

*Agile Tools for Real-World Data*

# Python for Data Analysis



Copyright © 2012 O'Reilly Media. All rights reserved.

**O'REILLY®**

McKinney, Wes. Python for Data Analysis : Data Wrangling with Pandas, NumPy, and IPython, O'Reilly Media, 2012. ProQuest Ebook Central,  
<http://ebookcentral.proquest.com/lib/kcl/detail.action?docID=1041908>.

Created from kcl on 2019-04-30 04:51:41.

*Wes McKinney*

# Python for Data Analysis

Looking for complete instructions on manipulating, processing, cleaning, and crunching structured data in Python? This hands-on book is packed with practical case studies that show you how to effectively solve a broad set of data analysis problems, using several Python libraries—including NumPy, pandas, matplotlib, and IPython.

Written by Wes McKinney, the main author of the pandas library, *Python for Data Analysis* also serves as a practical, modern introduction to scientific computing in Python for data-intensive applications. It's ideal for analysts new to Python and for Python programmers new to scientific computing.

- Use the IPython interactive shell as your primary development environment
- Learn basic and advanced NumPy (Numerical Python) features
- Get started with data analysis tools in the pandas library
- Use high-performance tools to load, clean, transform, merge, and reshape data
- Create scatter plots and static or interactive visualizations with matplotlib
- Apply the pandas groupby facility to slice, dice, and summarize datasets
- Work with time series data in many different forms
- Learn how to solve problems in web analytics, social sciences, finance, and economics, through detailed examples

Wes McKinney is the main author of pandas, the popular open source Python library for data analysis. Wes is an active speaker and participant in the Python and open source communities. He worked as a quantitative analyst at AQR Capital Management and Python consultant before founding DataPad, a data analytics company, in 2013. He graduated from MIT with an S.B. in Mathematics.

*“The scientific and data analysis communities have been waiting for this book for years: loaded with concrete practical advice, yet full of insight about how all the pieces fit together. It should become a canonical reference for technical computing in Python for years to come.”*

—Fernando Pérez

research scientist at UC Berkeley,  
one of the originators of IPython

US \$39.99

CAN \$41.99

ISBN: 978-1-449-31979-3



5 3 9 9 9



Twitter: @oreillymedia  
facebook.com/oreilly

O'REILLY®  
oreilly.com

---

# Python for Data Analysis

*Wes McKinney*

O'REILLY®

Beijing • Cambridge • Farnham • Köln • Sebastopol • Tokyo

## Python for Data Analysis

by Wes McKinney

Copyright © 2013 Wes McKinney. All rights reserved.

Printed in the United States of America.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (<http://my.safaribooksonline.com>). For more information, contact our corporate/institutional sales department: 800-998-9938 or [corporate@oreilly.com](mailto:corporate@oreilly.com).

**Editors:** Julie Steele and Meghan Blanchette

**Production Editor:** Melanie Yarbrough

**Copieditor:** Teresa Exley

**Proofreader:** BIM Publishing Services

**Indexer:** BIM Publishing Services

**Cover Designer:** Karen Montgomery

**Interior Designer:** David Futato

**Illustrator:** Rebecca Demarest

October 2012: First Edition.

### Revision History for the First Edition:

2012-10-05 First release

2013-05-17 Second release

See <http://oreilly.com/catalog/errata.csp?isbn=9781449319793> for release details.

Nutshell Handbook, the Nutshell Handbook logo, and the O'Reilly logo are registered trademarks of O'Reilly Media, Inc. *Python for Data Analysis*, the cover image of a golden-tailed tree shrew, and related trade dress are trademarks of O'Reilly Media, Inc.

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and O'Reilly Media, Inc., was aware of a trademark claim, the designations have been printed in caps or initial caps.

While every precaution has been taken in the preparation of this book, the publisher and author assume no responsibility for errors or omissions, or for damages resulting from the use of the information contained herein.

ISBN: 978-1-449-31979-3

[LSI]

1368653545