

Inhalt

Vorwort	XVII
Danksagung	XVIII
1 Einleitung	1
1.1 Die richtige Wahl	1
1.2 Aufbau des Buches	2
1.3 Dieses Buch und die Werkzeuge dahinter	3
1.4 Download der Beispiele.....	3
1.5 Über den Autor	4
1.6 Anregungen und Kritik	4
2 Numerisches Programmieren	5
2.1 Überblick	5
2.2 Zusammenhang zwischen Python, NumPy, Matplotlib, SciPy und Pandas	6
2.3 Python – eine Alternative zu MATLAB	7
3 Installation von NumPy, Matplotlib, Pandas und JupyterLab	9
3.1 Einleitung	9
3.2 Installation mit conda und Miniconda	10
3.3 Installation mit pip	11
3.4 Starten von JupyterLab	11
3.5 Warum JupyterLab?	12

Teil I	NumPy	13
4	NumPy Einführung	15
4.1	Überblick	15
4.1.1	Was ist NumPy?	15
4.1.2	Ein einfaches Beispiel	16
4.2	Vergleich NumPy-Datenstrukturen und Listen	17
4.2.1	Zentrale Unterschiede	17
4.2.2	Speicherbedarf	18
4.2.3	Zeitvergleich zwischen Listen und NumPy-Arrays	21
5	Erzeugung und Struktur von Arrays	23
5.1	Dimensionen	23
5.1.1	Nulldimensionale Arrays in NumPy	23
5.1.2	Eindimensionales Array	24
5.1.3	Zwei- und mehrdimensionale Arrays	24
5.2	Gestalt eines Arrays	25
5.3	Indizierung und Teilbereichsoperator	26
5.4	Dreidimensionale Arrays	33
5.5	Array-Erzeugungsfunktionen	35
5.5.1	arange	36
5.5.2	linspace	38
5.6	Arrays mit Nullen und Einsen	39
5.7	Einheitsmatrix	41
5.7.1	Die identity-Funktion	41
5.7.2	Die eye-Funktion	42
5.8	Datentypen	43
5.9	Arrays kopieren	45
5.9.1	numpy.copy(A) und A.copy()	45
5.9.2	Zusammenhängend gespeicherte Arrays	46
5.10	Aufgaben	48
6	Datentyp-Objekt: dtype	51
6.1	dtype	51
6.2	Strukturierte Arrays	53

6.3	Ein- und Ausgabe von strukturierten Arrays	56
6.4	Unicode-Strings in Arrays	58
6.5	Umbenennen von Spaltennamen	59
6.6	Spaltenwerte austauschen	59
6.7	Komplexeres Beispiel	60
6.8	Aufgaben	62
7	Arrays kombinieren und umformen.....	63
7.1	Reduktion und Umformung von Arrays	64
7.1.1	flatten	64
7.1.2	ravel	65
7.1.3	Unterschiede zwischen ravel und flatten.....	65
7.1.4	reshape	66
7.2	Hinzufügen von Dimensionen	68
7.3	Konkatenation und Stapelung von Arrays	69
7.3.1	concatenate	69
7.3.2	stack.....	72
7.3.3	dstack.....	74
7.3.4	vstack.....	78
7.3.5	hstack.....	79
7.4	dsplit	81
7.5	Arrays wiederholen mit tile.....	82
7.6	Aufgaben	85
8	Numerische Operationen auf NumPy-Arrays.....	87
8.1	Operationen mit Skalaren	87
8.2	Operationen zwischen und auf Arrays.....	89
8.3	Matrizenmultiplikation und Skalarprodukt	90
8.3.1	Definition der dot-Funktion.....	90
8.3.2	Beispiele zur dot-Funktion	91
8.3.3	Das dot-Produkt im dreidimensionalen Fall	92
8.4	Vergleichsoperatoren	97
8.5	Logische Operatoren.....	98
8.6	Broadcasting	99

8.6.1	Zeilenweises Broadcasting	100
8.6.2	Spaltenweises Broadcasting.....	103
8.6.3	Broadcasting von zwei eindimensionalen Arrays	106
8.7	Distanzmatrix	106
8.8	ufuncs	108
8.8.1	Anwendung von ufuncs	108
8.8.2	Parameter für Rückgabewerte bei ufuncs	110
8.8.3	accumulate	113
8.8.4	reduce	115
8.8.5	outer	116
8.8.6	at	117
8.9	Aufgaben	117
9	Statistik und Wahrscheinlichkeiten	119
9.1	Einführung	119
9.2	Auf dem random-Modul aufbauende Funktionen	120
9.2.1	Echte Zufallszahlen	121
9.2.2	Erzeugen einer Liste von Zufallszahlen	122
9.2.3	Zufällige Integer-Zahlen	123
9.2.4	Stichproben oder Auswahlen	124
9.2.5	Zufallsintervalle	125
9.2.6	Seed oder Startwert	125
9.2.7	Gewichtete Zufallsauswahl	127
9.2.8	Stichproben mit Python	129
9.2.9	Kartesische Auswahl	132
9.2.10	Kartesisches Produkt	132
9.2.11	Kartesische Auswahl: cartesian_choice	132
9.2.12	Gauss'sche Normalverteilung	135
9.2.13	Übung mit Binärsender	138
9.3	Das random-Untermodule von NumPy	141
9.3.1	Integers und Floats zufällig erzeugen	141
9.3.2	numpy.random.choice	143
9.3.3	numpy.random.random_sample.....	145
9.4	Synthetische Verkaufszahlen	147
9.5	Aufgaben	148

10	Boolesche Maskierung und Indizierung	151
10.1	Fancy-Indizierung	153
10.2	Indizierung mit einem Integer-Array.....	154
10.3	nonzero und where	154
10.4	Beispielanwendungen mit np.where.....	156
10.5	Aufgaben	158
11	Lesen und Schreiben von Daten-Dateien	159
11.1	Textdateien speichern mit savetxt.....	160
11.2	Textdateien laden mit loadtxt.....	162
11.2.1	loadtxt ohne Parameter	162
11.2.2	Spezielle Trenner.....	162
11.2.3	Selektives Einlesen von Spalten	163
11.2.4	Datenkonvertierung beim Einlesen	163
11.3	tofile	165
11.4	fromfile	166
11.5	Empfohlene Methoden	168
11.6	Eine weitere Möglichkeit: genfromtxt	169
Teil II	Matplotlib	171
12	Einführung	173
12.1	Ein erstes Beispiel	174
12.2	Format-Parameter von plot	175
12.3	Mehrere Datenreihen mit Achsentiteln	177
13	Objektorientiert plotten	179
13.1	Erzeugung einer Figure und Axes	181
13.2	Achsenbeschriftungen und Titel	183
13.3	Die Plot-Methode	184
13.4	Wertebereiche der Achsen	185
13.5	Plotten mehrerer Funktionen.....	188
13.6	Streudiagramme	189
13.7	Flächen einfärben	192
13.8	Aufgaben	195

14	Mehrfache Plots und Doppelachsen	197
14.1	Subplots mit subplot	198
14.2	Flexible Layouts mit GridSpec	206
14.3	Doppelachsen	213
14.4	Aufgaben	215
15	Achsen- und Skalenteilung	217
15.1	Achsen und Spines	217
15.2	Achsenbeschriftungen ändern	222
15.3	Justierung der Tick-Beschriftungen	223
16	Legenden und Annotationen	225
16.1	Legende hinzufügen	225
16.2	Annotations / Anmerkungen	229
16.3	Aufgaben	236
17	Konturplots	237
17.1	Erstellen eines Maschengitters	238
17.2	Funktionen auf Meshgrids	240
17.3	contour ohne meshgrid	242
17.4	Linienstil und Farben anpassen	242
17.5	Gefüllte Konturen	244
17.6	Individuelle Farben	245
17.7	Schwellen	246
17.8	Andere Grids	247
17.8.1	Meshgrid genauer	248
17.8.2	mgrid	250
17.8.3	ogrid	250
17.9	imshow	252
17.10	Aufgaben	253
18	Histogramme und Diagramme	255
18.1	Histogramme	256
18.2	Säulendiagramm	260
18.3	Balkendiagramme	262

18.4	Gruppierte Balkendiagramme	263
18.5	xkcd-Modus.....	267
18.6	Tortendiagramme	269
18.7	Stapeldiagramme	270
18.8	Aufgaben	271
 Teil III Pandas		273
19	Pandas:Series	275
19.1	Grundlagen der Series-Datenstruktur	276
19.2	Zugriff und Indizierung.....	279
19.3	Wertmanipulation mit apply	280
19.4	Series aus Dictionaries	281
19.5	NaN – Fehlende Daten.....	282
19.5.1	Fehlende Werte prüfen	283
19.5.2	Zusammenhang zwischen NaN und None	284
19.5.3	Fehlende Daten filtern	285
19.5.4	Fehlende Daten auffüllen	285
19.5.5	Vergleich verschiedener Interpolationsmethoden.....	288
19.6	Aufgaben	289
20	DataFrame	291
20.1	Ein erstes Beispiel	292
20.2	Zusammenhang zu Series	293
20.3	Manipulation der Spaltennamen	294
20.4	DataFrames aus Dictionaries	295
20.5	Zugriff auf Spalten	298
20.6	Selektion von Zeilen	298
20.6.1	loc	298
20.6.2	query.....	300
20.7	Modifikation von DataFrames	302
20.7.1	Spalten einfügen	303
20.7.2	Spalten austauschen.....	307

20.7.3	Zeilen austauschen	308
20.7.4	Einzelne Werte mittels at und iat ändern.....	308
20.8	Index ändern.....	309
20.8.1	Umsortierung von Spalten und Index	310
20.8.2	Spalten umbenennen	312
20.8.3	Spalte als Index verwenden.....	312
20.9	Summen und kumulative Summen	313
20.9.1	Leere Spalten und nachträgliches Befüllen	315
20.10	Sortierung.....	317
20.11	Aufgaben	318
21	Styling	321
21.1	Einleitung	321
21.2	Daten und Präsentation trennen.....	322
21.3	Die .style-Eigenschaft	323
21.3.1	Grundlegende Formatierung mit .format	323
21.4	Maximalwerten in Zeilen und Spalten	324
21.5	Anwenden eines Farbverlaufs	325
21.5.1	Balkendiagramme innerhalb von Zellen anwenden	326
21.6	Aufgaben	328
22	Dateiverarbeitung	331
22.1	DSV- / CSV-Dateien	332
22.1.1	CSV- und DSV-Dateien lesen.....	332
22.1.2	Schreiben von CSV-Dateien	334
22.1.3	Einfaches Beispiel	334
22.1.4	Umfangreicheres Beispiel	336
22.1.5	Beispiel mit einer Nicht-Standard-csv-Datei	338
22.2	Lesen und Schreiben von JSON-Dateien	340
22.3	Lesen und Schreiben von Excel-Dateien.....	341
22.4	Aufgaben	342
23	Pandas: groupby	345
23.1	Groupby mit Series.....	346
23.2	Arbeitsweise von groupby	348

23.3	GroupBy mit DataFrames	350
23.3.1	GroupBy mit Funktion	352
23.3.2	Beispiel mit Datei.....	354
23.4	Aufgaben	355
24	Pivot-Tabellen	359
24.1	Pivot-Funktion in Pandas	360
24.2	Pivot-Aufruf ohne Werte für ‚values‘	362
24.3	Die Funktion pivot_table in Pandas	363
24.4	Pivoting auf den Titanic-Daten	365
24.5	Aufgaben	368
25	Umgang mit NaN	369
25.1	‚nan‘ in Python.....	369
25.2	NaN in Pandas	370
25.2.1	Beispiel mit NaNs.....	373
25.3	dropna() verwenden	375
25.4	Aufgaben	377
26	Binning	379
26.1	Einführung	379
26.2	Binning mit Pandas	380
26.2.1	Binning mit cut	380
26.2.2	Erzeugen eines IntervalIndex-Objektes	382
26.2.3	Mehr zu pd.cut	383
26.2.4	Speicheroptimierung mit Categorical.....	384
26.2.5	Binnings mit Labels	385
26.3	Aufgaben	386
27	Mehrstufige Indizierung	387
27.1	Einführung	387
27.2	Mehrstufig indizierte Series-Objekte	388
27.3	Mehrstufige Indizierung durch Listen-Multiplikation	390
27.4	Weitere Möglichkeiten zur Index-Erzeugung	391
27.5	Zugriffsmöglichkeiten	392

27.6	Dreistufige Indizes	396
27.7	Zusammenhang zu DataFrames	397
27.7.1	Manueller Weg mit <code>pd.concat</code>	398
27.7.2	<code>unstack</code> und <code>stack</code>	399
27.8	Vertauschen mehrstufiger Indizes.....	402
27.9	Aufgaben	404
28	Datenvisualisierung mit Pandas	407
28.1	Einführung	407
28.2	Liniendiagramm in Pandas	408
28.2.1	Series.....	408
28.2.2	DataFrames	410
28.2.3	Sekundärachsen (Twin Axes).....	413
28.2.4	Mehrere Y-Achsen	414
28.2.5	Spalten mit Zeichenketten (Strings) in Floats wandeln.....	416
28.3	Balkendiagramme in Pandas	417
28.3.1	Ein einfaches Beispiel.....	417
28.3.2	Balkengrafik für die Programmiersprachennutzung	418
28.3.3	Farbgebung einer Balkengrafik	420
28.4	Tortendiagramme in Pandas	421
28.4.1	Ein einfaches Beispiel.....	421
28.5	Flächenplot mit <code>area</code>	423
28.6	Aufgaben	424
29	Zeit und Datum	425
29.1	Einführung	425
29.2	Python-Standardmodule für Zeitdaten	426
29.2.1	Die <code>date</code> -Klasse	426
29.2.2	Die <code>time</code> -Klasse	428
29.3	Die <code>datetime</code> -Klasse	429
29.4	Unterschied zwischen Zeiten	431
29.4.1	Wandlung von <code>datetime</code> -Objekten in Strings	432
29.4.2	Wandlung mit <code>strptime</code>	433
29.5	Ausgabe in Landessprache.....	434
29.6	<code>datetime</code> -Objekte aus Strings erstellen.....	435

30	Zeitreihen.....	437
30.1	Einführung	437
30.2	Zeitreihen und Python	438
30.3	Datumsbereiche erstellen	440
30.4	Datumsbereiche mit Uhrzeitangaben	443
30.5	Aufgaben	444
 Teil IV Anwendungen		445
31	Techniken der Bildverarbeitung	447
31.1	Einführung	447
31.2	Bilder laden und anzeigen	448
31.3	Histogramme der Farbwerte	450
31.4	Bildausschnitte	452
31.5	Geometrische Transformationen	452
31.6	Filterung.....	454
31.7	Aufhellen und Abtönen von Bildern	459
31.8	Kachelung.....	467
31.9	Wasserzeichen mit np.where	468
31.10	Ein weiteres Beispiel für Wasserzeichen mit np.where	470
31.11	Aufgaben	473
32	Finanzverwaltung mit Pandas	475
32.1	Haushaltsbuch.....	475
32.1.1	Haushaltsbuch mit CSV-Datei.....	476
32.1.2	Excel-Haushaltsbuch mit Kontenplan	480
32.1.3	Auswertung des Excel-Haushaltsbuches	482
32.2	Einnahmeüberschussrechnung	483
32.2.1	Journaldatei	484
32.2.2	Analyse und Visualisierung der Daten.....	485
32.2.3	Steuersummen	490

Teil V	Lösungen zu den Aufgaben	493
33	Lösungen zu den Aufgaben	495
33.1	Lösungen zu Kapitel 5 (Erzeugung und Struktur von Arrays)	495
33.2	Lösungen zu Kapitel 6 (Datentyp-Objekt: dtype)	497
33.3	Lösungen zu Kapitel 7 (Arrays kombinieren und umformen)	499
33.4	Lösungen zu Kapitel 8 (Numerische Operationen auf NumPy-Arrays)	502
33.5	Lösungen zu Kapitel 9 (Statistik und Wahrscheinlichkeiten)	505
33.6	Lösungen zu Kapitel 10 (Boolesche Maskierung und Indizierung)	510
33.7	Lösungen zu Kapitel 13 (Objektorientiert plotten)	512
33.8	Lösungen zu Kapitel 14 (Mehrfache Plots und Doppelachsen)	515
33.9	Lösungen zu Kapitel 16 (Legenden und Annotationen)	517
33.10	Lösungen zu Kapitel 17 (Konturplots)	519
33.11	Lösungen zu Kapitel 18 (Histogramme und Diagramme)	523
33.12	Lösungen zu Kapitel 19 (Pandas:Series)	527
33.13	Lösungen zu Kapitel 20 (DataFrame)	531
33.14	Lösungen zu Kapitel 21 (Styling)	536
33.15	Lösungen zu Kapitel 22 (Dateiverarbeitung)	538
33.16	Lösungen zu Kapitel 23 (Pandas: groupby)	541
33.17	Lösungen zu Kapitel 24 (Pivot-Tabellen)	547
33.18	Lösungen zu Kapitel 25 (Umgang mit NaN)	548
33.19	Lösungen zu Kapitel 26 (Binning)	549
33.20	Lösungen zu Kapitel 27 (Mehrstufige Indizierung)	550
33.21	Lösungen zu Kapitel 28 (Datenvisualisierung mit Pandas)	555
33.22	Lösungen zu Kapitel 30 (Zeitreihen)	557
33.23	Lösungen zu Kapitel 31 (Techniken der Bildverarbeitung)	558
	Stichwortverzeichnis	561