Numeric PYTHON

Python Data Analysis with NumPy, Pandas, and Matplotlib

DAS VORWORT

Preface

Data-driven methods have become a central force in contemporary software development. Concepts such as *Big Data* and *Machine Learning* now play a decisive role across research, industry, and applied engineering. As a result, both institutions and individual practitioners face a fundamental question: which programming language is best suited for data-oriented and numerical applications?

In recent years, Python has consistently emerged as one of the most widely used programming languages in this domain. Its rapid development has established it as a core technology for data analysis, scientific computing, and artificial intelligence. While Python was not originally designed for numerical computation, its ecosystem has evolved to fill this gap with remarkable success.

This success is largely attributable to powerful extension libraries such as `NumPy`, `SciPy`, `Matplotlib`, and `Pandas`. Together, these tools have transformed Python into a robust platform that rivals specialized numerical and statistical software packages, both in academic research and in industrial practice.

This book provides a structured and accessible introduction to numerical and data-oriented programming with Python, with a particular focus on `NumPy`, `Matplotlib`, and `Pandas`. It is aimed at readers who already have a basic command of Python and are ready to deepen their skills in scientific and data-driven applications. The book is therefore well suited for students, researchers, engineers, and professionals seeking to extend their Python knowledge into the realm of numerical computing and data analysis.

The present volume is based on the successful German-language book *Numerisches Python* by the same author. The first edition was published on 18 June 2021, and a third, thoroughly revised edition appeared on 14 November 2025. This English-language edition is a translation of the third edition and includes several necessary adaptations for an international readership.

Brigitte Bauer-Schiewek, Editor in January 2026

Acknowledgments

Writing a book requires not only experience and expertise, but above all one essential thing: time. Time beyond the ordinary – time that is often made possible by the support and understanding of one's family. My special thanks therefore go to my wife, Karola, who accompanied me throughout the creation of this book – from the first to the third edition – with great understanding and active support.

A heartfelt thank you also goes to the many participants of my Python courses. Through their questions, feedback, and suggestions, I was able to continuously refine both my didactic approach and the technical concepts presented in this book. I am equally grateful to the many users of my online tutorials at *www.python-kurs.eu* and *www.python-course.eu*, especially those who contacted me with constructive comments and questions.

My sincere thanks also go to Hanser Verlag, which made the publication of this book possible. While this is the first English edition, it is in fact a translation of the third edition of a successful German-language book. I am especially grateful to Ms. Brigitte Bauer-Schiewek (Computer Book Program Planning) for the consistently excellent collaboration, as well as to Kristin Rothe (Editorial Assistant for Computer Books) for her support during the production process.

Finally, I would like to express my sincere gratitude to Python itself, to its creator Guido van Rossum, and to the countless contributors who continue to develop, maintain, and improve this remarkable programming language. The Python community, with its openness, commitment to quality, and emphasis on readability and elegance, has played a decisive role in Python's success and longevity.

I encountered Python for the first time roughly eighteen years ago, and it was an immediate and lasting fascination. Since then, Python has accompanied me throughout

my professional career and personal projects alike, shaping the way I think about programming, problem solving, and software design. Its clarity, versatility, and expressive power continue to inspire me, and this book would not exist without the language and the people behind it.

Bernd Klein, Singen in January 2026

1 Introduction

1.1 The Right Choice

Choosing the right programming language for everyday professional work is of great importance. This choice depends on many factors – and quite often, you don't really have a choice at all. The language is often predetermined by the company, the team, or the specific project. Increasingly, however, developers have the good fortune of being able to work with Python.

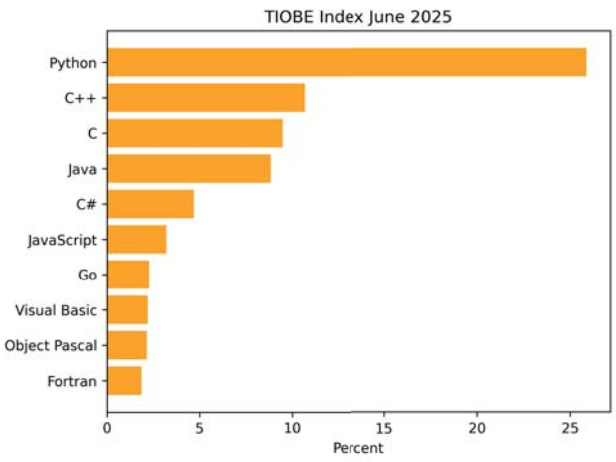

Figure 1.1 Top 10 Programming Languages

If you look at current surveys of the most popular programming languages, Python almost always ranks among the top – mostly at number one. In the TIOBE Index[1)], Python now clearly leads all other languages. Back in 2018, when the first edition of this book was published, Python was

1 The TIOBE Index, published monthly since 2001 by the Dutch company TIOBE Software BV, ranks programming languages by popularity. The ranking is based on the frequency of search queries containing the language's name in search engines like Google, Bing, or Yahoo!. The index does not measure lines of code or technical quality, but rather the language's visibility on the web.

still in third place in the TIOBE ranking. Even then, it was inducted into the TIOBE Hall of Fame – as the language with the highest popularity growth that year. Since October 2021, Python has held the top spot in the TIOBE Index without challenge.

Of course, it's nice to know that the language you prefer is also popular with many others – and perhaps even widely used in your industry. But one of the most important questions remains: Can your own projects be developed more easily and effectively in Python than in other languages? What we mean by "more easily and effectively" includes aspects like development time, runtime performance, maintainability, and so on.

Programming languages are like shoes: there's no one-size-fits-all. No single pair works equally well for formal events, the office, sports, and hiking. Python, however, is a language that can be used flexibly in most domains – one of the reasons for its great success. But above all, Python owes its rise to the powerful modules NumPy, SciPy, Matplotlib, and Pandas. These make it especially easy to solve numerical problems, thanks to Python's clear and readable syntax. Moreover, NumPy provides data structures that are 10 to 100 times faster than implementations in pure Python or many other languages. Since these modules are largely written in C, they achieve speeds close to native C programs.

1.2 Structure of the Book

This book is about Python and its excellent capabilities for tackling numerical problems – specifically, the modules that have become indispensable for topics like "Big Data" and "Machine Learning". The focus is on using the libraries `NumPy`, `Matplotlib`, and `Pandas`.

The book assumes basic knowledge of Python and is aimed at readers who have already gained some initial experience with the language, for example through an introductory course, online tutorial, or comparable textbook. It is intended for those who are familiar with core Python concepts such as variables, control structures, functions, and modules, and who now wish to extend their skills toward data-oriented and numerical programming.

This book offers a practice-oriented introduction to numerical programming with Python and is divided into several logically structured parts:

Part I: NumPy Begins with the fundamentals of numerical programming and demonstrates how to work efficiently with arrays using `NumPy`. Topics such as array creation, indexing, data types, mathematical operations, broadcasting, and statistical evaluations are systematically covered.

Part II: Matplotlib Introduces the basics of data visualization with `Matplotlib`. The range of presentation options is illustrated, from simple plots to complex diagrams with multiple axes or contour plots.

Part III: Pandas Provides an introduction to working with tabular data. The data structures `Series` and `DataFrame`, handling missing values, grouping, pivot tables, and time and date functions are covered. It also shows how to read and write data from various sources such as `CSV`, `Excel`, or `JSON` files.

Part IV: Applications Presents concrete use cases with `pandas` – including image processing, financial analysis, and analysis of energy production in Germany.

Part V: Solutions Contains the solutions to the exercises in the book for self-assessment.

The book is supplemented by an introductory chapter on the installation and setup of the required libraries.

1.3 This Book and the Tools Behind It

This book was created entirely using LaTeX and `pythontex`. This combines the typographic strengths of LaTeX – such as precise formatting, consistent layout, automatic tables of contents, cross-references, and bibliography management – with the dynamic capabilities of Python. Thanks to `pythontex`, Python code can be embedded directly into the document and executed during compilation. Results such as numerical output, tables, diagrams, or interactive content appear automatically at the correct location in the book.

With syntax highlighting, reproducible code execution, and direct integration into the text, a particularly transparent and consistent didactic workflow is achieved – ideal for a book on data analysis and visualization with Python.

1.4 Download the Examples

All examples used in the book are available for download at

http://www.python-kurs.eu/books/numerical_python

A list of corrections is also available there.

1.5 About the Author

He completed his studies in computer science in 1988 with a diploma from Saarland University. Until 2007, he worked as a software developer in the industry. Since then, he has been active internationally as a lecturer and trainer in software development and programming languages – focusing on the Python programming language since 2009.

He collaborates with universities, research institutions, and companies both in Germany and abroad. Since 2008, he has run the online learning platforms *https://python-kurs.eu* and *https://python-course.eu*, which are used by millions of learners worldwide every year.

Published books:

- Klein, Bernd: *Einführung in Python 3: Für Ein- und Umsteiger*. 4., vollständig überarbeitete Auflage, Carl Hanser Verlag GmbH & Co. KG, Munich, 2021.
  ISBN 978-3-446-46379-0
- Klein, Bernd: *Python–Grundlagen | eLearning*. 1. Auflage, Carl Hanser Verlag GmbH & Co. KG, Munich, 2023.
  ISBN 978-3-446-47992-0
- Klein, Bernd: *Numerisches Python: Arbeiten mit NumPy, Matplotlib und Pandas*. 3., aktualisierte Auflage, Carl Hanser Verlag GmbH & Co. KG, Munich, 2025.
  ISBN 978-3-446-48549-5
- Klein, Bernd; Klein, Philip: *Funktionale Programmierung mit Python*. 1. Auflage, Carl Hanser Verlag GmbH & Co. KG, Munich, 2025.
  ISBN 978-3-446-48191-6

1.6 Suggestions and Feedback

If you notice any inaccuracies or errors in the book, feel free to send an email directly to the author: `bernd.klein@python-kurs.eu`.

This also applies to suggestions, comments, or requests regarding the book.

We will take all feedback into account in future editions.

I wish all readers as much enjoyment reading this book as I had writing it. Have fun exploring, programming, and discovering!

Bernd Klein, January 2026