LEHRBUCH

Bernd Heesen

Data Science und Statistik mit R

Anwendungslösungen für die Praxis

Springer Gabler

Inhaltsverzeichnis

1 Einleitung 1
1.1 Data Science und Big Data 2
1.2 Business Intelligence 4
1.3 Engpass Data-Science-Kompetenz 5
Literatur 6

2 Data Science 7
2.1 Standards für das Projektmanagement 8
2.2 Standards für die Visualisierung 9
2.3 Standards für den Datenaustausch 11
Literatur 14

3 Data Science mit R 15
3.1 Installation einer R-Umgebung 16
3.1.1 R installieren 16
3.1.2 R-Tools installieren 22
3.1.3 RStudio installieren 25
3.2 R Grundlagen 30
3.2.1 Starten und Beenden von R 30
3.2.2 Arbeiten mit Packages 32
3.2.3 Installation des Pakets datascience 33
3.3 R-Syntax 34
3.3.1 Datentypen 35
3.3.2 Datenstrukturen 41
3.3.3 Operationen 64
3.3.4 Funktionen 85
3.3.5 Kontrollstrukturen 92
3.4 Import und Export von Daten 97
3.4.1 Umgang mit Dateien 97
3.4.2 Import von Daten 98
3.4.3 Export von Daten 100

3.5 Transformation von Daten . . . 106
3.5.1 Konventionen für die Transformation . . . 107
3.5.2 Funktionen für die Transformation . . . 111
3.6 Visualisierung . . . 132
3.6.1 Farben . . . 150
3.6.2 Landkarten . . . 160
3.6.3 Interaktive Landkarten . . . 170
3.6.4 Dynamische Abbildungen . . . 174
3.7 Fallstudie . . . 175
Literatur . . . 209

4 Statistik . . . 211
4.1 Deskriptive Statistik . . . 214
4.1.1 Lagemaße . . . 216
4.1.2 Streuungsmaße . . . 219
4.1.3 Univariate Statistik und Verteilungen . . . 228
4.1.4 Bivariate und multivariate Statistik . . . 234
4.1.5 Visualisierung . . . 251
4.2 Wahrscheinlichkeitsrechnung . . . 262
4.2.1 Zufallsvariablen . . . 263
4.2.2 Kombinationen . . . 275
4.2.3 Zufallszahlen und Zufallswerte . . . 278
4.2.4 Wahrscheinlichkeitsverteilungen . . . 280
4.2.5 Z-Standardisierung . . . 290
4.2.6 Bedingte Wahrscheinlichkeit . . . 297
4.2.7 Bernoulli-Experiment . . . 302
4.3 Induktive Statistik . . . 305
4.3.1 Stichproben . . . 306
4.3.2 Konfidenzintervalle . . . 316
4.3.3 Hypothesentests . . . 323
4.3.4 Regression . . . 350
4.3.5 Clustering . . . 360
Literatur . . . 365

5 Ausblick Machine Learning & Künstliche Intelligenz . . . 367
Literatur . . . 369

Anhang . . . 371