



# Python Data Science Handbook: Essential Tools for Working with Data

*Jake Vanderplas*

[Download now](#)

[Read Online ➔](#)

# Python Data Science Handbook: Essential Tools for Working with Data

*Jake Vanderplas*

## Python Data Science Handbook: Essential Tools for Working with Data Jake Vanderplas

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

## Python Data Science Handbook: Essential Tools for Working with Data Details

Date : Published November 21st 2016 by O'Reilly Media (first published March 25th 2016)

ISBN :

Author : Jake Vanderplas

Format : Paperback 541 pages

Genre : Computer Science, Programming, Science, Technical, Coding



[Download Python Data Science Handbook: Essential Tools for Worki ...pdf](#)



[Read Online Python Data Science Handbook: Essential Tools for Wor ...pdf](#)

**Download and Read Free Online Python Data Science Handbook: Essential Tools for Working with Data Jake Vanderplas**

## **From Reader Review Python Data Science Handbook: Essential Tools for Working with Data for online ebook**

### **Ye Lin Kyaw says**

It is broad and deep enough for the beginners and experienced users who migrate from other platforms

---

### **James Mason says**

Extremely well written. Just the right level of depth. It was useful to work through bit by bit to gain a general understanding and practice, and I'm sure it will also be useful as a desktop reference. I was inspired throughout to look at my data in new ways and apply new, modern methods to the data in order to obtain more robust results and hopefully uncover things about it that I simply would not have otherwise. Most of that happened in the machine learning (final) chapter. I appreciated the attention to aesthetics in visualizations in earlier chapters, especially the one on matplotlib. And I also really appreciated the first chapter on IPython and the various ways you can write your code, though I wish it had a little more breadth in terms of the available options and justifications for why you might use, e.g., Jupyter notebooks as opposed to Atom/Ipython console. I also wish that there were more astronomy examples since that is the author's and my area of study. Despite those minor qualms, 5 stars!

Note that the goodreads subtitle is incorrect. It should be: Essential Tools for Working with Data.

---

### **Loc Nguyen says**

Useful book, especially like pandas package for data manipulation.

---

### **Alexeyqu says**

[https://github.com/alexeyqu/random\\_co...](https://github.com/alexeyqu/random_co...)

---

### **Jeremy says**

The book is written as a Jupyter notebook, and is available for free on Github:

<https://github.com/jakevdp/PythonData...>

Books written as Jupyter Notebooks are simply wonderful. They should become the default medium for learning new materials related to computer science and mathematics.

Regarding the book itself, it fits more in the "practical knowledge" category, which is totally fine since it's a handbook. Being exposed to the different methods and tools is great. There is however no real theoretical

---

explanations behind the tools themselves or details about their implementation, but the reader can freely refer to extra materials if needed.

---

### **Mikkel Hartmann says**

I read this book after having worked as a data scientist for about a year and a half. Most of my work had focused on machine learning, so I had picked up Numpy, Pandas, and Matplotlib along the way. This approach left some glaring holes in my usage of these modules. After having read this book I can see that there has been a couple of things I have been doing wrong -- or at least very ineffectively. So reading this book was definitely a good idea.

I especially appreciated the chapters on Numpy and Pandas (~180 pages). Particularly the proper usage of indexing (eg. timestamps as indices) and multi-indexing for hierachal structure. Both chapters also contain advice on how to speed up the code when needed.

Generally, I really liked this book and will definitely add it to our library at work so I can reference it and lend it to our students and interns.

---

### **Matt Heavner says**

The python data science handbook is the best python tutorial I have read. It is "an overview of python if you want to be a data scientist" - the breadth and depth on specific tools (matplotlib & beyond, pandas, and sci-kit, as well as ipython & jupyter notebooks) is perfect for a data science application. This is definitely addressing the "computer skills" third of the data science Venn diagram (not much on mathematics or subject matter expertise). Recommended for learning python or having as a reference.

---

### **Hays Hutton says**

Liked how it goes in depth into NumPy and then Pandas. Sometimes a "little" too API based but that makes it practical in some respects.

---

### **Shashank Shrivastava says**

To the point and with examples in notebook format, it was easy to follow and understand.

---

### **Kenta Suzuki says**

A good introduction to how to manipulate and handle the data, useful for the people who code in python and want a foundation of data science.

This book explains how to use especially numpy, pandas, matplotlib, data wrangling part.

---

This book also contains light (not so rigorous) machine learning techniques.

---

### **Terran M says**

This book is not as good as R for Data Science: Import, Tidy, Transform, Visualize, and Model Data, but if you are constrained or committed to using Python, it is the best available alternative as of 2018. Chapters 1 through 3 on ipython, Numpy, and Pandas are excellent. Chapter 4 on Matplotlib is disappointing, but that's because Matplotlib is itself a weak and obsolete tool; the book acknowledges that fact and cannot fix it. I do not care for Chapter 5, which attempts too much and delivers too little (for example, the "in depth" treatment of linear regression is all of 2 pages). I suggest that you stop at the end of Chapter 4 and instead move on to Introduction to Machine Learning with Python: A Guide for Data Scientists.

---

### **Ryan says**

Outstanding.

---

### **Ray says**

This is really an amazing technical resource. Vanderplas manages to keep his content extraordinarily practical and grounded, without being irreverent to the theory like so many lower-quality modern data science texts are. As a contributor to the Python data software libraries such as Scikit-learn, the author is eminently qualified to give a tour of their inner workings. Finally, the book is self-aware of where it lacks depth, and does an excellent job in referring readers to further resources.

---

### **Faheemsadiki says**

its amazing book

---