

# Numerisches Python

Arbeiten mit NumPy, Matplotlib und Pandas

# DAS INHALTS- VERZEICHNIS

» Hier geht's  
direkt  
zum Buch

# Inhalt

<b>Vorwort</b> .....	<b>XV</b>
<b>Danksagung</b> .....	<b>XVI</b>
<b>1 Einleitung</b> .....	<b>1</b>
1.1 Die richtige Wahl .....	1
1.2 Aufbau des Buches .....	2
1.3 Python-Installation .....	2
1.4 Download der Beispiele .....	3
1.5 Anregungen und Kritik .....	3
<b>2 Numerisches Programmieren mit Python</b> .....	<b>4</b>
2.1 Begriffsbestimmung .....	4
2.2 Zusammenhang zwischen Python, NumPy, Matplotlib, SciPy und Pandas .....	5
2.3 Python, eine Alternative zu MATLAB .....	6
<b>Teil I NumPy</b> .....	<b>7</b>
<b>3 NumPy – Einführung</b> .....	<b>9</b>
3.1 Überblick .....	9
3.2 Vergleich NumPy-Datenstrukturen und Python .....	10
3.3 Ein einfaches Beispiel .....	10
3.4 Grafische Darstellung der Werte .....	11
3.5 Speicherbedarf .....	12
3.6 Zeitvergleich zwischen Python-Listen und NumPy-Arrays .....	15
<b>4 Arrays in NumPy erzeugen</b> .....	<b>17</b>
4.1 Erzeugung äquidistanter Intervalle .....	17
4.1.1 arange .....	17
4.1.2 linspace .....	19

4.1.3	Nulldimensionale Arrays in NumPy .....	20
4.1.4	Eindimensionales Array .....	21
4.1.5	Zwei- und mehrdimensionale Arrays .....	21
4.2	Gestalt eines Arrays .....	22
4.3	Indizierung und Teilbereichsoperator .....	23
4.4	Dreidimensionale Arrays .....	28
4.5	Arrays mit Nullen und Einsen .....	31
4.6	Arrays kopieren .....	32
4.6.1	numpy.copy(A) und A.copy() .....	32
4.6.2	Zusammenhängend gespeicherte Arrays .....	32
4.7	Einheitsmatrix .....	34
4.7.1	Die identity-Funktion .....	34
4.7.2	Die eye-Funktion .....	35
4.8	Datentypen .....	36
4.9	Aufgaben .....	38
<b>5</b>	<b>Datentyp-Objekt: dtype .....</b>	<b>40</b>
5.1	dtype .....	40
5.2	Strukturierte Arrays .....	42
5.3	Ein- und Ausgabe von strukturierten Arrays .....	45
5.4	Unicode-Strings in Arrays .....	46
5.5	Umbenennen von Spaltennamen .....	47
5.6	Spaltenwerte austauschen .....	47
5.7	Komplexeres Beispiel .....	48
5.8	Aufgaben .....	49
<b>6</b>	<b>Dimensionsänderungen, Konkatenationen, Stapeln .....</b>	<b>51</b>
6.1	Reduktion und Reshape von Arrays .....	51
6.1.1	flatten .....	51
6.1.2	ravel .....	52
6.1.3	Unterschiede zwischen ravel und flatten .....	53
6.1.4	reshape .....	54
6.2	Weitere Dimensionen hinzufügen .....	55
6.3	Konkatenation von Arrays .....	56
6.4	Vektoren stapeln .....	58
6.4.1	stack-Funktion .....	58
6.4.2	dstack-Funktion .....	60
6.4.3	vstack .....	63
6.4.4	hstack .....	64
6.5	„Fliesen“ mit tile .....	66

<b>7</b>	<b>Numerische Operationen auf NumPy-Arrays</b>	<b>68</b>
7.1	Operatoren und Skalare	68
7.2	Arithmetische Operationen auf zwei Arrays	70
7.3	Matrizenmultiplikation und Skalarprodukt	71
7.3.1	Definition der dot-Funktion	71
7.3.2	Beispiele zur dot-Funktion	71
7.3.3	Das dot-Produkt im dreidimensionalen Fall	72
7.4	Vergleichsoperatoren	78
7.5	Logische Operatoren	79
7.6	Broadcasting	79
7.6.1	Zeilenweises Broadcasting	80
7.6.2	Spaltenweises Broadcasting	83
7.6.3	Broadcasting von zwei eindimensionalen Arrays	85
7.7	Distanzmatrix	86
7.8	ufuncs	87
7.8.1	Anwendung von ufuncs	88
7.8.2	Parameter für Rückgabewerte bei ufuncs	89
7.8.3	accumulate	91
7.8.4	reduce	93
7.8.5	outer	94
7.8.6	at	94
7.9	Aufgaben	95
<b>8</b>	<b>Statistik und Wahrscheinlichkeiten</b>	<b>96</b>
8.1	Einführung	96
8.2	Auf dem random-Modul aufbauende Funktionen	97
8.2.1	Echte Zufallszahlen	97
8.2.2	Erzeugen einer Liste von Zufallszahlen	98
8.2.3	Zufällige Integer-Zahlen	99
8.2.4	Stichproben oder Auswahlen	100
8.2.5	Zufallsintervalle	100
8.2.6	Seed oder Startwert	101
8.2.7	Gewichtete Zufallsauswahl	102
8.2.8	Stichproben mit Python	105
8.2.9	Kartesische Auswahl	106
8.2.10	Kartesisches Produkt	107
8.2.11	Kartesische Auswahl: cartesian_choice	107

8.2.12	Gauss'sche Normalverteilung.....	110
8.2.13	Übung mit Binärsender .....	112
8.3	Das random-Untersmodul von NumPy .....	115
8.3.1	Integers und Floats zufällig erzeugen .....	115
8.3.2	<code>numpy.random.choice</code> .....	117
8.3.3	<code>numpy.random.random_sample</code> .....	118
8.4	Synthetische Verkaufszahlen .....	119
8.5	Aufgaben .....	121
<b>9</b>	<b>Boolesche Maskierung und Indizierung .....</b>	<b>123</b>
9.1	Fancy-Indizierung .....	125
9.2	Indizierung mit einem Integer-Array.....	125
9.3	<code>nonzero</code> und <code>where</code> .....	126
9.4	<code>flatnonzero</code> und <code>count_nonzero</code> .....	127
9.5	Aufgaben .....	127
<b>10</b>	<b>Lesen und Schreiben von Daten-Dateien .....</b>	<b>128</b>
10.1	Text-Dateien speichern mit <code>saveetxt</code> .....	128
10.2	Text-Dateien laden mit <code>loadtxt</code> .....	130
10.2.1	<code>loadtxt</code> ohne Parameter.....	130
10.2.2	Spezielle Trenner .....	130
10.2.3	Selektives Einlesen von Spalten .....	131
10.2.4	Datenkonvertierung beim Einlesen .....	131
10.3	<code>tofile</code> .....	133
10.4	<code>fromfile</code> .....	134
10.5	Best Practice, um Daten zu laden und zu speichern .....	135
10.6	Und noch ein anderer Weg: <code>genfromtxt</code> .....	136
10.7	<code>recfromcsv</code> .....	136
<b>Teil II</b>	<b>Matplotlib.....</b>	<b>137</b>
<b>11</b>	<b>Einführung .....</b>	<b>139</b>
11.1	Ein erstes Beispiel.....	140
11.2	Der Formatparameter von <code>pyplot.plot</code> .....	141
11.3	Bezeichnungen für die Achsen.....	143

<b>12</b>	<b>Objekt-Hierarchie</b> .....	<b>145</b>
12.1	Erzeugung einer Figure und Axes .....	147
12.2	Achsenbeschriftungen und Titel .....	148
12.3	Die Plot-Methode .....	149
12.4	Wertebereiche der Achsen .....	150
12.5	Plotten mehrerer Funktionen .....	152
12.6	Streudiagramme .....	153
12.7	Flächen einfärben .....	155
<b>13</b>	<b>Mehrfache Plots und Doppelachsen</b> .....	<b>158</b>
13.1	Mehrere Abbildungen und Achsen .....	158
<b>14</b>	<b>Gridspec in Matplotlib</b> .....	<b>166</b>
<b>15</b>	<b>Achsen- und Skalenteilung</b> .....	<b>174</b>
15.1	Achsenverschiebungen und Achsenbezeichnungen .....	174
15.2	Achsenbeschriftungen ändern .....	178
15.3	Justierung der Tick-Beschriftungen .....	179
<b>16</b>	<b>Legenden und Annotationen</b> .....	<b>180</b>
16.1	Legende hinzufügen .....	180
16.2	Annotations/Anmerkungen .....	183
16.3	Aufgaben .....	190
<b>17</b>	<b>Konturplots</b> .....	<b>191</b>
17.1	Erstellen eines Maschengitters .....	191
17.2	Funktionen auf Meshgrids .....	193
17.3	contour ohne meshgrid .....	196
17.4	Linienstil und Farben anpassen .....	196
17.5	Gefüllte Konturen .....	198
17.6	Individuelle Farben .....	199
17.7	Schwellen .....	200
17.8	Andere Grids .....	201
17.8.1	Meshgrid genauer .....	201
17.8.2	mgrid .....	203
17.8.3	ogrid .....	203
17.9	imshow .....	205
17.10	Aufgaben .....	206

<b>18</b>	<b>Histogramme und Diagramme</b>	<b>208</b>
18.1	Histogramme	209
18.2	Säulendiagramm	213
18.3	Balkendiagramme	214
18.4	Gruppierte Balkendiagramme	215
18.5	xkcd-Modus	219
18.6	Tortendiagramme	221
18.7	Stapeldiagramme	222
18.8	Aufgaben	223
<b>Teil III Pandas</b>		<b>225</b>
<b>19</b>	<b>Pandas: Einführung</b>	<b>227</b>
19.1	Datenstrukturen	228
19.2	Series	228
19.2.1	Indizierung	231
19.2.2	pandas.Series.apply	232
19.2.3	Zusammenhang zu Dictionaries	232
19.3	NaN – Fehlende Daten	233
19.3.1	Die Methoden <code>isnull()</code> und <code>notnull()</code>	234
19.3.2	Zusammenhang zwischen NaN und None	235
19.3.3	Fehlende Daten filtern	236
19.3.4	Fehlende Daten auffüllen	236
19.4	Aufgaben	237
<b>20</b>	<b>DataFrame</b>	<b>238</b>
20.1	Zusammenhang zu Series	238
20.2	Manipulation der Spaltennamen	239
20.3	DataFrames aus Dictionaries	240
20.4	Zugriff auf Spalten	241
20.5	Selektion von Zeilen	242
20.5.1	<code>loc</code>	242
20.5.2	<code>query</code>	243
20.6	Modifikation von DataFrames	245
20.6.1	Spalten einfügen	245
20.6.2	Spalten austauschen	249
20.6.3	Zeilen austauschen	250
20.6.4	Einzelne Werte mittels <code>at</code> und <code>iat</code> ändern	250

20.7	Index ändern .....	251
20.7.1	Umsortierung der Spalten und des Index .....	252
20.7.2	Spalten umbenennen .....	253
20.7.3	Spalte in Index umfunktionieren .....	253
20.8	Summen und kumulative Summen .....	254
20.9	Sortierung .....	257
20.10	DataFrame und verschachtelte Dictionaries .....	258
20.11	Aufgaben .....	259
<b>21</b>	<b>Dateiverarbeitung .....</b>	<b>261</b>
21.1	DSV-/CSV-Dateien .....	261
21.1.1	CSV- und DSV-Dateien lesen .....	262
21.1.2	Schreiben von CSV-Dateien .....	264
21.2	Lesen und Schreiben von JSON-Dateien .....	269
21.3	Lesen und Schreiben von Excel-Dateien .....	269
21.4	Aufgaben .....	270
<b>22</b>	<b>Pandas: groupby .....</b>	<b>272</b>
22.1	groupby mit Series .....	272
22.2	Arbeitsweise von groupby .....	274
22.3	groupby mit DataFrames .....	276
22.3.1	groupby mit Funktionen .....	277
22.3.2	Weitere Anwendungen zu groupby .....	279
22.4	Aufgaben .....	283
<b>23</b>	<b>Pivot-Tabellen .....</b>	<b>285</b>
23.1	Pivot-Funktion in Pandas .....	285
23.2	Pivot-Aufruf ohne Werte für values .....	288
23.3	Pivoting auf den Titanic-Daten .....	289
23.4	Aufgaben .....	291
<b>24</b>	<b>Umgang mit NaN .....</b>	<b>292</b>
24.1	nan in Python .....	292
24.2	NaN in Pandas .....	293
24.2.1	Beispiel mit NaNs .....	295
24.3	dropna() verwenden .....	297
24.4	Aufgaben .....	298



<b>25</b>	<b>Binning</b> .....	<b>299</b>
25.1	Einführung.....	299
25.2	Binning mit Pandas.....	302
25.2.1	Binning mit cut .....	302
25.2.2	Erzeugen eines IntervalIndex-Objektes.....	303
25.2.3	Mehr zu pd.cut.....	304
25.2.4	Categorical.....	305
25.2.5	Binnings mit Labels .....	305
<b>26</b>	<b>Mehrstufige Indizierung</b> .....	<b>306</b>
26.1	Einführung.....	306
26.2	Mehrstufig indizierte Series-Objekte.....	307
26.3	Alternative Möglichkeiten .....	307
26.4	Zugriffsmöglichkeiten.....	308
26.5	Dreistufige Indizes .....	311
26.6	Zusammenhang zu DataFrames .....	312
26.6.1	Der harte direkte Weg.....	312
26.6.2	unstack und stack .....	313
26.7	Vertauschen mehrstufiger Indizes.....	316
26.8	Aufgaben .....	317
<b>27</b>	<b>Datenvisualisierung mit Pandas</b> .....	<b>319</b>
27.1	Einführung.....	319
27.2	Liniendiagramm in Pandas .....	320
27.2.1	Series .....	320
27.2.2	DataFrames .....	322
27.2.3	Sekundärachsen (Twin Axes) .....	325
27.2.4	Mehrere Y-Achsen .....	326
27.2.5	Spalten mit Zeichenketten (Strings) in Floats wandeln .....	327
27.3	Balkendiagramme in Pandas .....	329
27.3.1	Ein einfaches Beispiel.....	329
27.3.2	Balkengrafik für die Programmiersprachennutzung .....	329
27.3.3	Farbgebung einer Balkengrafik .....	331
27.4	Kuchendiagramme in Pandas .....	331
27.4.1	Ein einfaches Beispiel.....	331
27.5	Flächenplot mit area .....	333
27.6	Aufgaben .....	334

<b>28</b>	<b>Zeit und Datum</b> .....	<b>335</b>
28.1	Einführung.....	335
28.2	Python-Standardmodule für Zeit-Daten .....	336
28.2.1	Die date-Klasse .....	336
28.2.2	Die time-Klasse .....	337
28.3	Die datetime-Klasse .....	338
28.4	Unterschied zwischen Zeiten .....	340
28.4.1	Wandlung von datetime-Objekten in Strings .....	341
28.4.2	Wandlung mit strftime.....	341
28.5	Ausgabe in Landessprache .....	342
28.6	datetime-Objekte aus Strings erstellen .....	343
<b>29</b>	<b>Zeitreihen</b> .....	<b>345</b>
29.1	Einführung.....	345
29.2	Zeitreihen und Python .....	345
29.3	Datumsbereiche erstellen .....	348
29.4	Datumsbereiche mit Uhrzeiten .....	350
29.5	Aufgaben .....	351
<b>Teil IV</b>	<b>Anwendungen</b> .....	<b>353</b>
<b>30</b>	<b>Techniken der Bildverarbeitung</b> .....	<b>355</b>
30.1	Einführung.....	355
30.2	Bilder im misc-Paket .....	356
30.3	Eigene Bilder.....	358
30.4	Histogramme der Farbwerte .....	359
30.5	Bilderausschnitte .....	360
30.6	Geometrische Transformationen .....	360
30.7	Filtern.....	362
30.8	Bilder aufhellen und abtönen .....	366
30.9	Kachelung.....	374
30.10	Wasserzeichen.....	376
30.11	Aufgaben .....	378
<b>31</b>	<b>Finanzverwaltung mit Pandas</b> .....	<b>379</b>
31.1	Haushaltsbuch .....	379
31.1.1	Haushaltsbuch mit CSV-Datei .....	380
31.1.2	Erzeugen eines Excel-Haushaltsbuches.....	382
31.1.3	Auswertung des Excel-Haushaltsbuches .....	384

31.2	Einnahmenüberschussrechnung .....	385
31.2.1	Journaldatei .....	386
31.2.2	Analyse und Visualisierung der Daten .....	387
31.2.3	Steuersummen .....	393

## **Teil V Lösungen zu den Aufgaben ..... 397**

### **32 Lösungen zu den Aufgaben ..... 399**

32.1	Lösungen zu <a href="#">Kapitel 4 (Arrays in NumPy erzeugen)</a> .....	399
32.2	Lösungen zu <a href="#">Kapitel 5 (Datentyp-Objekt: dtype)</a> .....	400
32.3	Lösungen zu <a href="#">Kapitel 7 (Numerische Operationen auf NumPy-Arrays)</a> .....	403
32.4	Lösungen zu <a href="#">Kapitel 8 (Statistik und Wahrscheinlichkeiten)</a> .....	404
32.5	Lösungen zu <a href="#">Kapitel 9 (Boolesche Maskierung und Indizierung)</a> .....	409
32.6	Lösungen zu <a href="#">Kapitel 13 (Mehrfache Plots und Doppelachsen)</a> .....	410
32.7	Lösungen zu <a href="#">Kapitel 16 (Legenden und Annotationen)</a> .....	412
32.8	Lösungen zu <a href="#">Kapitel 17 (Konturplots)</a> .....	414
32.9	Lösungen zu <a href="#">Kapitel 18 (Histogramme und Diagramme)</a> .....	418
32.10	Lösungen zu <a href="#">Kapitel 19 (Pandas: Einführung)</a> .....	421
32.11	Lösungen zu <a href="#">Kapitel 20 (DataFrame)</a> .....	422
32.12	Lösungen zu <a href="#">Kapitel 21 (Dateiverarbeitung)</a> .....	426
32.13	Lösungen zu <a href="#">Kapitel 23 (Pivot-Tabellen)</a> .....	429
32.14	Lösungen zu <a href="#">Kapitel 22 (Pandas: groupby)</a> .....	431
32.15	Lösungen zu <a href="#">Kapitel 24 (Umgang mit NaN)</a> .....	434
32.16	Lösungen zu <a href="#">Kapitel 26 (Mehrstufige Indizierung)</a> .....	435
32.17	Lösungen zu <a href="#">Kapitel 27 (Datenvisualisierung mit Pandas)</a> .....	439
32.18	Lösungen zu <a href="#">Kapitel 29 (Zeitreihen)</a> .....	441
32.19	Lösungen zu <a href="#">Kapitel 30 (Techniken der Bildverarbeitung)</a> .....	441

### **Stichwortverzeichnis ..... 443**