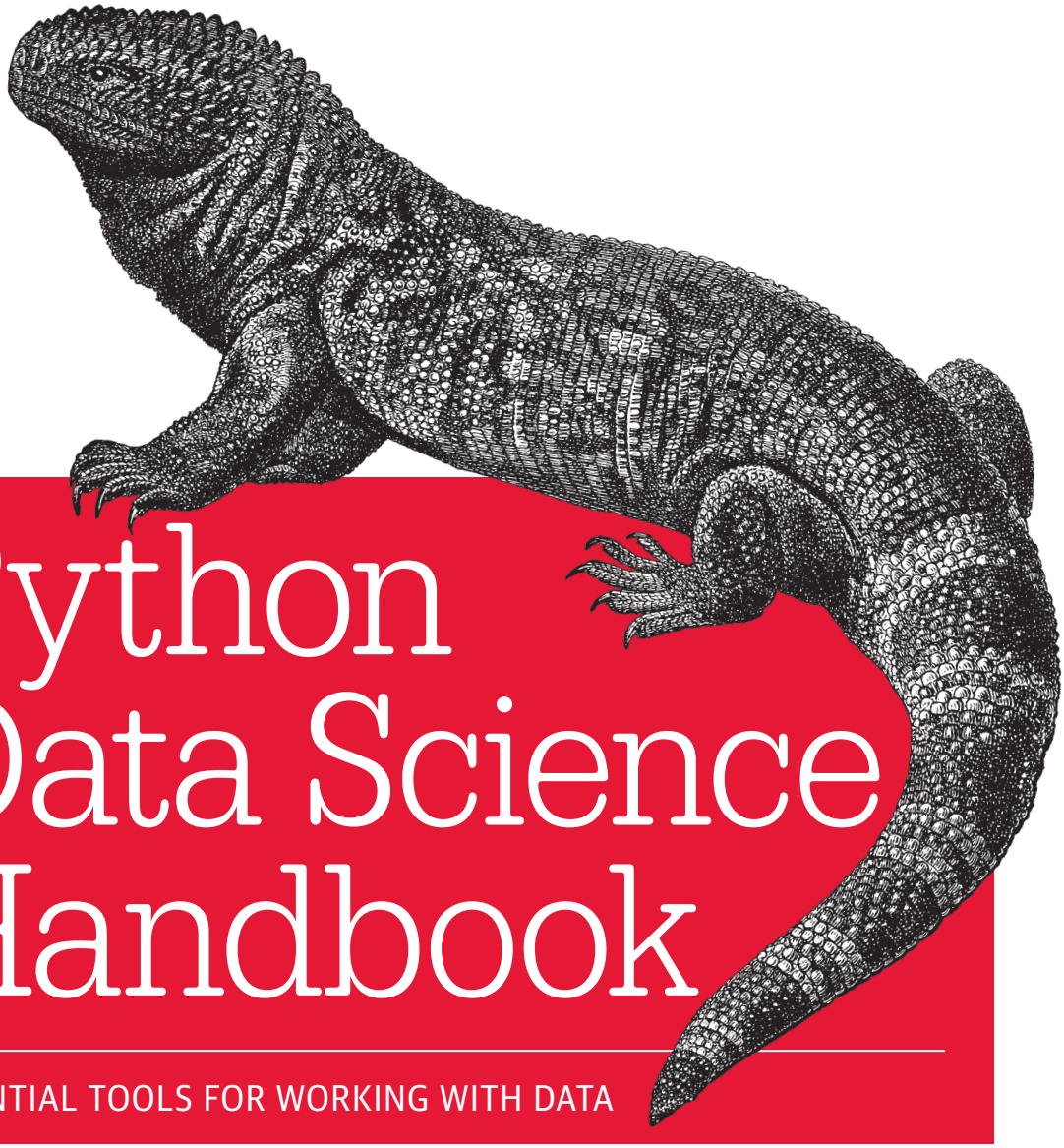
O'REILLY®



Python Data Science Handbook

ESSENTIAL TOOLS FOR WORKING WITH DATA

powered by

jupyter

Jake VanderPlas

O'REILLY®

Python Data Science Handbook

For many researchers, Python is a first-class tool mainly because of its libraries for storing, manipulating, and gaining insight from data. Several resources exist for individual pieces of this data science stack, but only with the *Python Data Science Handbook* do you get them all—IPython, NumPy, Pandas, Matplotlib, Scikit-Learn, and other related tools.

Working scientists and data crunchers familiar with reading and writing Python code will find this comprehensive desk reference ideal for tackling day-to-day issues: manipulating, transforming, and cleaning data; visualizing different types of data; and using data to build statistical or machine learning models. Quite simply, this is the must-have reference for scientific computing in Python.

With this handbook, you'll learn how to use:

- **IPython and Jupyter:** provide computational environments for data scientists using Python
- **NumPy:** includes the *ndarray* for efficient storage and manipulation of dense data arrays in Python
- **Pandas:** features the *DataFrame* for efficient storage and manipulation of labeled/columnar data in Python
- **Matplotlib:** includes capabilities for a flexible range of data visualizations in Python
- **Scikit-Learn:** for efficient and clean Python implementations of the most important and established machine learning algorithms

"If you want to learn data science with Python, this book is a fantastic starting point. I've used it with great success to teach computer science and statistics majors. Jake goes far beyond the basics of open source tools; he also explains the underlying concepts, patterns, and abstractions of data science using clear language and approachable explanations."

—Brian Granger
Associate Professor of Physics, Cal Poly; cofounder of Project Jupyter

Jake VanderPlas, a long-time user and developer of the Python scientific stack, currently works as an interdisciplinary research director at the University of Washington. He conducts his own astronomy research, and spends time advising and consulting with local scientists from a wide range of fields.

PYTHON / DATA

US $59.99 CAN $68.99

ISBN: 978-1-491-91205-8

Twitter: @oreillymedia
facebook.com/oreilly

Python Data Science Handbook

Essential Tools for Working with Data

Jake VanderPlas

Beijing • Boston • Farnham • Sebastopol • Tokyo

O'REILLY®

Python Data Science Handbook
by Jake VanderPlas

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://oreilly.com/safari*). For more information, contact our corporate/institutional sales department: 800-998-9938 or *corporate@oreilly.com*.

Editor: Dawn Schanafelt
Production Editor: Kristen Brown
Copyeditor: Jasmine Kwityn
Proofreader: Rachel Monaghan
Indexer: WordCo Indexing Services, Inc.
Interior Designer: David Futato
Cover Designer: Karen Montgomery
Illustrator: Rebecca Demarest

December 2016: First Edition

Revision History for the First Edition
2016-11-17: First Release

See *http://oreilly.com/catalog/errata.csp?isbn=9781491912058* for release details.

978-1-491-91205-8

[LSI]

Table of Contents