

Inhalt

Vorwort	XI
----------------------	-----------

Teil I Erkunden

1 Datenvisualisierung mit ggplot2	3
Einführung	3
Erste Schritte	4
Zuordnungen visueller Eigenschaften	7
Häufige Probleme	13
Facetten	14
Geometrische Objekte	16
Statistische Transformationen	23
Positionsanpassungen	28
Koordinatensysteme	32
Die »Layered Grammar of Graphics«	35
2 Workflow: Grundlagen	37
Grundlagen der Codierung	37
Was macht einen Namen aus?	38
Funktionen aufrufen	39
3 Datentransformation mit dplyr	41
Einführung	41
Zeilen mit filter() filtern	43
Zeilen mit arrange() anordnen	47
Spalten mit select() auswählen	49
Neue Variablen mit mutate() hinzufügen	51
Gruppierte Zusammenfassungen mit summarize()	55
Gruppierte Veränderungen (und Filter)	69

4	Workflow: Skripte	73
	Code ausführen	74
	RStudio-Fehlerdiagnose	75
5	Explorative Datenanalyse	77
	Einführung	77
	Fragen	78
	Variation	79
	Fehlende Werte	87
	Kovariation	89
	Muster und Modelle	101
	ggplot2-Aufrufe	104
	Mehr lernen	104
6	Workflow: Projekte	105
	Was ist real?	105
	Wo ist Ihre Analyse untergebracht?	106
	Pfade und Verzeichnisse	107
	RStudio-Projekte	108
	Zusammenfassung	110

Teil II Aufbereiten

7	Tibbles mit tibble	113
	Einführung	113
	Tibbles erzeugen	113
	Tibbles versus data.frame	115
	Teilgruppen	116
	Mit älterem Code arbeiten	117
8	Datenimport mit readr	119
	Einführung	119
	Erste Schritte	119
	Einen Vektor parsen	122
	Eine Datei parsen	130
	In eine Datei schreiben	135
	Andere Datentypen	137
9	Daten aufbereiten mit tidyr	139
	Einführung	139
	Aufbereitete Daten	140
	Ausbreiten und Zusammenziehen	143

Aufteilen und Vereinigen	147
Fehlende Werte	151
Fallstudie	153
Unaufbereitete Daten	158
10 Relationale Daten mit dplyr	159
Einführung	159
nycflights13	160
Schlüssel	162
Verändernde Verknüpfungen	165
Filternde Verknüpfungen	175
Verknüpfungsprobleme	178
Mengenoperationen	178
11 Strings mit stringr	181
Einführung	181
Grundlagen von Strings	181
Musterübereinstimmung mit regulären Ausdrücken	185
Tools	192
Andere Mustertypen	203
Andere Verwendungen von regulären Ausdrücken	206
stringi	206
12 Faktoren mit forcats	207
Einführung	207
Faktoren erzeugen	207
General Social Survey	209
Faktorenreihenfolge ändern	211
Faktorstufen modifizieren	216
13 Datum und Uhrzeit mit lubridate	219
Einführung	219
Datums-/Zeitwerte erzeugen	220
Datums-/Zeitkomponenten	224
Zeiträume	230
Zeitzone	235

Teil III Programmieren

14 Pipes mit magrittr	243
Einführung	243
Alternativen zu Pipes	243

Wann man Pipes nicht verwenden sollte	247
Andere Tools von magrittr	248
15 Funktionen	251
Einführung	251
Wann sollte man eine Funktion schreiben?	252
Funktionen – nützlich für Menschen und Computer	255
Bedingte Ausführung	258
Funktionsargumente	262
Rückgabewerte	267
Umgebung	269
16 Vektoren	271
Einführung	271
Grundlagen von Vektoren	272
Wichtige Typen atomarer Vektoren	273
Atomare Vektoren verwenden	276
Rekursive Vektoren (Listen)	281
Attribute	286
Erweiterte Vektoren	287
17 Iteration mit purrr	291
Einführung	291
for-Schleifen	292
Varianten von for-Schleifen	294
Die map-Funktionen	302
Fehlerverarbeitung	306
Zuordnungen über mehrere Argumente	308
Walk	311
Andere Muster von for-Schleifen	312

Teil IV Modellieren

18 Grundlagen der Modellierung mit modelr	321
Einführung	321
Ein einfaches Modell	323
Modelle visualisieren	331
Formel und Modellfamilien	335
Fehlende Werte	348
Andere Modellfamilien	348

19 Modelle erstellen	351
Einführung	351
Warum sind Diamanten geringer Qualität teurer als höherwertige? ...	352
Was beeinflusst die Anzahl der täglichen Flüge?	361
Mehr über Modelle lernen	373
20 Viele Modelle mit purrr und broom	375
Einführung	375
Gapminder	376
Listenspalten	386
Listenspalten erzeugen	388
Listenspalten vereinfachen	392
Daten bereinigen mit broom	395

Teil V Kommunizieren

21 R Markdown	399
Einführung	399
R-Markdown-Grundlagen	400
Textformatierung mit Markdown	403
Codebereiche	404
Fehlerbehebung	410
YAML-Header	410
Mehr lernen	413
22 Grafik für Kommunikation mit ggplot2	415
Einführung	415
Beschriftungen	416
Anmerkungen	419
Skalen	425
Zoomen	436
Themen	438
Diagramme speichern	440
Mehr lernen	443
23 R-Markdown-Formate	445
Einführung	445
Ausgabeoptionen	446
Dokumente	446
Notebooks	447
Präsentationen	448

Dashboards	448
Interaktivität	450
Websites	452
Andere Formate	453
Mehr lernen	453
24 R-Markdown-Workflow	455
Index	457