

Thinking in Pandas

How to Use the Python Data Analysis Library the Right Way

Understand and implement big data analysis solutions in pandas with an emphasis on performance. This book strengthens your intuition for working with pandas, the Python data analysis library, by exploring its underlying implementation and data structures.

Thinking in Pandas introduces the topic of big data and demonstrates concepts by looking at exciting and impactful projects that pandas helped to solve. From there, you will learn to assess your own projects by size and type to see if pandas is the appropriate library for your needs. Author Hannah Stepanek explains how to load and normalize data in pandas efficiently, and reviews some of the most commonly used loaders and several of their most powerful options. You will then learn how to access and transform data efficiently, what methods to avoid, and when to employ more advanced performance techniques. You will also go over basic data access and munging in pandas and the intuitive dictionary syntax. Choosing the right DataFrame format, working with multi-level DataFrames, and how pandas might be improved upon in the future are also covered.

By the end of the book, you will have a solid understanding of how the pandas library works under the hood. Get ready to make confident decisions in your own projects by utilizing pandas—the right way.

What You Will Learn

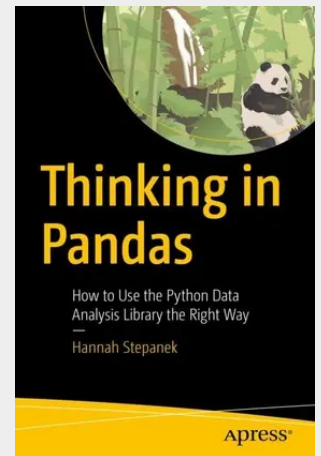
- Understand the underlying data structure of pandas and why it performs the way it does under certain circumstances
- Discover how to use pandas to extract, transform, and load data correctly with an emphasis on performance
- Choose the right DataFrame so that the data analysis is simple and efficient.
- Improve performance of pandas operations with other Python libraries

Who This Book Is For Software engineers with basic programming skills in Python keen on using pandas for a big data analysis project. Python software developers interested in big data.

Understand and implement big data analysis solutions in pandas with an emphasis on performance. This book strengthens your intuition for working with pandas, the Python data analysis library, by exploring its underlying implementation and data structures. *Thinking in Pandas* introduces the topic of big data and demonstrates concepts by looking at exciting and impactful projects that pandas helped to solve. From there, you will learn to assess your own projects by size and type to see if pandas is the appropriate library for your needs. Author Hannah Stepanek explains how to load and normalize data in pandas efficiently, and reviews some of the most commonly used loaders and several of their most powerful options. You will then learn how to access and transform data efficiently, what methods to avoid, and when to employ more advanced performance techniques. You will also go over basic data access and munging in pandas and the intuitive dictionary syntax. Choosing the right DataFrame format, working with multi-level DataFrames, and how pandas might be improved upon in the future are also covered. By the end of the book, you will have a solid understanding of how the pandas library works under the hood. Get ready to make confident decisions in your own projects by utilizing pandas—the right way.

What You Will Learn Understand the underlying data structure of pandas and why it performs the way it does under certain circumstances Discover how to use pandas to extract, transform, and load data correctly with an emphasis on performance Choose the right DataFrame so that the data analysis is simple and efficient. Improve performance of pandas operations with other Python libraries

Who This Book Is For Software engineers with basic programming skills in Python keen on using pandas for a big data analysis project. Python software developers interested in big data.



48,14 €

44,99 € (zzgl. MwSt.)

Lieferfrist: bis zu 10 Tage

Artikelnummer: 9781484258385

Medium: Buch

ISBN: 978-1-4842-5838-5

Verlag: Apress

Erscheinungstermin: 06.06.2020

Sprache(n): Englisch

Auflage: 1. Auflage 2020

Produktform: Kartoniert

Gewicht: 312 g

Seiten: 186

Format (B x H): 155 x 235 mm

