

Python for Scientists

Third Edition

The third edition of this practical introduction to Python has been thoroughly updated, with all code migrated to Python 3 and made available as Jupyter notebooks. The notebooks are available online with executable versions of all the book's content (and more).

The text starts with a detailed introduction to the basics of the Python language, without assuming any prior knowledge. Building upon each other, the most important Python packages for numerical math (NumPy), symbolic math (SymPy), and plotting (Matplotlib) are introduced, with brand new chapters covering numerical methods (SciPy) and data handling (Pandas). Further new material includes guidelines for writing efficient Python code and publishing code for other users.

Simple and concise code examples, revised for compatibility with Python 3, guide the reader and support the learning process throughout the book. Readers from the quantitative sciences, whatever their background, will be able to quickly acquire the skills needed for using Python effectively.

JOHN M. STEWART was Emeritus Reader in Gravitational Physics at the University of Cambridge, and a Life Fellow at King's College, Cambridge, before his death in 2016. He was the author of *Non-equilibrium Relativistic Kinetic Theory* (Springer, 1971) and *Advanced General Relativity* (Cambridge, 1991), and he translated and edited Hans Stephani's *General Relativity* (Cambridge, 1990).

MICHAEL MOMMERT is Assistant Professor for Computer Vision at the University of St. Gallen, Switzerland, where he combines computer vision and Earth observation to implement efficient learning methods for a wide range of use cases. Before, he was a solar system astronomer and actively wrote scientific open-source code for this community.

Cambridge University Press & Assessment
978-1-009-01480-9 — Python for Scientists
John M. Stewart , Michael Mommert
Frontmatter
[More Information](#)

Python for Scientists

Third Edition

JOHN M. STEWART

University of Cambridge

MICHAEL MOMMERT

University of St. Gallen, Switzerland



Cambridge University Press & Assessment
978-1-009-01480-9 — Python for Scientists
John M. Stewart, Michael Mommert
Frontmatter
[More Information](#)



CAMBRIDGE
UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom
One Liberty Plaza, 20th Floor, New York, NY 10006, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India
103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment,
a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of
education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781009014809

DOI: 10.1017/9781009029728

First and Second editions © John M. Stewart 2014, 2017

Third edition © Cambridge University Press & Assessment 2023

This publication is in copyright. Subject to statutory exception and to the provisions
of relevant collective licensing agreements, no reproduction of any part may take
place without the written permission of Cambridge University Press & Assessment.

First published 2014

Second edition 2017

Third edition 2023

Printed in the United Kingdom by CPI Group Ltd, Croydon CR0 4YY

A catalogue record for this publication is available from the British Library

A Cataloging-in-Publication data record for this book is available from the Library of Congress

ISBN 978-1-009-01480-9 Paperback

Additional resources for this publication at www.cambridge.org/9781009014809.

Cambridge University Press & Assessment has no responsibility for the persistence
or accuracy of URLs for external or third-party internet websites referred to in this
publication and does not guarantee that any content on such websites is, or will remain,
accurate or appropriate.

Contents

<i>Preface</i>	<i>page xi</i>
1 Introduction	1
1.1 Python for Scientists	1
1.2 Scientific Software	1
1.3 About This Book	4
1.4 References	4
2 About Python	5
2.1 What Is Python?	5
2.1.1 A Brief History of Python	6
2.1.2 The Zen of Python	7
2.2 Installing Python	8
2.2.1 Anaconda and Conda	9
2.2.2 Pip and PyPI	10
2.3 How Python Works	11
2.4 How to Use Python	11
2.4.1 The Python Interpreter	11
2.4.2 IPython and Jupyter	12
2.4.3 Integrated Development Environments	16
2.4.4 Cloud Environments	17
2.5 Where to Find Help?	17
2.6 References	19
3 Basic Python	22
3.1 Typing Python	22
3.2 Objects and Identifiers	23
3.3 Namespaces and Modules	26
3.4 Numbers	28
3.4.1 Integers	28
3.4.2 Real Numbers	28
3.4.3 Booleans	29
3.4.4 Complex Numbers	30

3.5	Container Objects	31
3.5.1	Lists	32
3.5.2	List Indexing	33
3.5.3	List Slicing	34
3.5.4	List Mutability	35
3.5.5	List Functions	36
3.5.6	Tuples	38
3.5.7	Strings	39
3.5.8	Dictionaries	41
3.5.9	Sets	42
3.6	Python <code>if</code> Statements	42
3.7	Loop Constructs	44
3.7.1	The <code>for</code> Loop	44
3.7.2	The <code>while</code> Loop	46
3.7.3	The <code>continue</code> Statement	46
3.7.4	The <code>break</code> Statement	47
3.7.5	List Comprehensions	48
3.8	Functions	49
3.8.1	Syntax and Scope	49
3.8.2	Positional Arguments	54
3.8.3	Keyword Arguments	54
3.8.4	Arbitrary Number of Positional Arguments	54
3.8.5	Arbitrary Number of Keyword Arguments	55
3.8.6	Anonymous Functions	55
3.9	Python Input/Output	56
3.9.1	Keyboard Input	56
3.9.2	The <code>print()</code> Function	56
3.9.3	File Input/Output	59
3.10	Error Handling	60
3.10.1	Traceback	60
3.10.2	Errors, Exceptions, and Warnings	62
3.11	Introduction to Python Classes	63
3.12	The Structure of Python	66
3.13	A Python Style Guide	67
3.14	References	68
4	NumPy: Numerical Math	72
4.1	Arrays	72
4.1.1	One-Dimensional Arrays	73
4.1.2	Basic Array Arithmetic	77
4.1.3	Two (and More)-Dimensional Arrays	79
4.1.4	Broadcasting	84
4.1.5	Array Manipulations	86

4.2	Working with Arrays	94
4.2.1	Mathematical Functions and Operators	94
4.2.2	Sums and Products	96
4.2.3	Comparing Arrays	96
4.2.4	Advanced Array Indexing	99
4.2.5	Sorting and Searching	100
4.3	Constants	101
4.4	Random Numbers	102
4.5	Simple Statistics	104
4.6	Polynomials	107
4.6.1	Converting Data to Coefficients	107
4.6.2	Converting Coefficients to Data	107
4.6.3	Manipulating Polynomials in Coefficient Form	108
4.7	Linear Algebra	108
4.7.1	Basic Operations on Matrices	108
4.7.2	Matrix Arithmetic	110
4.7.3	Solving Linear Systems of Equations	111
4.8	File Input/Output	112
4.8.1	Text File Input/Output	112
4.8.2	Binary File Input/Output	115
4.9	Special Array Types	116
4.9.1	Masked Arrays	116
4.9.2	Structured Arrays	117
4.10	References	118
5	SciPy: Numerical Methods	120
5.1	Special Functions	120
5.2	Constants	121
5.3	Numerical Integration	122
5.3.1	Integrating over Functions	122
5.3.2	Integrating over Sampled Values	124
5.4	Optimization and Root Search	124
5.4.1	Local Univariate Optimization	124
5.4.2	Local Multivariate Optimization	125
5.4.3	Function Fitting	127
5.4.4	Root Search	129
5.5	Numerical Interpolation	131
5.5.1	Univariate Interpolation	131
5.5.2	Multivariate Interpolation	133
5.6	Linear Algebra	134
5.6.1	Matrix Operations	135
5.7	Statistics	138
5.7.1	Univariate Continuous Probability Distributions	139
5.7.2	Multivariate and Discrete Probability Distributions	141

5.7.3	Correlation Tests	144
5.7.4	Distribution Tests	145
5.8	Ordinary Differential Equations	147
5.8.1	Initial Value Problems	148
5.8.2	Boundary Value Problems	151
5.9	SciKits: A Whole New World	153
5.10	References	153
6	Matplotlib: Plotting	156
6.1	Getting Started: Simple Figures	156
6.1.1	Frontends	156
6.1.2	Backends	157
6.1.3	A Simple Figure	157
6.2	Object-Oriented Matplotlib	159
6.3	Customizing Plots	161
6.3.1	Figure Size	162
6.3.2	Axis Range and Scaling	162
6.3.3	Ticks	163
6.3.4	Grid	164
6.3.5	Legend	164
6.3.6	Transparency	164
6.3.7	Text and Annotations	164
6.3.8	Mathematical Formulae	165
6.3.9	Colors	166
6.4	Cartesian Plots	169
6.4.1	Line Plots	169
6.4.2	Scatter Plots	170
6.4.3	Error Bars	171
6.4.4	Plotting Filled Areas	172
6.4.5	Bar Plots	173
6.5	Polar Plots	174
6.6	Plotting Images	175
6.7	Contour Plots	178
6.8	Compound Figures	179
6.9	Multidimensional Visualization	181
6.9.1	The Reduction to Two Dimensions	181
6.9.2	3D Plots	181
6.10	References	185
7	SymPy: Symbolic Math	187
7.1	Symbols and Functions	187
7.2	Conversions from Python to SymPy and Vice Versa	190

7.3	Matrices and Vectors	191
7.4	Some Elementary Calculus	193
	7.4.1 Differentiation	193
	7.4.2 Integration	193
	7.4.3 Series and Limits	196
7.5	Equality, Symbolic Equality and Simplification	197
7.6	Solving Equations	199
	7.6.1 Equations with One Independent Variable	200
	7.6.2 Linear Equations with More than One Independent Variable	201
	7.6.3 More General Equations	203
	7.6.4 Solving Ordinary Differential Equations	204
	7.6.5 Solving Partial Differential Equations	207
7.7	Plotting from within SymPy	207
7.8	References	212
8	Pandas: Data Handling	214
8.1	Series	214
8.2	DataFrames	216
	8.2.1 Axis Labels and Indexes	218
	8.2.2 Accessing Data	220
	8.2.3 Modifying Data	224
	8.2.4 Dealing with Missing Data	231
8.3	Specific Types of Data	234
	8.3.1 Categorical Data	234
	8.3.2 Textual Data	235
	8.3.3 Dates and Times	237
8.4	Functions	238
8.5	Data Visualization	241
8.6	File Input/Output	243
8.7	Pandas Hints	246
8.8	References	247
9	Performance Python	250
9.1	How to Write Efficient Python Code	250
	9.1.1 Measuring performance	250
	9.1.2 Optimization Starts before Coding	253
	9.1.3 Optimizing Basic Python	254
	9.1.4 NumPy	256
9.2	Parallelization	259
	9.2.1 Multithreading	260
	9.2.2 Multiprocessing	263
9.3	What Else?	264
9.4	References	265

10 Software Development Tools	267
10.1 Version Control	267
10.1.1 git	268
10.2 Create Your Own Python Module	274
10.3 Publish Your Code	276
10.3.1 GitHub	277
10.3.2 Python Package Index	277
10.4 References	278
<i>Index</i>	280

Preface

I bought a copy of *Python for Scientists* at a conference booth in 2016, looking for an affordable and easily readable textbook for a Python course I was teaching at that time. I was intrigued by how straightforwardly even complex things were explained in this book. It was a perfect match for my course and my students, despite my impression that the book was rather heavily focused on mathematical applications.

It has been five years since the second edition of *Python for Scientists* was released. This is a long time in the life cycle of a programming language that is still under active development. It was definitely time for an update.

Unfortunately, John is no longer with us to provide this update himself. Instead, I was honored that this task was offered to me, and I could not decline.

Besides updating the Python code examples shown in the book, I took the opportunity to also update the content of the book with the goal of making it accessible to a broader audience of scientists, especially those with a quantitative focus in their work. This includes a more in-depth discussion of numerical mathematics with NumPy (Chapter 4) and SciPy (Chapter 5), plotting capabilities with Matplotlib (Chapter 6), and, for the first time, data handling with Pandas (Chapter 8), performance computing with Python (Chapter 9), and an outline of software development techniques that are useful to scientists (Chapter 10). However, in order to keep the book reasonably short and affordable, other content, such as the detailed treatment of ordinary and partial differential equations, had to be significantly shortened or removed altogether – Python packages for dealing with such problems exist, but their discussion is beyond the scope of this beginner book.

I sincerely hope the third edition of *Python for Scientists* will be a useful companion on your long journey to becoming a scientific programmer.

Michael Mommert
St. Gallen, November 2022

Cambridge University Press & Assessment
978-1-009-01480-9 — Python for Scientists
John M. Stewart , Michael Mommert
Frontmatter
[More Information](#)
