1. Record Nr. UNINA9910483273203321

Autore	Unpingco Jose <1969->
Titolo	Python programming for data analysis / / Jose Unpingco
Pubbl/distr/stampa	Cham, Switzerland : , : Springer, , [2021] ©2021
ISBN	3-030-68952-2
Edizione	[1st ed. 2021.]
Descrizione fisica	1 online resource (XII, 263 p. 134 illus., 123 illus. in color.)
Disciplina	005.133
Soggetti	Python (Computer program language)
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	Introduction -- Basic Language -- Basic Data Structures -- Basic Programming -- File Input/Output -- Dealing with Errors -- Power Python Features to Master -- Advanced Language Features -- Using modules -- Object oriented programming -- Debugging from Python -- Using Numpy – Numerical Arrays in Python -- Data Visualization Using Python -- Bokeh for Web-based Visualization -- Getting Started with Pandas -- Some Useful Python-Fu -- Conclusion.
Sommario/riassunto	This textbook grew out of notes for the ECE143 Programming for Data Analysis class that the author has been teaching at University of California, San Diego, which is a requirement for both graduate and undergraduate degrees in Machine Learning and Data Science. This book is ideal for readers with some Python programming experience. The book covers key language concepts that must be understood to program effectively, especially for data analysis applications. Certain low-level language features are discussed in detail, especially Python memory management and data structures. Using Python effectively means taking advantage of its vast ecosystem. The book discusses Python package management and how to use third-party modules as well as how to structure your own Python modules. The section on object-oriented programming explains features of the language that facilitate common programming patterns. After developing the key Python language features, the book moves on to third-party modules that are foundational for effective data analysis, starting with Numpy.

The book develops key Numpy concepts and discusses internal Numpy array data structures and memory usage. Then, the author moves onto Pandas and details its many features for data processing and alignment. Because strong visualizations are important for communicating data analysis, key modules such as Matplotlib are developed in detail, along with web-based options such as Bokeh, Holoviews, Altair, and Plotly. The text is sprinkled with many tricks-of-the-trade that help avoid common pitfalls. The author explains the internal logic embodied in the Python language so that readers can get into the Python mindset and make better design choices in their codes, which is especially helpful for newcomers to both Python and data analysis. To get the most out of this book, open a Python interpreter and type along with the many code samples.