

# Auf einen Blick

---

<b>Über den Autor</b> .....	<b>9</b>
<b>Einführung</b> .....	<b>21</b>
<b>Teil I: Das Handwerkszeug</b> .....	<b>25</b>
<b>Kapitel 1:</b> R: Was R kann und wie R das macht .....	27
<b>Kapitel 2:</b> Mit Packages arbeiten .....	49
<b>Kapitel 3:</b> Daten grafisch darstellen .....	59
<b>Teil II: Interaktion mit dem Anwender</b> .....	<b>95</b>
<b>Kapitel 4:</b> Mit einem Browser arbeiten .....	97
<b>Kapitel 5:</b> Dashboards .....	125
<b>Teil III: Maschinelles Lernen</b> .....	<b>157</b>
<b>Kapitel 6:</b> Werkzeuge und Daten für Projekte für maschinelles Lernen .....	159
<b>Kapitel 7:</b> Entscheidungen, Entscheidungen, Entscheidungen .....	181
<b>Kapitel 8:</b> Auf in den Wald voller randomisierter Bäume .....	199
<b>Kapitel 9:</b> Unterstützen Sie Ihren lokalen Vektor .....	215
<b>Kapitel 10:</b> K-Means-Clusteranalyse .....	237
<b>Kapitel 11:</b> Neuronale Netze .....	255
<b>Teil IV: Große Datensätze</b> .....	<b>271</b>
<b>Kapitel 12:</b> Marketing erforschen .....	273
<b>Kapitel 13:</b> Aus der Stadt, die niemals schläft .....	291
<b>Teil V: Karten und Bilder</b> .....	<b>307</b>
<b>Kapitel 14:</b> Daten auf Karten darstellen .....	309
<b>Kapitel 15:</b> Spaß mit Bildern .....	321
<b>Teil VI: Der Top-Ten-Teil</b> .....	<b>335</b>
<b>Kapitel 16:</b> Mehr als zehn Packages für Ihre R-Projekte .....	337
<b>Kapitel 17:</b> Mehr als zehn nützliche Informationsquellen .....	343
<b>Stichwortverzeichnis</b> .....	<b>347</b>

# Inhaltsverzeichnis

---

<b>Über den Autor</b> .....	<b>9</b>
Widmung .....	9
<b>Einführung</b> .....	<b>21</b>
Über dieses Buch .....	21
Teil 1: Das Handwerkszeug .....	21
Teil 2: Interaktion mit dem Anwender .....	22
Teil 3: Maschinelles Lernen .....	22
Teil 4: Große Datensätze .....	22
Teil 5: Karten und Bilder .....	22
Teil 6: Der Top-Ten-Teil .....	22
Was Sie nicht lesen müssen .....	22
Törichte Annahmen über den Leser .....	23
Symbole, die in diesem Buch verwendet werden .....	23
Wie es weitergeht .....	23
<b>TEIL I</b>	
<b>DAS HANDWERKSZEUG</b> .....	<b>25</b>
<b>Kapitel 1</b>	
<b>R: Was R kann und wie R das macht</b> .....	<b>27</b>
R herunterladen .....	27
RStudio herunterladen .....	28
Eine Session mit R .....	31
Das Arbeitsverzeichnis .....	31
Jetzt geht es richtig los .....	32
R-Funktionen .....	36
Benutzerdefinierte Funktionen .....	37
Kommentare .....	38
R-Strukturen .....	38
Vektoren .....	38
Numerische Vektoren .....	39
Matrizen .....	41
Listen .....	43
Datensätze (Data Frames) .....	44
for-Schleifen und if-Anweisungen .....	47
<b>Kapitel 2</b>	
<b>Mit Packages arbeiten</b> .....	<b>49</b>
Packages installieren .....	49
Daten untersuchen .....	51
Anfang und Ende .....	51

## 14 Inhaltsverzeichnis

Fehlende Daten .....	51
Teilmengen (Sub Sets) .....	52
R-Formeln .....	53
Weitere Packages .....	54
tidyverse erforschen .....	55

### Kapitel 3

#### **Daten grafisch darstellen** ..... **59**

Diagramme mit dem R-Basispaket erstellen .....	59
Histogramme .....	59
Dichteplots .....	61
Säulendiagramme .....	63
Säulen gruppieren .....	66
Schneller Projektvorschlag .....	67
Kreisdiagramme .....	69
Streudiagramme .....	69
Matrix von Streudiagrammen .....	71
Kastengrafik: Kästchen und Antennen .....	72
Zu ggplot2 aufsteigen .....	73
Wie es funktioniert .....	74
Histogramme .....	75
Säulendiagramme .....	78
Gruppierte Säulendiagramme .....	79
Noch einmal Gruppierungen .....	81
Streudiagramme .....	84
Streudiagramme mit Pfiff .....	84
Matrix von Streudiagrammen .....	88
Boxplots .....	90

## TEIL II

### **INTERAKTION MIT DEM ANWENDER** ..... **95**

#### Kapitel 4

#### **Mit einem Browser arbeiten** ..... **97**

The Shiny! .....	97
Ihr erstes shiny-Projekt erstellen .....	98
Die Benutzeroberfläche .....	100
Der Server .....	102
Abschließende Schritte .....	103
Reaktiv werden .....	104
Mit ggplot arbeiten .....	107
Den Server ändern .....	108
Noch ein paar Änderungen .....	110
Mit ggplot reaktiv werden .....	112
Ein weiteres shiny-Projekt .....	114
Die Version mit dem R-Basispaket .....	115
Die ggplot-Version .....	121
Projektvorschlag .....	123

<b>Kapitel 5</b>	
<b>Dashboards</b>	<b>125</b>
Das Package shinydashboard	125
Dashboard-Layouts verstehen	126
Erste Schritte für die Benutzeroberfläche	126
Die Benutzeroberfläche zusammenstellen: Boxen, Boxen, Boxen ...	127
In Spalten ausrichten.	134
Ein cleverer Trick: Tabs für besseren Überblick.	136
Projektvorschlag: Statistiken ergänzen.	140
Projektvorschlag: Fügen Sie valueBoxen in tabPanels ein.	140
Die Seitenleiste verwenden	142
Die Benutzerschnittstelle	143
Der Server	145
Projektvorschlag: Neue Position für den Schieberegler.	147
Mit Grafiken interagieren	149
Klicks, Doppelklicks, Rahmenauswahl.	149
Warum das alles?	152
Projektvorschlag: Mit airquality experimentieren.	155

### TEIL III MASCHINELLES LERNEN **157**

<b>Kapitel 6</b>	
<b>Werkzeuge und Daten für Projekte für maschinelles Lernen</b>	<b>159</b>
Das UCI (University of California-Irvine) ML Repository.	160
Einen UCI-Datensatz herunterladen	160
Die Daten aufräumen	162
Die Daten untersuchen	163
Beziehungen in den Daten untersuchen	166
Einführung in das Package Rattle	171
Rattle für iris verwenden	173
Daten einlesen und weiter untersuchen	173
Cluster in den Daten finden	175

<b>Kapitel 7</b>	
<b>Entscheidungen, Entscheidungen, Entscheidungen</b>	<b>181</b>
Komponenten von Entscheidungsbäumen	181
Wurzel und Blätter	182
Baumerstellung	183
Entscheidungsbäume in R	183
In R den Baum wachsen lassen	184
Den Baum in R zeichnen	185
Entscheidungsbäume in Rattle	187
Den Baum erstellen	188
Den Baum zeichnen	190
Den Baum auswerten	190

Projekt: Ein komplexerer Entscheidungsbaum . . . . .	191
Die Daten: Car Evaluation . . . . .	192
Datenexploration . . . . .	193
Den Baum erstellen und zeichnen . . . . .	194
Den Baum auswerten . . . . .	195
Schneller Projektvorschlag: Den Komplexitätsparameter verstehen. . . . .	196
Projektvorschlag: Titanic. . . . .	197

## Kapitel 8

### **Auf in den Wald voller randomisierter Bäume . . . . . 199**

Einen Random Forest wachsen lassen . . . . .	199
Random Forests in R . . . . .	201
Den Forest erstellen. . . . .	201
Den Forest auswerten . . . . .	202
Genauer hinsehen . . . . .	203
Fehler als Grafik darstellen. . . . .	205
Die Wichtigkeit (importance) als Grafik darstellen . . . . .	206
Projekt: Glas identifizieren . . . . .	207
Die Daten. . . . .	208
Die Daten in Rattle bekommen . . . . .	209
Die Daten untersuchen . . . . .	210
Den Random Forest wachsen lassen . . . . .	210
Die Ergebnisse visualisieren . . . . .	212
Projektvorschlag: Pilze identifizieren . . . . .	213

## Kapitel 9

### **Unterstützen Sie Ihren lokalen Vektor . . . . . 215**

Ein paar Daten, mit denen Sie arbeiten können . . . . .	215
Eine Teilmenge verwenden. . . . .	216
Eine Grenze definieren . . . . .	216
Stützvektoren verstehen. . . . .	217
Trennbarkeit: Sie ist üblicherweise nicht linear . . . . .	219
Support Vector Machines in R . . . . .	222
e1071 einsetzen . . . . .	222
kernlab einsetzen . . . . .	227
Projekt: Parteien im Repräsentantenhaus. . . . .	229
Die Daten einlesen. . . . .	230
Die Daten untersuchen . . . . .	230
Die SVM erstellen . . . . .	231
Die SVM auswerten . . . . .	232
Projektvorschlag: Noch einmal Titanic. . . . .	235

## Kapitel 10

### **K-Means-Clusteranalyse . . . . . 237**

Wie es funktioniert . . . . .	237
K-Means-Clustering in R . . . . .	239
Die Daten vorbereiten und analysieren . . . . .	239
Die Ausgabe verstehen . . . . .	240
Die Cluster visualisieren . . . . .	241

Die optimale Anzahl an Clustern ermitteln .....	242
Schneller Projektvorschlag: Die Sepalen hinzufügen .....	245
Projekt: Glas-Cluster .....	247
Die Daten .....	247
Rattle starten und die Daten untersuchen .....	248
Clusteranalyse vorbereiten .....	249
Clusteranalyse durchführen .....	249
Über Rattle hinaus .....	250
Projektvorschlag: Ein paar schnelle Projekte .....	252
Datenpunkte und Cluster visualisieren .....	252
Die optimale Clusteranzahl .....	252
Variablen hinzufügen .....	253

## Kapitel 11

### Neuronale Netze .....

**255**

Netzwerke im Nervensystem .....	255
Künstliche neuronale Netze .....	256
Überblick .....	256
Eingabeschicht und verdeckte Schicht .....	257
Ausgabeschicht .....	258
Wie alles funktioniert .....	258
Neuronale Netze in R .....	259
Ein neuronales Netz für den Datenframe iris erstellen .....	259
Das neuronale Netz als Grafik darstellen .....	261
Das Netz evaluieren .....	262
Schneller Projektvorschlag: Schon wieder diese Sepalen .....	263
Projekt: Banknoten .....	263
Die Daten .....	263
Ein schneller Blick voraus .....	264
Rattle vorbereiten .....	265
Das Netz auswerten .....	267
Über Rattle hinaus: Das Netz visualisieren .....	267
Projektvorschläge: Experimente mit Rattle .....	269

## TEIL IV

### GROßE DATENSÄTZE .....

**271**

## Kapitel 12

### Marketing erforschen .....

**273**

Projekt: Einzelhandelsdaten analysieren .....	273
Die Daten .....	274
RFM in R .....	275
RFM und maschinelles Lernen .....	282
K-Means-Clustering .....	282
Mit Rattle arbeiten .....	283
Tiefer in die Cluster eintauchen .....	285
Die Cluster und die Klassen .....	287
Schneller Projektvorschlag .....	289
Projektvorschlag: Ein anderer Datensatz .....	289

**Kapitel 13****Aus der Stadt, die niemals schläft** ..... 291

Den Datensatz untersuchen	291
Aufwärmen	292
Ein schneller Blick und genaues Hinsehen	292
Pipelining, Filtern und Gruppieren	293
Visualisieren	295
Mit »join« zusammenführen	296
Schneller Projektvorschlag: Namen der Airlines	299
Projekt: Verspätete Abflüge	299
Eine Variable hinzufügen: Wochentag	299
Schneller Projektvorschlag: Unterschiede zwischen Wochentagen analysieren	300
Verspätung, Wochentag und Airport	301
Verspätung und Flugdauer	305
Projektvorschlag: Verspätung und Wetter	306

**TEIL V****KARTEN UND BILDER** ..... 307**Kapitel 14****Daten auf Karten darstellen** ..... 309

Projekt: Die Airports von Wisconsin	309
Sich um die Voraussetzungen kümmern	309
Die geografischen Daten von Wisconsin abrufen	310
Die geografischen Daten der Flughäfen ermitteln	311
Die Flughäfen auf der Karte des Bundesstaates ausgeben	313
Schneller Projektvorschlag: Andere Quelle für die geografischen Daten der Airports	314
Projektvorschlag 1: Karte eines anderen Bundesstaates der USA	315
Projektvorschlag 2: Karte der USA	315
Die Hauptstädte der Bundesstaaten	317
Die Flughäfen einzeichnen	318

**Kapitel 15****Spaß mit Bildern** ..... 321

Ein Bild aufpolieren: Es ist magick!	321
Die Grafik einlesen	322
Drehen, horizontal und vertikal spiegeln	323
Anmerkungen	324
Transformationen kombinieren	325
Schneller Projektvorschlag: Drei F	326
Bilder kombinieren	326
Animationen	327
Ihre eigenen Morphs erstellen	328

Projekt: Zwei Legenden auf der Suche nach einer Legende .....	329
Stan und Ollie herunterladen .....	330
Die Jungs mit dem Hintergrund kombinieren .....	330
So funktioniert <code>image_apply()</code> .....	331
Zurück zur Animation .....	332
Projektvorschlag: Kombinieren Sie eine Animation und ein Diagramm .....	332
<b>TEIL VI</b>	
<b>DER TOP-TEN-TEIL .....</b>	<b>335</b>
<b>Kapitel 16</b>	
<b>Mehr als zehn Packages für Ihre R-Projekte .....</b>	<b>337</b>
Maschinelles Lernen .....	337
Datenbanken .....	338
Karten .....	338
Bildbearbeitung .....	340
Textanalyse .....	340
<b>Kapitel 17</b>	
<b>Mehr als zehn nützliche Informationsquellen .....</b>	<b>343</b>
Mit Benutzern interagieren .....	343
Maschinelles Lernen .....	344
Datenbanken .....	344
Karten und Bilder .....	345
<b>Stichwortverzeichnis .....</b>	<b>347</b>