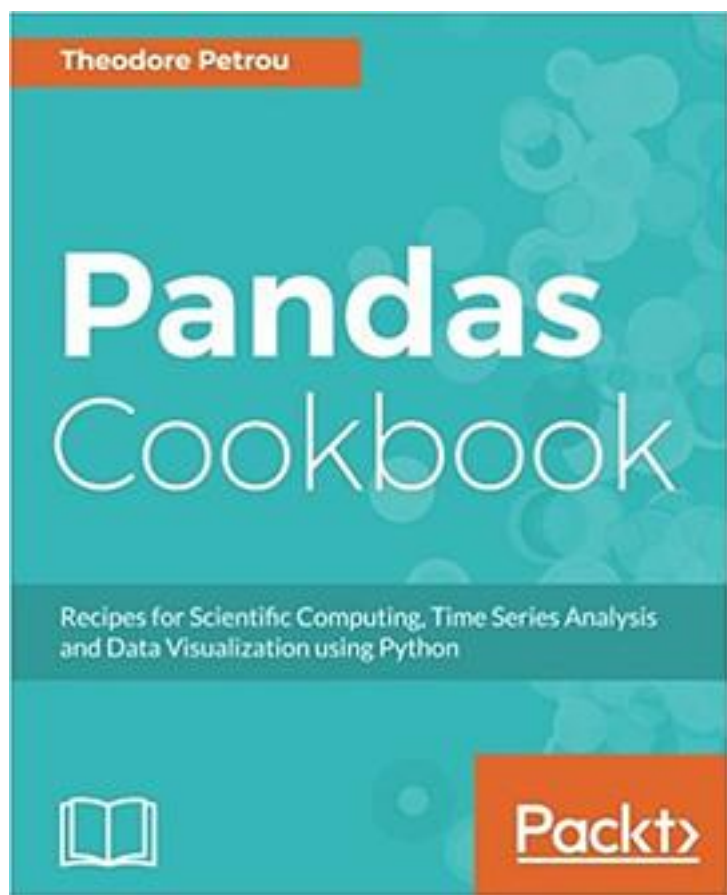# Pandas Cookbook

Pandas Cookbook_下载链接1_

著者:Theodore Petrou

出版者:Packt Publishing

出版时间:2017-10-23

装帧:Paperback

isbn:9781784393878

Key Features

Use the power of pandas to solve most complex scientific computing problems with ease

Leverage fast, robust data structures in pandas to gain useful insights from your data

Practical, easy to implement recipes for quick solutions to common problems in data using pandas

Book Description

This book will provide you with unique, idiomatic, and fun recipes for both fundamental and advanced data manipulation tasks with pandas. Some recipes focus on achieving a deeper understanding of basic principles, or comparing and contrasting two similar operations. Other recipes will dive deep into a particular dataset, uncovering new and unexpected insights along the way.

The pandas library is massive, and it's common for frequent users to be unaware of many of its more impressive features. The official pandas documentation, while thorough, does not contain many useful examples of how to piece together multiple commands like one would do during an actual analysis. This book guides you, as if you were looking over the shoulder of an expert, through practical situations that you are highly likely to encounter.

Many advanced recipes combine several different features across the pandas library to generate results.

What you will learn

Master the fundamentals of pandas to quickly begin exploring any dataset

Isolate any subset of data by properly selecting and querying the data

Split data into independent groups before applying aggregations and transformations to each group

Restructure data into tidy form to make data analysis and visualization easier

Prepare real-world messy datasets for machine learning

Combine and merge data from different sources through pandas SQL-like operations

Utilize pandas unparalleled time series functionality

Create beautiful and insightful visualizations through pandas direct hooks to Matplotlib and Seaborn

作者介绍:

About the Author

Theodore Petrou is a data scientist and the founder of Dunder Data, a professional educational company focusing on exploratory data analysis. He is also the head of Houston Data Science, a meetup group with more than 2,000 members that has the primary goal of getting local data enthusiasts together in the same room to practice data science. Before founding Dunder Data, Ted was a data scientist at Schlumberger, a large oil services company, where he spent the vast majority of his time exploring

data.

Some of his projects included using targeted sentiment analysis to discover the root cause of part failure from engineer text, developing customized client/server dashboarding applications, and real-time web services to avoid the mispricing of sales items. Ted received his masters degree in statistics from Rice University, and used his analytical skills to play poker professionally and teach math before becoming a data scientist. Ted is a strong supporter of learning through practice and can often be found answering questions about pandas on Stack Overflow.

目录: Table of Contents
1. Pandas Foundations
2. Essential DataFrame Operations
3. Beginning Data Analysis
4. Selecting Subsets of Data
5. Boolean Indexing
6. Index Alignment
7. Grouping for Aggregation, Filtration and Transformation
8. Restructuring Data into Tidy Form
9. Joining multiple pandas objects
10. Time Series
11. Visualization
· · · · · · ([收起](#))

[Pandas Cookbook_下载链接1](#)

标签

Python

Pandas

Programming

数据分析

计算机

编程

评论

1.内容翔实，但按照顺序读起来有些枯燥，最好的学习方式应该是把不同的分析操作融合在数据分析的案例中，未来还需要在案例分析中印证和反复练习。2.pandas里面的数据结构介绍深度不够，需要去看官方文档。

实在

Nice

以dataframe和series两大数据结构为核心，把各自相关的attributes和methods结合案例讲述得非常清晰。 第8章非常有用，讲怎么把raw data通过stack, melt, pivot等方式整理成Hadley Wickham标准下的tidy data。

很实在了，要深入的话可以进一步看官方文档

工具书

最美不过数据框———引自 艾新波

Pandas Cookbook_下载链接1_

书评

[Pandas Cookbook_下载链接1_](#)