

1	Einleitung	1
	Wovon handelt dieses Buch?	1
	Warum Python zur Datenanalyse?	2
	Essenzielle Python-Bibliotheken	4
	Installation und Einrichtung	7
	Community und Konferenzen	10
	Navigation durch dieses Buch	10
	Typografische Konventionen	12
	Benutzung von Codebeispielen	13
	Danksagungen	13
2	Einführende Beispiele	15
	1.USA.gov-Daten von bit.ly	16
	MovieLens 1M Data Set	25
	US-Babynamen von 1880 bis 2010	31
	Fazit und der Weg vor Ihnen	43
3	IPython: Eine interaktive Rechen- und Entwicklungsumgebung	45
	IPython-Grundlagen	46
	Verwenden der Befehlschronik	58
	Interaktion mit dem Betriebssystem	61
	Werkzeuge zur Softwareentwicklung	63
	IPython HTML Notebook	74
	Tipps zur produktiven Codeentwicklung mit IPython	75
	Fortgeschrittene Features von IPython	78
	Danksagung	80

4	Grundlagen von NumPy: Arrays und vektorisierte Berechnung	81
	Das ndarray von NumPy: ein mehrdimensionales Array-Objekt	82
	Universelle Funktionen: Schnelle elementweise Array-Funktionen	97
	Datenverarbeitung mit Arrays	99
	Dateiein- und -ausgabe bei Arrays	106
	Lineare Algebra	107
	Erzeugen von Zufallszahlen	109
	Beispiel: Random Walks	110
5	Erste Schritte mit pandas	115
	Einführung in die Datenstrukturen von pandas	116
	Essenzielle Funktionalität	126
	Zusammenfassen und Berechnen deskriptiver Statistik	141
	Behandlung fehlender Daten	147
	Hierarchisches Indizieren	152
	Weitere Themen in pandas	157
6	Laden und Speichern von Daten sowie Dateiformate	161
	Lesen und Schreiben von Daten im Textformat	161
	Binäre Datenformate	177
	Interagieren mit HTML und Web-APIs	179
	Interaktion mit Datenbanken	180
7	Datenaufbereitung: Säubern, Transformieren, Verknüpfen und Umformen	183
	Kombinieren und Verknüpfen von Datensätzen	183
	Umformen und Transponieren	196
	Transformieren von Daten	201
	Manipulation von Strings	213
	Fallstudie: Die USDA-Nahrungsmitteldatenbank	220
8	Plotten und Visualisieren	227
	Kurze Einführung in die matplotlib-API	228
	Plotten von Funktionen in pandas	241
	Karten zeichnen: Visualisieren der Daten aus der Erdbebenkrise auf Haiti	252
	Das Ökosystem der Visualisierungstools in Python	258
9	Aggregation von Daten und Gruppenoperationen	261
	GroupBy-Mechanismen	262
	Aggregation von Daten	270
	Gruppenweise Operationen und Transformationen	275

Pivot-Tabellen und Kreuztabellierung	285
Fallstudie: Die Datenbank des US-Bundeswahlausschusses von 2012	288
10 Zeitreihen	299
Datentypen und Werkzeuge für Datum und Zeit	300
Grundlagen von Zeitreihen	303
Datumsbereiche, Frequenzen und Verschiebungen	307
Berücksichtigung von Zeitzonen	313
Perioden und Arithmetik von Perioden	317
Resampling und Konvertieren von Frequenzen	323
Plotten von Zeitreihen	330
Funktionen mit gleitenden Fenstern	332
Notizen zu Rechengeschwindigkeit und Speichernutzung	338
11 Anwendungen auf Daten aus Finanzwelt und Ökonomie	341
Themen aus der Datenklempnerei	341
Gruppentransformationen und Analyse	353
Weitere Anwendungsbeispiele	358
12 NumPy für Fortgeschrittene	367
Interns von ndarray-Objekten	367
Fortgeschrittene Manipulation von Arrays	369
Broadcasting	377
Fortgeschrittene Nutzung von ufuncs	382
Strukturierte und Record-Arrays	385
Mehr zum Thema Sortieren	387
Die Klasse Matrix in NumPy	392
Ein- und Ausgabe von Arrays für Fortgeschrittene	393
Tipps für höhere Leistung	395
Anhang: Grundlagen der Programmiersprache Python	399
Index	449