

<<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

出版时间：机械工业出版社

作者：Wes McKinney

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，请支持正版图书。

更多资源请访问：<http://www.tushu007.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，广受用户好评。

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

<<利用Python进行数据分析>>

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>>

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:<http://www.tushu007.com>