

# 《利用Python进行数据分析》

## 图书基本信息

书名：《利用Python进行数据分析》

13位ISBN编号：9787111436737

10位ISBN编号：7111436733

出版时间：2013-11-18

出版社：机械工业出版社

作者：Wes McKinney

页数：464

译者：唐学韬

版权说明：本站所提供下载的PDF图书仅提供预览和简介以及在线试读，请支持正版图书。

更多资源请访问：[www.tushu000.com](http://www.tushu000.com)

# 《利用Python进行数据分析》

## 内容概要

# 《利用Python进行数据分析》

## 作者简介

Wes McKinney 资深数据分析专家，对各种Python库（包括NumPy、pandas、matplotlib以及IPython等）等都有深入研究，并在大量的实践中积累了丰富的经验。撰写了大量与Python数据分析相关的经典文章，被各大技术社区争相转载，是Python和开源技术社区公认的权威人物之一。开发了用于数据分析的著名开源Python库——pandas，广获用户好评。在创建Lambda Foundry（一家致力于企业数据分析的公司）之前，他曾是AQR Capital Management的定量分析师。

# 《利用Python进行数据分析》

## 书籍目录

### 目录

前言	1
第1章 准备工作	5
本书主要内容	5
为什么要使用Python进行数据分析	6
重要的Python库	7
安装和设置	10
社区和研讨会	16
使用本书	16
致谢	18
第2章 引言	20
来自bit.ly的1.usa.gov数据	21
MovieLens 1M数据集	29
1880—2010年间全美婴儿姓名	35
小结及展望	47
第3章 IPython：一种交互式计算和开发环境	48
IPython基础	49
内省	51
使用命令历史	60
与操作系统交互	63
软件开发工具	66
IPython HTML Notebook	75
利用IPython提高代码开发效率的几点提示	77
高级IPython功能	79
致谢	81
第4章 NumPy基础：数组和矢量计算	82
NumPy的ndarray：一种多维数组对象	83
通用函数：快速的元素级数组函数	98
利用数组进行数据处理	100
用于数组的文件输入输出	107
线性代数	109
随机数生成	111
范例：随机漫步	112
第5章 pandas入门	115
pandas的数据结构介绍	116
基本功能	126
汇总和计算描述统计	142
处理缺失数据	148
层次化索引	153
其他有关pandas的话题	158
第6章 数据加载、存储与文件格式	162
读写文本格式的数据	162
二进制数据格式	179
使用HTML和Web API	181
使用数据库	182
第7章 数据规整化：清理、转换、合并、重塑	186
合并数据集	186

- 重塑和轴向旋转 200
- 数据转换 204
- 字符串操作 217
- 示例：USDA食品数据库 224
- 第8章 绘图和可视化 231
  - matplotlib API入门 231
  - pandas中的绘图函数 244
  - 绘制地图：图形化显示海地地震危机数据 254
- Python图形化工具生态系统 260
- 第9章 数据聚合与分组运算 263
  - GroupBy技术 264
  - 数据聚合 271
  - 分组级运算和转换 276
  - 透视表和交叉表 288
  - 示例：2012联邦选举委员会数据库 291
- 第10章 时间序列 302
  - 日期和时间数据类型及工具 303
  - 时间序列基础 307
  - 日期的范围、频率以及移动 311
  - 时区处理 317
  - 时期及其算术运算 322
  - 重采样及频率转换 327
  - 时间序列绘图 334
  - 移动窗口函数 337
  - 性能和内存使用方面的注意事项 342
- 第11章 金融和经济数据应用 344
  - 数据规整化方面的话题 344
  - 分组变换和分析 355
  - 更多示例应用 361
- 第12章 NumPy高级应用 368
  - ndarray对象的内部机理 368
  - 高级数组操作 370
  - 广播 378
  - ufunc高级应用 383
  - 结构化和记录式数组 386
  - 更多有关排序的话题 388
  - NumPy的matrix类 393
  - 高级数组输入输出 395
  - 性能建议 397
- 附录A Python语言精要 401

# 《利用Python进行数据分析》

## 精彩短评

- 1、有时间会结合应用好好再读一遍。
- 2、这种书,最好有数据时的再去操作效果好些,一次性读完的话容易忘记。对于numpy、pandas有一些介绍。总感觉R进行数据处理还是更和谐些。
- 3、Pandas作者详实细致介绍了Python在数据处理方面的工作方式和方法,确实不同于我常用的SAS。学习期间,常常惊叹于Python的强大,后悔没有早点接触,非常认同那句话“人生苦短,我用Python”,而想要用Python进行数据处理,本书是目前的不二选择!
- 4、该书很需要更新,pandas进化太快了
- 5、入门书籍,pandas,numpy
- 6、罗列各种工具,不具备系统性
- 7、陆陆续续花了半年读了下来,实践下来发现,恩,貌似还用不着(逃。。。
- 8、翻一遍学习下
- 9、那种读完之后,在工作中天天当字典来查的书。但貌似pandas更新很快,rolling\_mean这种写法很快不支持了。搞得我天天加代码屏蔽warnings。
- 10、对于快速入门还可以
- 11、写得很清楚!另外第12章回应了我的很多好奇。冲着本书作者是pandas作者本人,加一星。
- 12、pd的作者写的内容作为入门很不错要是排版更清晰就更好了
- 13、学pandas如果英文有障碍,则这本书可能是最佳的了;英文没障碍就还是官方文档最佳。pandas的官方文档也很给力,提供odf下载,打印出来看不用对屏幕,很多库都没那么厚道,非得在线看。
- 14、比较实用也比较易懂
- 15、pandas库的操作指南,结合ipython;不过数据分析的坑太大了,其重点不再编程语言上,我还是老老实实学python后端吧
- 16、虽然这本书是讲Pandas的,但是对我印象最深刻的还是他关于NumPy的章节,比其他很多关于NumPy的资料好多了。
- 17、chapter 3 主旨:ipython 很好,大家快来用啊~2333  
numpy begin~  
不过听说这是pandas的实验手册?  
对了,译者吐槽也是醉醉哒
- 18、工具书,有点枯燥,不适合入门
- 19、不错。
- 20、pandas 工具书,如果手上有数据,用起来会很爽
- 21、书挺不错的,算是pandas的入门吧
- 22、介绍性质
- 23、这本书是Pandas的模块作者写的书,比较全,也很好入门。附录中的python语言精要部分,适合多看几遍。里面言简意赅的总结了python这门语言最高频使用到的一些语言要点。
- 24、pandas作者
- 25、入门书,思路最重要,等到spark的时候也是df这一套。除了里边给的示例数据下载不方便外没啥毛病。
- 26、入门级别的书。大概稍微懂点python都可以直接来看。其实主要就讲了两部分,numpy跟pandas,而且讲的还不全。
- 27、无论是作者还是译者,一定是敲了足够多的代码,才有功底写(译)成了该书。  
简单概括,IPython和pandas都值你花时间得去试试。  
最后,引自本书P401 知识是一座宝库,而实践就是开启这座宝库的钥匙。----Thomas Fuller
- 28、只是看书了,没有写代码。用的不多。
- 29、Pandas使用手册啊.可以一翻
- 30、介绍 numpy,scipy,pandas, matplotlib的. 必读.
- 31、2017第29本,这本书完全可以不看,pandas和canopy都改得面目全非了,完全不能使用;1.上手就

## 《利用Python进行数据分析》

在知乎上找几个案例照着做几遍就好 2. 查函数清单去官方手册上找

32、打吐了啊再来一遍 2017.3.6 又重新打了一遍 彻底作为工具书使用

33、Python涉及到的，可以做的绝对不止这些，但是如果只是要学会数据分析大概也是够了。这本书最棒的一点就是有实际例子且能找到对应的源数据文件，否则都不想动手吧

34、pandas 作者的书，当然主要都是围绕 pandas 讲的

35、我的Pandas数据分析入门书，唯一的缺点可能是赶不上库的更新速度，纸质书的通病

36、科普类给个大框架很给力。

37、混乱！！

38、pandas作者写的。浏览过一遍，现在发现pandas数据框已经成为很多公共接口的标准或默认数据类型了。这本书非常适合后续查阅。

39、在看的时候无时无刻都能感受到译者在翻译的时候的诚意。其实这个书更适合当做工具书来看。

40、较为基础

41、pandas的作者，python Data Science入门书籍

42、作为学习数据分析入门书很不错

43、pandas，实际过程我基本靠百度和pandas官网document 0.19.1搞定了全部，因为实在抽不出时间完整地先学再用。不过看此书收获更大，知识成体系。另外我不接受IPython的[In]、[Out]格式，使用纯python+库+Sublime

44、读的最久最认真的一本，感觉看完这一本，知识点全部掌握的话，对于数据挖掘方向的Python知识应该够用了，力荐！

45、不管怎么说，数据仓库与数据挖掘这门课算是告一段落了，一天写了32页论文。。。累

46、写的不错，看完基本就会用pandas了

47、pandas使用说明与示例,就是版本有点老了

48、作者是pandas作者，之前在AQR工作的

49、不错的入门书，对各个库都有简要介绍。可作为panda库的工具书。

50、本书作者也是pandas的作者，而在数据分析、数据处理中，pandas是使用频率非常高的一个库，于我目前的工作而言，90%都是使用它。本书主要介绍了pandas和numpy两个库的详细用法，也涉及到一点图表的制作，但内容很。由于pandas在不断更新，所以书中的用法有些是已经被丢弃了的。在使用的时候会有提醒。书中讲到了很多各个模块、各个函数的用法。读者需要根据自己的需要去学习和研究。最后，本书是难得的少见的python数据分析方面的好书了

# 《利用Python进行数据分析》

精彩书评



# 《利用Python进行数据分析》

## 版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:[www.tushu000.com](http://www.tushu000.com)