

Thomas Rahlf

# Datenvisualisierung mit R

111 Beispiele

2., erweiterte Auflage

 **Springer** Spektrum

# Inhaltsverzeichnis

<b>1</b>	<b>Daten für alle</b> . . . . .	1
1.1	Datenvisualisierung zwischen Wissenschaft und Journalismus . . .	1
1.2	Warum R? . . . . .	3
1.3	Das Konzept des Datendesigns . . . . .	4

## Teil I Grundlagen und Technik

<b>2</b>	<b>Aufbau und technische Voraussetzungen</b> . . . . .	9
2.1	Begriffe und Elemente . . . . .	9
2.2	Gestaltungsraster . . . . .	9
2.3	Perzeption . . . . .	12
2.4	Schriften . . . . .	16
2.4.1	Fonts . . . . .	18
2.4.2	Freie Schriften . . . . .	19
2.5	Symbole . . . . .	20
2.5.1	Symbolfonts . . . . .	22
2.5.2	Symbole im SVG-Format . . . . .	24
2.6	Farbe . . . . .	24
2.6.1	Farbmodelle . . . . .	24
2.6.2	Farbe in statistischen Abbildungen . . . . .	26
<b>3</b>	<b>Umsetzung in R</b> . . . . .	29
3.1	Installation . . . . .	29
3.2	Grundkonzepte in R . . . . .	30
3.2.1	Datenstrukturen . . . . .	31
3.2.2	Import von Daten . . . . .	35
3.3	Grafikkonzepte in R . . . . .	43
3.3.1	Paper-Pencil-Prinzip des Base Graphics System: High-Level- und Low-Level-Funktionen . . . . .	48
3.3.2	Einstellung von Grafikparametern . . . . .	51

3.3.3	Randeinstellungen von Abbildungen und Grafiken . . . . .	58
3.3.4	Mehrfachgrafiken: Panels mit <code>mfrow</code> und <code>mfcol</code> . . . . .	59
3.3.5	Komplexere Anordnungen mit <code>layout</code> . . . . .	60
3.3.6	Schrifteinbindung . . . . .	63
3.3.7	Ausgabe mit <code>cairo_pdf</code> . . . . .	65
3.3.8	Unicode in Abbildungen . . . . .	66
3.3.9	Farbeeinstellungen . . . . .	70
3.4	R-Pakete und -Funktionen in diesem Buch . . . . .	72
3.4.1	Pakete . . . . .	73
3.4.2	Funktionen . . . . .	77
3.4.3	Schematische Vorgehensweise . . . . .	85
<b>4</b>	<b>Über R hinaus</b> . . . . .	87
4.1	Ergänzungen mit LaTeX . . . . .	87
4.2	Manuelle Nachbearbeitung und Symbolfonterstellung in Inkscape . . . . .	92
4.2.1	Nachbearbeitung . . . . .	92
4.2.2	Symbolfonterstellung . . . . .	94
<b>5</b>	<b>Zu den Beispielen</b> . . . . .	99
5.1	Versuch einer Systematik . . . . .	99
5.2	Die Skripte zum Laufen bringen . . . . .	101

## Teil II Beispiele

<b>6</b>	<b>Kategoriale Daten</b> . . . . .	105
6.1	Balken- und Säulendiagramme . . . . .	105
6.1.1	Balkendiagramm einfach . . . . .	106
6.1.2	Balkendiagramm für Mehrfachantworten – die ersten beiden Antwortkategorien . . . . .	111
6.1.3	Balkendiagramm für Mehrfachantworten – alle Antwortkategorien . . . . .	116
6.1.4	Balkendiagramm für Mehrfachantworten – alle Antwortkategorien, Variante . . . . .	119
6.1.5	Balkendiagramm für Mehrfachantworten – alle Antwortkategorien (Panel) . . . . .	121
6.1.6	Balkendiagramm für Mehrfachantworten – Symbole für Individuen . . . . .	124
6.1.7	Balkendiagramm für Mehrfachantworten – alle Antwortkategorien, gruppiert . . . . .	127
6.1.8	Säulendiagramm mit zweizeiliger Beschriftung . . . . .	133
6.1.9	Säulendiagramm mit 45-Grad-Beschriftung . . . . .	135
6.1.10	Profildiagramm für Mehrfachantworten – Mittelwerte der Antworten . . . . .	136
6.1.11	Dotchart für drei Variablen . . . . .	138

6.1.12	Säulendiagramm mit Anteilen . . . . .	142
6.2	Kreis- und Radialdiagramme . . . . .	145
6.2.1	Einfaches Kreisdiagramm . . . . .	146
6.2.2	Kreisdiagramme, Beschriftung innen (Panel) . . . . .	148
6.2.3	Sitzverteilung (Panel) . . . . .	150
6.2.4	Spie Chart . . . . .	153
6.2.5	Radialpolygone (Panel) . . . . .	156
6.2.6	Radialpolygone (Panel) – andere Spaltenanordnung . . . . .	158
6.2.7	Radialpolygone übereinander . . . . .	159
6.3	Grafiktabellen . . . . .	160
6.3.1	Vereinfachtes Gantt-Diagramm . . . . .	161
6.3.2	Vereinfachtes Gantt-Diagramm – Farben nach Personen . . . . .	165
6.3.3	Bumpchart . . . . .	167
6.3.4	Heatmap . . . . .	170
6.3.5	Mosaikplot (Panel) . . . . .	173
6.3.6	Ballonplot . . . . .	175
6.3.7	Treemap . . . . .	177
6.3.8	Treemaps für zwei Ebenen (Panel) . . . . .	180
6.4	Netzwerkbeziehungen . . . . .	184
6.4.1	Ungerichtetes Netzwerk . . . . .	185
6.4.2	Chord Diagram . . . . .	190
6.4.3	Gerichtetes Netzwerk . . . . .	195
6.4.4	Riverplot . . . . .	198
6.4.5	Heatmap für Beziehungen . . . . .	201
6.4.6	Multiple Barchart . . . . .	204
<b>7</b>	<b>Verteilungen</b> . . . . .	<b>207</b>
7.1	Histogramme und Boxplots . . . . .	207
7.1.1	Histogramme übereinander . . . . .	207
7.1.2	Säulendiagramme mit Colorbrewer gefärbt (Panel) . . . . .	209
7.1.3	Histogramme (Panel) . . . . .	213
7.1.4	Boxplots für Gruppen – absteigend sortiert . . . . .	215
7.1.5	Boxplots für Gruppen – absteigend sortiert, Vergleich zweier Erhebungen . . . . .	219
7.2	(Bevölkerungs-)Pyramiden . . . . .	223
7.2.1	Pyramide mit mehreren Farben . . . . .	225
7.2.2	Pyramiden – Betonung der äußeren Bereiche (Panel) . . . . .	228
7.2.3	Pyramiden – Betonung der inneren Bereiche (Panel) . . . . .	231
7.2.4	Pyramiden mit eingezeichneter Linie (Panel) . . . . .	234
7.2.5	Pyramide mit Zusammenfassungen . . . . .	235
7.2.6	Balkendiagramme als Pyramiden (Panel) . . . . .	238
7.3	Ungleichheit . . . . .	241
7.3.1	Einfache Lorenzkurve . . . . .	242
7.3.2	Lorenzkurven übereinander . . . . .	244
7.3.3	Lorenzkurven (Panel) . . . . .	246

7.3.4	Vergleich von Einkommensanteilen mit Balkendiagramm (Quintile) . . . . .	249
7.3.5	Vergleich von Einkommensanteilen mit Balkendiagramm (Dezile) . . . . .	251
7.3.6	Vergleich von Einkommensanteilen mit Panel-Balkendiagramm (Quintile) . . . . .	253
<b>8</b>	<b>Zeitreihen</b> . . . . .	<b>257</b>
8.1	Kurze Zeitreihen . . . . .	257
8.1.1	Säulendiagramm für Entwicklungen . . . . .	257
8.1.2	Säulendiagramm mit Anteilen für Wachstumsentwicklungen	260
8.1.3	Quartalswerte als Säulen . . . . .	263
8.1.4	Quartalswerte als Linien mit Werte-Beschriftungen . . . . .	265
8.1.5	Kurze Zeitreihen übereinander . . . . .	267
8.2	Flächen unter und zwischen Zeitreihen . . . . .	269
8.2.1	Flächen zwischen zwei Zeitreihen . . . . .	269
8.2.2	Fläche als Korridor mit Zeitreihen (Panel) . . . . .	271
8.2.3	Prognoseintervalle (Panel) . . . . .	274
8.2.4	Prognoseintervalle Index (Panel) . . . . .	278
8.2.5	Zeitreihen mit gestapelten Flächen . . . . .	280
8.2.6	Flächen unterhalb einer Zeitreihe . . . . .	283
8.2.7	Zeitreihen mit Trend (Panel) . . . . .	286
8.3	Darstellung von Tages-, Wochen- und Monatswerten . . . . .	289
8.3.1	Tageswerte mit Beschriftungen . . . . .	289
8.3.2	Tageswerte mit Beschriftungen und Wochensymbolen (Panel) . . . . .	291
8.3.3	Tageswerte mit Monatsbeschriftung . . . . .	296
8.3.4	Zeitreihen aus Wochenwerten (Panel) . . . . .	298
8.3.5	Monatswerte (Panel) . . . . .	301
8.3.6	Monatswerte mit Monatsbeschriftung . . . . .	303
8.3.7	Monatswerte mit Monatsbeschriftung (Layout) . . . . .	306
8.4	Sonderfälle und Spezielles . . . . .	309
8.4.1	Zeitreihen als Streudiagramm (Panel) . . . . .	309
8.4.2	Zeitreihen mit fehlenden Werten . . . . .	311
8.4.3	Saisonspannweiten (Panel) . . . . .	315
8.4.4	Saisonspannweiten übereinander . . . . .	317
8.4.5	Saisonfigur (Seasonal Subseries Plot) mit Datentabelle . . . . .	319
8.4.6	Zeitliche Spannweiten . . . . .	323
<b>9</b>	<b>Streudiagramme</b> . . . . .	<b>325</b>
9.1	Varianten . . . . .	327
9.1.1	Streudiagramm Variante 1: Vier Quadranten farblich unterschieden . . . . .	327
9.1.2	Streudiagramm Variante 2: Ausreißer farblich hervorgehoben . . . . .	330

- 9.1.3 Streudiagramm Variante 3: Bereiche farblich hervorgehoben 333
- 9.1.4 Streudiagramm Variante 4: Eingezeichnete Ellipse . . . . . 335
- 9.1.5 Streudiagramm Variante 5: Verbundene Punkte . . . . . 338
- 9.2 Sonderfälle und Spezielles . . . . . 340
  - 9.2.1 Streudiagramm mit wenigen Punkten . . . . . 340
  - 9.2.2 Streudiagramm mit selbst definierten Symbolen . . . . . 343
  - 9.2.3 Karte von Deutschland als Streudiagramm . . . . . 346
- 10 Karten . . . . . 349**
  - 10.1 Einführende Beispiele . . . . . 349
    - 10.1.1 Karten von Deutschland: Ortsnetzbereiche und Postleitzahlengebiete . . . . . 349
    - 10.1.2 Gefilterte Postleitzahlenkarte . . . . . 352
    - 10.1.3 Europakarte Nuts 2006 (Ausschnitt) . . . . . 354
  - 10.2 Punkte, Diagramme und Symbole in Karten . . . . . 356
    - 10.2.1 Karte von Deutschland mit ausgewählten Orten und Umriss (Panel) . . . . . 356
    - 10.2.2 Karte von Deutschland mit ausgewählten Orten (Kreisdiagramme) und Umriss . . . . . 359
    - 10.2.3 Karte von Deutschland mit ausgewählten Orten (Säulen) und Umriss . . . . . 362
    - 10.2.4 Karte von Deutschland als dreidimensionales Streudiagramm . . . . . 365
    - 10.2.5 Karte von Nordrhein-Westfalen mit ausgewählten Orten (Symbole) und Umriss . . . . . 369
    - 10.2.6 Karte von Tunesien mit selbst definierten Symbolen . . . . . 371
  - 10.3 Choroplethenkarten . . . . . 374
    - 10.3.1 Choroplethenkarte von Deutschland auf Kreisebene . . . . . 375
    - 10.3.2 Choroplethenkarte von Deutschland auf Kreisebene (Panel) 377
    - 10.3.3 Choroplethenkarte von Europa auf Länderebene . . . . . 383
    - 10.3.4 Choroplethenkarte von Europa auf Länderebene (Panel) . . 386
    - 10.3.5 Weltchoroplethenkarte: Regionen . . . . . 389
  - 10.4 Sonderfälle und Spezielles . . . . . 391
    - 10.4.1 Weltkarte mit Orthodromen . . . . . 392
    - 10.4.2 Stadtkarten mit OpenStreetMap-Daten (Panel) . . . . . 394
    - 10.4.3 Georeferenzierte Karte im Rasterformat . . . . . 400
    - 10.4.4 Cartogram (Panel) . . . . . 407
- 11 Illustratives . . . . . 411**
  - 11.1 Tabelle mit Symbolen der Schrift „Symbol Signs“ . . . . . 411
  - 11.2 Radialsäulendiagramme mit Beschriftung (Panel) . . . . . 413
  - 11.3 Radialsäulendiagramme ohne Beschriftung (Panel) . . . . . 420
  - 11.4 Radialsäulendiagramm (Poster) . . . . . 423
  - 11.5 Nacht-Karte von Deutschland als Streudiagramm . . . . . 428
  - 11.6 Streudiagramm Gampminder . . . . . 430

11.7 Karte von Napoleons Rußlandfeldzug von 1812/13 von Charles Joseph Minard, 1869 . . . . .	435
<b>12 Interaktive Visualisierung mit JavaScript: Highcharts und Mapael .</b>	<b>439</b>
12.1 Streudiagramm in Highcharts . . . . .	441
12.2 Zeitreihe in Highcharts . . . . .	449
12.3 Choroplethenkarten mit Mapael . . . . .	456
12.3.1 Installation der Javascript-Bibliotheken von Mapael . . . . .	458
<b>Anhang A. Verwendete Daten . . . . .</b>	<b>477</b>
<b>Literatur . . . . .</b>	<b>483</b>
<b>Sachverzeichnis . . . . .</b>	<b>487</b>