	Beschreibung der Stichprobe .....	406
10.2.2	Analyse der Fragestellung .....	407
10.3	Unterscheiden sich die Mittelwerte mehrerer Gruppen? .....	408
10.3.1	Grafische und numerische Beschreibung .....	409
10.3.2	Analyse der Fragestellung .....	410
10.3.3	Post-hoc-Tests .....	413
10.4	Unterscheidet sich die Lage einer Variablen zwischen mehreren Gruppen? .....	413
10.5	Wie wirken zwei kategoriale Variablen kombiniert auf Mittelwerte? .....	416
10.5.1	Numerische und grafische Beschreibung .....	416
10.5.2	Analyse der Fragestellung .....	418
10.5.3	Modellselektion .....	420
10.6	Hängen Chancen von einer oder mehreren Variablen ab? .....	423
10.6.1	Logistische Regression .....	424
10.6.2	Logistische Regression mit mehreren erklärenden Variablen .....	427
10.7	Unterscheiden sich Chancen und Odds-Ratios zwischen Gruppen? .....	430
10.7.1	Vergleich von Odds in mehreren Gruppen? .....	431
10.7.2	Vergleich von Odds-Ratios in mehreren Gruppen? .....	433
10.8	Zusammenfassung der Konzepte .....	436
10.9	Übungen .....	437
10.10	R-Befehle im Überblick .....	438
<b>Teil V</b>	<b>Multivariate Daten</b>	<b>441</b>
<b>Kapitel 11</b>	<b>Dimensionsreduktion</b>	<b>443</b>
11.1	Kann man Komplexität multidimensionaler Daten reduzieren? .....	444
11.1.1	Grundlagen der Hauptkomponentenanalyse .....	445
11.1.2	Anwendung der Hauptkomponentenanalyse .....	455
11.2	Weiterverwenden der Ergebnisse einer Hauptkomponentenanalyse .....	466
11.3	Zusammenfassung der Konzepte .....	470
11.4	Übungen .....	471
11.5	Vertiefung: Extraktion der Hauptkomponenten für zwei Variablen .....	472
11.6	R-Befehle im Überblick .....	475
<b>Kapitel 12</b>	<b>Gruppierung von Beobachtungen</b>	<b>477</b>
12.1	Wie entdeckt man Gruppen ähnlicher Beobachtungen? .....	479
12.1.1	Distanz- und Ähnlichkeitsmaße .....	480
12.1.2	Hierarchische Clusterverfahren .....	482
12.1.3	Outputteile .....	485
12.1.4	Anwendung auf die Demografiekennzahlen .....	489
12.1.5	Teilungsverfahren .....	491
12.1.6	Speichern der Clusterzugehörigkeit .....	492

12.2	Wie findet man Cluster in den Variablen? .....	493
12.3	Wie findet man Cluster in großen Datensätzen? .....	494
12.3.1	Centroid-Verfahren .....	495
12.3.2	Outputteile des Verfahrens .....	495
12.3.3	Analyse des Outputs .....	497
12.4	Wie können kategoriale Variablen einbezogen werden? .....	499
12.4.1	Distanzmatrix .....	500
12.4.2	Fuzzy-Verfahren .....	500
12.4.3	Outputteile .....	501
12.4.4	Analyse des Outputs .....	502
12.5	Zusammenfassung der Konzepte .....	503
12.6	Übungen .....	504
12.7	R-Befehle im Überblick .....	505
<b>Kapitel 13 Mehrdimensionale kategoriale Daten</b>		<b>507</b>
13.1	Modelle für Kreuztabellen .....	508
13.1.1	Modelle und erwartete Häufigkeiten .....	509
13.1.2	Modellhierarchie .....	514
13.1.3	Loglineare Modelle .....	515
13.2	Welche Modelle beschreiben die Eintragungen einer Kreuztabelle gut? ..	515
13.2.1	Pearson- und Likelihood-Ratio-Statistik .....	516
13.2.2	Modellberechnung in R .....	517
13.2.3	Modellauswahl .....	520
13.3	Höherdimensionale loglineare Modelle .....	521
13.3.1	Dreidimensionale Kreuztabellen .....	522
13.3.2	Loglineare Modelle in drei Dimensionen .....	522
13.3.3	Modellauswahl .....	524
13.3.4	Interpretation .....	526
13.4	Weitere Beispiele .....	527
13.4.1	Ein Modell mit einer Dreifachwechselwirkung .....	528
13.4.2	Ein Modell ohne Dreifachwechselwirkung .....	529
13.4.3	Ein Modell mit nur einer Zweifachwechselwirkung .....	532
13.4.4	Modelle ohne Wechselwirkungen .....	534
13.5	Abschließende Bemerkungen .....	534
13.5.1	Berechnung von Odds und Odds-Ratios .....	534
13.5.2	Modelle mit mehr als drei Variablen .....	535
13.5.3	Vergleich mit logistischer Regression .....	536
13.6	Zusammenfassung der Konzepte .....	536
13.7	Übungen .....	537
13.8	R-Befehle im Überblick .....	537

<b>Teil VI</b>	<b>Appendix</b>	<b>539</b>
<b>Literaturverzeichnis</b>		<b>541</b>
<b>R GUI (Windows)</b>		<b>543</b>
<b>R-Funktionen</b>		<b>545</b>
<b>Index</b>		<b>551</b>