

版权注意事项：

- 1、书籍版权归作者和出版社所有
- 2、本PDF仅限用于个人获取知识，进行私底下的知识交流
- 3、PDF获得者不得在互联网上以任何目的进行传播
- 4、如觉得书籍内容很赞，请购买正版实体书，支持作者
- 5、请于下载PDF后24小时内删除本PDF。



担心自己基础差学不会？
买过好几本书，Python还没入门？

毫无障碍学 Python

邓文渊 著

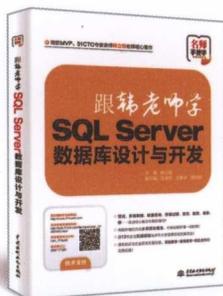
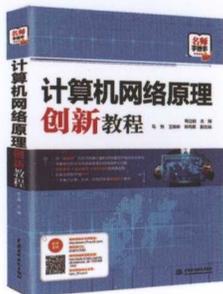
一点一例，Step by Step
一步一图，操作清晰明确
毫无障碍，简单超乎想像

网络爬虫、人脸识别、验证码识别、数据抓取、数据分析与图表制作、文件解析与文件操作、音频视频播放与控制、基本绘图、游戏设计、打包成.exe文件……

通过本书，你会发现，这些高营养技术瞬息变得易消化、好吸收！

邓文渊 像编写代码一样为初学者精确设计学习方法与内容的一线Python专家

优秀原创精品图书推荐



内 容 提 要

这不是一本简单教你学习Python语法、数据类型、程序结构的书。这也不是一本苦涩难懂、学而不会、例多无用的书。

本书从第一行Python代码开始，就让您轻松感受到学习Python的乐趣。本书经过一线Python开发大师的精心设计，巧妙地把Python的基础知识点融入到了流行、精彩但又经过精心处理的商业案例之中，突出了学习目标的同时，简化了学习难度。一些流行、重要但看似复杂的技术，如爬虫、人脸识别、网页实时数据抓取、文件自动化处理等，跟随本书您都可以轻松掌握并亲手实现。

本书基础讲解系统，案例设计突出精、新、小、实，特别适合于Python学习者自学使用或作为高校大数据、人工智能、云计算等相关专业的教材使用。

北京市版权局著作权合同登记图字：01-2017-4156

本书为基峰资讯股份有限公司独家授权出版发行的中文简体字版本。本书中文简体字版在中国大陆之专有出版权属于中国水利水电出版社所有。在没有得到本书原版出版者和本书出版者书面许可时，任何单位和个人不得擅自摘抄、复制本书的全部或一部分以任何方式包括（资料 and 出版物）进行传播。本书原版版权属基峰资讯股份有限公司。版权所有，侵权必究。

图书在版编目（CIP）数据

毫无障碍学Python / 邓文渊著. — 北京：中国水利水电出版社，2017.9
ISBN 978-7-5170-5834-2

I. ①毫… II. ①邓… III. ①软件工具—程序设计
IV. ①TP311.561

中国版本图书馆CIP数据核字(2017)第221280号

责任编辑：周春元 加工编辑：孙丹 封面设计：梁燕

书 名	毫无障碍学 Python HAOWUZHANGAI XUE Python
作 者	邓文渊 著
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路1号D座 100038) 网址: www.waterpub.com.cn E-mail: mchannel@263.net (万水) sales@waterpub.com.cn 电话: (010) 68367658 (营销中心)、82562819 (万水)
经 售	全国各地新华书店和相关出版物销售网点
排 版	北京万水电子信息有限公司
印 刷	三河市鑫金马印装有限公司
规 格	170mm×230mm 16开本 20.75印张 328千字
版 次	2017年9月第1版 2017年9月第1次印刷
印 数	0001—3000册
定 价	58.00元 (附1张DVD)

凡购买我社图书，如有缺页、倒页、脱页的，本社营销中心负责调换
版权所有·侵权必究

推荐序

——你会不会像我一样喜欢这本书？

●Python, 我对它好奇已很久

伴随着爬虫、云计算、大数据、人工智能等一波波狂欢风口，其中有一个让任何“码界新老民工”都无法忽视的存在，那就是Python。不就是一种开发语言吗？但是：

爬虫，为什么是Python，而不是C、Java……？

云计算，为什么是Python，而不是C、Java……？

大数据，为什么是Python，而不是C、Java……？

人工智能，为什么是Python，而不是C、Java……？

更让人诧异的是，不久前我与著名嵌入式开发专家杨国华教授一起讨论工作时了解到，Python在嵌入式开发领域也正在攻城略地，大有燎原之势。嵌入式开发领域不是C家的后院么？

作为一名专职的计算机图书编辑，于公于私，我觉得是时候要深入了解一下Python了。当然，基于我对自己深刻了解（缺脑），我采用了最笨但也最适合我的方法去了解Python，那就是亲自去学一学、用一用。所以，依照网上书店的销售排行买了我的第一本英文翻译版的Python书，结果枯涩且错误百出的翻译耗掉了我为数不多的耐心。我又选了些国产的Python书，或枯燥或无用或超出了我的理解能力……我的要求不高，你只要能让我顺顺利利地比照书上内容做出点东西来就行，如爬虫、抓个网页或者做个人脸识别什么的……

但作为编辑，而且是专业的计算机图书编辑，我对图书的这种状况虽然不满但还是能理解。我知道，想出版一本能让没有编程经验的人好学好懂又有趣的Python书，肯定是不容易的。

●我很喜欢这个简单

由于工作关系，台湾的著名计算机出版公司碁峰公司经常会给我发一些他们最

新的图书出版信息，其中我发现了一本讲解Python的书。我大概看了一下简介，发现这本书带有视频讲解，另外从目录中我还发现里面包含了很多我非常感兴趣的内容，如爬虫、网络实时数据抓取、数据分析与展示、人脸识别、验证码破解、Word文件操作、绘图、游戏开发等。

内容看起来很不错，所以利用工作之便，让对方给我寄了一本样书，当然是免费的。我并没有信心以我的基础能读懂这样一本书，也做好了该书只是用花哨的主题来哗众取宠的思想准备。

习惯性地，我先打开随书光盘的第一段视频，这段视频讲的是集成开发环境的安装。根据我的经验，如果只讲安装而不讲配置，可能就说明作者并无太多实战经验。但看完以后，我发现整个Python开发环境的安装的确是根本不需要进行任何配置操作的，如系统路径等。按照书上讲的安装过程，不会出现一点意外。我印象中之前几次尝试学习Java和.Net时，都是在开发环境的安装这一步直接栽了——安装开发环境的过程中不是缺这个就是少那个……

我很怀疑，这么简单的安装操作后我是否真的就可以编写Python程序了。按照书中的操作步骤，我开始把本书第一个简单得不能再简单的小程序代码敲了进去：

```
a=1
b=2
sum=a+b
print("总和"+=str(sum))
```

它是那么简单！没看到期待中的变量、语法、程序结构的讲解。

它是那么简单！没有印象中的导入包、头文件，甚至没有变量声明，没有印象中的编译、连接。

它是那么简单！从敲代码到理解代码到运行出结果，前后不到1分钟时间。

更加令我怀疑人生的是，我仅以若干年前在大学里学习C语言时遗留的一些模糊知识，竟然对几行代码完全不感觉困惑……

简单，正是我喜欢的风格。感觉我好像可以跟随这本书走得远一点。

● 怀疑地走到最后

一步一步，终于到了我最激动的爬虫制作，此时依然担心我是否真的能按书上所讲亲手爬下一个网页数据，哪怕是一整个网页。事实上，这个过程中也确实遇到了问题，案例中所讲的网站地址是台湾地址，在大陆无法直接访问。百度了半天，

解决的办法其实超级简单，把代码中的网址随便换成一个大陆可以访问的网址就可以了……

我竟然做了一个爬虫出来!

又经过后面几个案例，我逐渐意识到本书精妙所在。这的确不是一本哗众取宠、为追求市场热点而编写的书：每个知识点都配有一个小得不能再小但可以独立运行的案例；知识点和案例的先后顺序经过了严格设计，形成了完美的由浅入深但绝不越界的内容结构；每章的综合案例尤其值得称赞，综合案例虽然长一点，却是恰到好处地由前面讲到基本点组成。当然最重要的是，我的的确确亲手将所有实验都做了出来，而且能够完全理解。

难以想象，我跟随全书竟然完成了爬虫、网页实时数据抓取、数据分析与图表制作、文件解析及操作、人脸识别、验证码识别、绘图、游戏以及Python文件的打包等所有案例。

期间，我甚至没有功夫去抱怨别扭的台湾话。

●以权谋私

台湾的暮峰公司免费送我这本书的目的是让我评估这本书是否可以在大陆出版，但我沉浸在学习与收获的快感之中，以至于都忘了给对方回复一个邮件。直到后来暮峰公司再次来信小心翼翼地（我们是甲方）询问此书是否可在大陆出版。

我是如此地喜欢这本书，没有假惺惺地摆一下谱就与对方签订了本书的出版合同。

接下来的工作更加忙碌，为了让读者早日能见到这本我最喜爱的Python书，我同步开始了本书出版准备工作：把繁体版翻译成简体版；把整书的全部程序再运行一遍，重抓实验图片从而把图片中的繁体字转为简体字；把光盘中所有代码中的繁体字变成简体字；把原书案例代码中注释不够详细的地方逐行加了注释（我习惯于假设读者都是我这种水平，注释也许啰嗦，但万一某个读者真与我水平一样甚至还没有我水平高呢？）；把一些大陆不能访问的实验案例改为大家可以顺利进行实验的案例……唯一抱歉的是，光盘中的视频我没法重录，好在作者的普通话还可以。

所有这些工作，正常来说在出版社都有专门的人进行处理。我以权谋私全部亲手给包了。一是为了再次享受本书的精彩以对所学知识加以巩固，二是想把我在学习过程中觉得不顺畅的地方变得顺畅。

由于本人水平毕竟有限，错误之处请大家批评指正。书中遇到任何问题可加本

人QQ (172559140) 进行讨论。

●本书适合谁

本书特别适合于多次尝试学习Python、多次购买Python图书但都半途而废的读者。

本书特别适合于想学Python但却担心自己基础太差而学不会的读者。

本书特别适合于想用Python 做点真玩意出来的读者。

但本书不适合于Python资深码农。

希望本书能够真正助您步入Python殿堂。

本书读者、译者、策划编辑

周春元

学习资源说明

为了确保您在学习本书内容时能得到完整的学习效果，并能快速练习或观看范例效果，本书在光盘中提供了许多相关的配套内容供读者练习与参考。

光盘内容

1. **本书范例**：将各章范例的完成文件依章节名称放置各数据夹中。
2. **教学影片**：在应用程序开发过程中的许多学习重点，有时经由教学影片的导引，会胜过阅读大量的说明文字。作者特别针对书本中较为繁琐但在操作上十分重要的地方，录制成教学影片。读者可以依照影片里的操作，搭配书中的说明进行学习，相信会有事半功倍的效果。

有提供教学影片的章节，在目录会有一个影片图示，读者可以对照使用。

专属网站资源

为了加强读者服务，并持续更新书上相关的内容，我们特提供本系列丛书的相关网站资源，读者可以由文章列表中取得书中的勘误、更新或相关信息消息，更欢迎您加入我们的粉丝团，让所有信息一次到位不漏接。

藏经阁专栏 <http://blog.e-happy.com.tw/?tag=程序特训班>

程序特训班粉丝团 <https://www.facebook.com/eHappyTT>

注意事项

本光盘内容是提供给读者自我练习以及培训机构于教学时练习之用，版权分属于文渊阁工作室与提供原始程序档案的各公司所有，请勿复制本光盘做其他用途。

本书目录

推荐序
学习资源说明

Chapter	轻松配置 Python 开发环境	1
01	1.1 Python 编程语言简介	2
	1.1.1 Python 发展史.....	2
	1.1.2 Python 特色.....	3
	1.2 内置的 Anaconda 开发环境	4 
	1.2.1 安装 Anaconda 套件.....	4
	1.2.2 Anaconda Prompt 套件.....	6
	1.2.3 在命令窗口运行一个 Python 程序.....	9
	1.2.4 用 Anaconda Prompt 建立虚拟环境.....	9
	1.2.5 IPython 交互式命令窗口.....	12
	1.3 Spyder 编辑器	14 
	1.3.1 启动 Spyder 编辑器及文件管理.....	15
	1.3.2 Spyder 智能输入.....	17
	1.3.3 程序调试.....	17
	1.4 Jupyter Notebook 编辑器	19 
	1.4.1 启动 Jupyter Notebook 及建立文件.....	19
	1.4.2 Jupyter Notebook 智能输入.....	20
	1.4.3 用 Jupyter Notebook 运行程序.....	21
	1.4.4 在线运行 Python.....	22
Chapter	基本语法与结构控制	23
02	2.1 变量与数据类型	24
	2.1.1 变量.....	24
	2.1.2 变量命名规则.....	25

58	2.1.3 数值、布尔与字符串数据类型.....	26
58	2.1.4 print 及 type 命令.....	27
46	2.1.5 数据类型的转换.....	29
28	2.2 表达式.....	30
38	2.2.1 input 命令.....	30
48	2.2.2 算术运算符.....	31
17	2.2.3 关系运算符.....	31
75	2.2.4 逻辑运算符.....	32
57	2.2.5 复合赋值运算符.....	33
Chapter	2.3 条件语句.....	35
06	2.3.1 程序流程控制.....	35
77	2.3.2 单向条件 (if...).....	35
87	2.3.3 双向条件语句 (if...else).....	37
87	2.3.4 多向条件语句 (if...elif...else).....	38
87	2.3.5 条件嵌套.....	40
Chapter	循环、数据结构及函数.....	43
03	3.1 循环.....	44
19	3.1.1 列表 (List) 结构.....	44
89	3.1.2 range 函数.....	45
Chapter	3.1.3 for 循环.....	46
07	3.1.4 for 循环嵌套.....	48
102	3.1.5 break 及 continue 命令.....	49
100	3.1.6 for...if...else 循环.....	50
107	3.1.7 while 循环.....	52
101	3.2 列表、元组结构及字典结构.....	53
801	3.2.1 高级列表操作.....	54
	3.2.2 元组 (Tuple).....	56
	3.2.3 字典 (Dict).....	58
	3.2.4 高级字典操作.....	59

3.3 函数	62
3.3.1 自定义函数	62
3.3.2 不定参函数	64
3.3.3 变量有效范围	65
3.3.4 系统内置函数	66
3.3.5 导入包	69

Chapter 04

文件处理及 SQLite 数据库	71
4.1 文件和目录管理	72
4.1.1 os	72
4.1.2 os.path	73
4.1.3 os.walk	75
4.1.4 shutil 包	76
4.1.5 glob 包	77
4.2 open() 打开文件	78
4.2.1 用 open() 函数打开文件	78
4.2.2 文件操作	81
4.2.3 文件操作的应用	85
4.3 SQLite 数据库	91
4.3.1 管理 SQLite 数据库	91
4.3.2 用 SQLite Manager 创建 SQLite 数据库	93
4.3.3 使用 sqlite3 包	94
4.3.4 用 cursor 进行数据查询	97
4.3.5 SQLite 数据库实战	98

Chapter 05

Web 数据抓取与分析	105
5.1 网址解析	106
5.2 网页数据抓取	107
5.2.1 用 requests 抓取网页源代码	107
5.2.2 搜索指定字符串	108

5.2.3	用正则表达式抓取网页内容	109
5.3	小试网页分析	114
5.3.1	HTML 网页结构	114
5.3.2	从网页开发界面查看网页源代码	115
5.3.3	通过鼠标右键查看网页源代码	116
5.3.4	试用 BeautifulSoup 进行网页抓取与解析	117
5.4	牛刀初试——编写你自己的网络爬虫	121
5.4.1	抓取北京市 PM2.5 实时数据	121
5.4.2	爬取指定网站的图片	124

Chapter	实现自动网页测试	127
06	6.1 检查网站数据是否更新	128
6.1.1	用 hashlib 判别文件是否有过更新	128
6.1.2	用 md5 检查网站内容是否更新	128
6.1.3	再试牛刀：抓取网络公开数据	129
6.2	通过“任务计划程序”实现自动下载	133
6.3	用 Selenium 组件实现浏览器操作自动化	137
6.3.1	Selenium 组件	138
6.3.2	查找网页元素	140
6.3.3	应用：自动登录网站	142

Chapter	绘制图形	145
07	7.1 Matplotlib 组件	146
7.1.1	Matplotlib 基本绘图	146
7.1.2	plot 方法的参数及图形设置	147
7.1.3	在 Matplotlib 中显示中文	150
7.1.4	绘制柱状图及饼图	151
7.1.5	实战：爬取我国 1990 年到 2016 年 GDP 数据并绘图显示 ..	154
7.2	Bokeh 组件	156
7.2.1	Bokeh 基本绘图	156

	7.2.2 line 方法的参数及图形属性设置.....	158
	7.2.3 散点图.....	161
	7.2.4 实战：用 Bokeh 绘制我国 GDP 数据统计图.....	163
Chapter	实战：Word 文件操作.....	167
08	8.1 用 Win32com 组件操作 Word 文件.....	168
	8.1.1 实现新建文件并保存.....	168
	8.1.2 打开文件并显示文件内容.....	171
	8.1.3 设置范围的格式.....	172
	8.1.4 表格处理.....	174
	8.1.5 向 Word 文件中插入图片.....	176
	8.1.6 自动查找替换 Word 文件中的指定文字.....	177
	8.2 实战：菜单自动生成器及批处理替换文字.....	179 
	8.2.1 实战一：自动生成菜单 Word 文件.....	179
	8.2.2 实战二：批量替换 Word 文件中的文字.....	183
Chapter	实战：PM2.5 实时监测显示器.....	185
09	9.1 Pandas：强大的数据处理套件.....	186
	9.1.1 创建 DataFrame 数据.....	186
	9.1.2 读取 DataFrame 数据.....	188
	9.1.3 DataFrame 数据的修改及排序.....	192
	9.1.4 删除 DataFrame 数据.....	193
	9.1.5 导入数据.....	194
	9.1.6 绘制图形.....	197
	9.2 实战：PM2.5 实时监测显示器.....	198 
	9.2.1 应用程序总览.....	198
	9.2.2 PM2.5 公开数据.....	199
	9.2.3 用户数据接口配置.....	200
	9.2.4 事件处理及函数.....	202

Chapter

10

实战：人脸识别及验证码图片破解 205

10.1 OpenCV：人脸识别应用..... 206

10.1.1 用 OpenCV 读取和显示图形 206

10.1.2 保存图像文件 209

10.1.3 OpenCV 基本绘图 210

10.1.4 用 OpenCV 进行人脸识别 212

10.1.5 抓取脸部图形及保存 215

10.1.6 抓取摄像头视频图像 217

10.1.7 实战：通过人脸识别进行登录 219 

10.2 用 Tesseract 识别验证码 223

10.2.1 简单的 OCR-Tesseract 包 223

10.2.2 验证码识别的原理 225

10.2.3 实战：验证码破解 227 

Chapter

11

实战：Firebase 实时数据库应用 231

11.1 Firebase 实时数据库 232

11.1.1 Firebase 实时数据库简介 232

11.1.2 创建 Firebase 实时数据库 232

11.1.3 新增 Firebase 实时数据库资料 234

11.1.4 设置数据库权限 235

11.2 建立 Firebase 数据库连接 236

11.2.1 使用 python-firebase 包 236

11.2.2 firebase 对象的方法 236

11.3 实战：Firebase 版电子词典 245

11.3.1 英汉词典标准版 246

11.3.2 英汉词典进阶版 253

Chapter 12	实战：批量更改文件夹、文件名及文件查找 259
	12.1 文件管理应用 260
	12.1.1 实战：根据指定的编号保存文件 260
	12.1.2 实战：大批文件复制搬移及重新命名 262 
	12.1.3 实战：找出重复的照片 264
	12.1.4 实战：把图片文件改为相同大小 266
	12.2 在多个文件中查找指定的字符 268
	12.2.1 实战：在多个文本文件中查找 268 
	12.2.2 实战：在 Word 文件中查找指定字符 269
Chapter 13	实战：音乐播放器 273
	13.1 关于音乐与音乐的播放 274
	13.1.1 使用 pygame 包 274
	13.1.2 mixer 对象 274
	13.2 音效播放 274
	13.2.1 Sound 对象 274
	13.2.2 实战：制作一个音效播放器 275 
	13.3 音乐播放 278
	13.3.1 music 对象 278
	13.3.2 实战：制作 MP3 音乐播放器 278 
Chapter 14	实战：Pygame 游戏开发 283
	14.1 Pygame 入门 284
	14.1.1 Pygame 的基本架构 284
	14.1.2 基本绘图 286
	14.1.3 加载图片 288
	14.1.4 插入文本 290
	14.2 Pygame 动画——让图片动起来 290