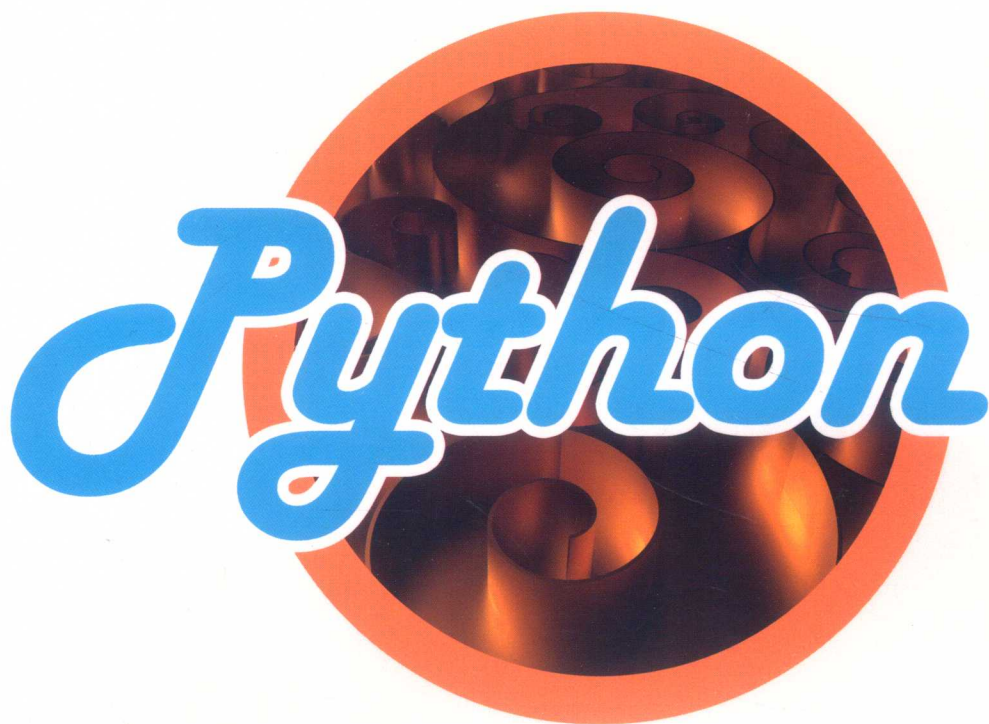


新工科·普通高校程序设计优秀教材



# Python编程指南

语法基础、网络爬虫、数据可视化与项目实战

关东升◎著

清华大学出版社





# Python编程指南

语法基础、网络爬虫、数据可视化与项目实践



关东升◎著

清华大学出版社

北京

## 内 容 简 介

本书是为高校师生学习 Python 编程语言而设计编著的教材。全书分为 20 章,其中包括绪论,搭建开发环境,第一个 Python 程序,Python 语法基础,数据类型,运算符,控制语句,数据结构,函数,面向对象编程,异常处理,常用模块,正则表达式,文件操作与管理,数据库编程,网络编程,wxPython 图形用户界面编程,Python 多线程编程,以及最后给出的两个实战项目——“项目实战 1:网络爬虫与爬取股票数据”和“项目实战 2:数据可视化与股票数据分析”。每一章后面都安排了若干同步练习题,并在附录 A 中提供了参考答案。

本书既可作为高等学校计算机软件技术课程的教材,也可作为社会培训机构的培训教程,还适用于广大 Python 初学者和 Python 开发的程序员等。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

### 图书在版编目(CIP)数据

Python 编程指南:语法基础\*、网络爬虫、数据可视化与项目实战/关东升著. —北京:清华大学出版社, 2019

ISBN 978-7-302-53133-3

I. ①P… II. ①关… 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(2019)第 110185 号

责任编辑:盛东亮 赵晓宁

封面设计:李召霞

责任校对:李建庄

责任印制:李红英

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, [c-service@tup.tsinghua.edu.cn](mailto:c-service@tup.tsinghua.edu.cn)

质量反馈:010-62772015, [zhiliang@tup.tsinghua.edu.cn](mailto:zhiliang@tup.tsinghua.edu.cn)

课件下载: <http://www.tup.com.cn>, 010-62795954

印 装 者:三河市君旺印务有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:23.5

字 数:572千字

版 次:2019年9月第1版

印 次:2019年9月第1次印刷

定 价:69.00元

产品编号:082637-01

## 作者简介

**关东升** 一个在 IT 领域摸爬滚打 20 多年的老程序员、培训师、作者。精通多种 IT 技术。参与设计和开发北京市公交一卡通大型项目，开发国家农产品追溯系统、金融系统微博等移动客户端项目，并在 App Store 发布多款游戏和应用软件。近期为中国移动、中国联通、南方航空、中石油、工商银行、平安银行和天津港务局等企事业单位授课。著有《iOS 开发指南》《从零开始学 Swift》《Android 网络游戏开发实战》《Cocos2d-x 实战》《JSP 网络程序设计》等 40 多本计算机书籍。



# 前言

## PREFACE

Python 语言诞生至今经历了将近 30 年时间,但是在前 20 年里,国内使用 Python 语言进行软件开发的程序员并不多,而在近 5 年,人们对 Python 语言的关注度迅速提高,这并不仅仅是因为 Python 语言非常优秀,而是 Python 语言满足了当下科学计算、人工智能、大数据和区块链等新技术的发展需要。Python 语言是一种“胶水”语言,具有丰富的动态特性、简单的语法结构和面向对象的编程特点,并拥有成熟而丰富的第三方库。Python 语言适合于除硬件开发之外的几乎所有领域的软件开发。

本书是智捷课堂开发的又一本立体化图书,配套了课件和服务等内容。本书每一章节知识点安排合理,讲解得非常细致,非常适合零基础读者学习 Python 语言,如果读者能够认真学习本书,将能独立开发 Python 网络爬虫和数据可视化项目。

### 本书服务网址

为了更好地为广大读者提供服务,我们专门为本书建立了一个服务网址 <http://www.zhijieketang.com/group/9>,希望读者对书中内容发表评论,提出宝贵意见。

### 下载源代码

书中包括了 200 多个完整的案例项目源代码,大家可以到本书网站 <http://www.zhijieketang.com/group/9> 免费注册并下载。

### 源代码目录结构

本书作为一本介绍 Python 编程的书,提供了很多示例源代码。下载本书源代码并解压,读者会看到如图 0-1 所示的目录结构。图 0-1 中的 chapter3~chapter20 是本书第 3 章~第 20 章配套示例源代码, ch6.6.1.py 文件则是第 6.6.1 节示例代码。

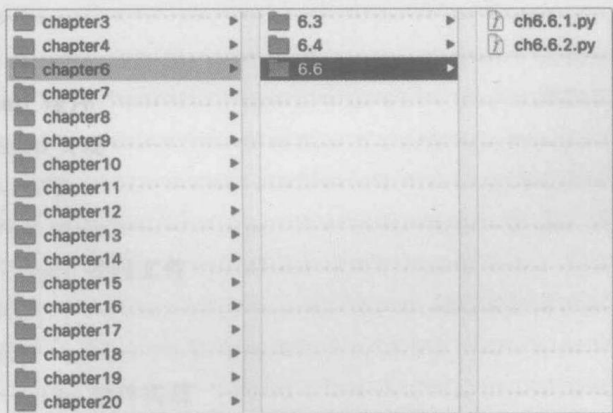


图 0-1 示例源代码目录结构

### 勘误与支持

如果读者发现了任何问题，均可以通过清华大学出版社电子信息读者服务 QQ 群 628808216 联系编辑，我们会在第一时间给予回复。

### 致谢

感谢清华大学出版社的盛东亮编辑给本书提出的宝贵意见。感谢智捷课堂团队的赵志荣、赵大羽、关锦华、闫婷娇、王馨然、关秀华、闫喜华和赵浩丞参与本书内容讨论和审核。感谢赵浩丞手绘了书中全部草图，并从专业的角度修改书中图片，力求将这些图片更加真实完美地奉献给广大读者。感谢我的家人容忍我的忙碌和对我的关心及照顾，使我能抽出时间，投入全部精力专心地编写此书。

由于时间仓促，书中难免存在不妥之处，敬请读者批评指正。

关东升

2019年7月

本书在编写过程中，参考了以下书籍和资料：

封面设计：李江英  
 责任编辑：李桂庄  
 责任印制：李江英

本书作为一本介绍 Python 编程的书，提供了从 Python 编程入门到进阶的完整知识体系。本书共分 20 章，主要内容包括：Python 语言基础、数据类型、运算符、控制语句、数据结构、函数、面向对象编程、异常处理、常用模块、正则表达式、文件操作与管理、数据库编程、网络编程、图形用户界面编程等。

地址：北京清华大学学研大厦A座  
 邮编：100084  
 社总机：010-62770175  
 投稿与读者服务：010-62776969  
 邮购服务：010-62776969  
 读者信箱：zgdx@tup.tsinghua.edu.cn  
 版权保护：010-62776969  
 印装者：三河市君旺印务有限公司  
 经销：全国新华书店  
 开本：185mm×280mm 印张：23.5  
 版次：2019年9月第1版  
 定价：69.00元

产品编号：082637-01

010-62786544  
 572千字  
 2019年9月第1次印刷

清华大学出版社

## 图书资源支持

感谢您一直以来对清华版图书的支持和爱护。为了配合本书的使用,本书提供配套的资源,有需求的读者请扫描下方二维码,在图书专区下载,也可以拨打电话或发送电子邮件咨询。

如果您在使用本书的过程中遇到了什么问题,或者有相关图书出版计划,也请您发邮件告诉我们,以便我们更好地为您服务。

### 我们的联系方式:

地址:北京市海淀区双清路学研大厦 A 座 701

邮编:100084

电话:010-62770175-4608

资源下载:<http://www.tup.com.cn>

客服邮箱:[tupjsj@vip.163.com](mailto:tupjsj@vip.163.com)

QQ: 2301891038 (请写明您的单位和姓名)

教学交流、课程交流



清华电子



扫一扫,获取最新目录

用微信扫一扫右边的二维码,即可关注清华大学出版社公众号“清华电子”。

## 本书涉及的主题

Python语法基础

数据类型

运算符

控制语句

数据结构

函数

面向对象编程

异常处理

常用模块

正则表达式

文件操作与管理

数据库编程

网络编程

图形用户界面编程

Python多线程编程



# 目录

## CONTENTS

<b>第 1 章 绪论</b> .....	1
1.1 Python 语言历史 .....	1
1.2 Python 语言设计哲学——Python 之禅 .....	2
1.3 Python 语言特点 .....	3
1.4 Python 语言应用前景 .....	4
1.5 如何获得帮助 .....	5
1.6 同步练习 .....	5
<b>第 2 章 搭建开发环境</b> .....	6
2.1 搭建 Python 环境 .....	6
2.2 PyCharm 开发工具 .....	9
2.2.1 下载和安装 .....	9
2.2.2 设置 Python 解释器 .....	11
2.3 Eclipse+PyDev 开发工具 .....	13
2.3.1 JDK 下载和安装 .....	14
2.3.2 设置环境变量 .....	15
2.3.3 Eclipse 下载和安装 .....	18
2.3.4 安装 PyDev 插件 .....	20
2.3.5 设置 Python 解释器 .....	24
2.3.6 设置 UTF-8 编码 .....	24
2.4 Visual Studio Code 开发工具 .....	26
2.5 文本编辑工具 .....	29
2.6 本章小结 .....	30
2.7 同步练习 .....	30
<b>第 3 章 第一个 Python 程序</b> .....	31
3.1 使用 Python Shell 实现 .....	31
3.2 使用 PyCharm 实现 .....	33
3.2.1 创建项目 .....	33
3.2.2 创建 Python 代码文件 .....	34
3.2.3 编写代码 .....	36
3.2.4 运行程序 .....	36
3.3 使用 Eclipse+PyDev 插件实现 .....	36
3.3.1 创建项目 .....	37
3.3.2 创建 Python 代码文件 .....	38

3.3.3	运行程序	39
3.4	使用 Visual Studio Code 实现	41
3.4.1	创建 Python 代码文件	41
3.4.2	运行程序	41
3.5	使用文本编辑工具+Python 解释器实现	43
3.5.1	编写代码	43
3.5.2	运行程序	43
3.6	代码解释	44
3.7	本章小结	45
3.8	同步练习	45
<b>第 4 章</b>	<b>Python 语法基础</b>	<b>47</b>
4.1	标识符和关键字	47
4.1.1	标识符	47
4.1.2	关键字	47
4.2	变量和常量	48
4.2.1	变量	48
4.2.2	常量	48
4.3	注释	49
4.4	语句	49
4.5	模块	50
4.6	包	52
4.6.1	创建包	52
4.6.2	导入包	52
4.7	本章小结	53
4.8	同步练习	53
<b>第 5 章</b>	<b>数据类型</b>	<b>54</b>
5.1	数字类型	54
5.1.1	整数类型	54
5.1.2	浮点类型	55
5.1.3	复数类型	55
5.1.4	布尔类型	55
5.2	数字类型相互转换	56
5.2.1	隐式类型转换	56
5.2.2	显式类型转换	57
5.3	字符串类型	57
5.3.1	字符串表示方式	58
5.3.2	字符串格式化	60
5.3.3	字符串查找	61
5.3.4	字符串与数字相互转换	63
5.4	本章小结	64
5.5	同步练习	64
<b>第 6 章</b>	<b>运算符</b>	<b>65</b>
6.1	算术运算符	65

6.1.1	一元运算符	65
6.1.2	二元运算符	65
6.2	关系运算符	66
6.3	逻辑运算符	68
6.4	位运算符	69
6.5	赋值运算符	71
6.6	其他运算符	72
6.6.1	同一性测试运算符	72
6.6.2	成员测试运算符	73
6.7	运算符优先级	74
6.8	本章小结	75
6.9	同步练习	75
<b>第7章 控制语句</b>		<b>76</b>
7.1	分支语句	76
7.1.1	if 结构	76
7.1.2	if-else 结构	77
7.1.3	elif 结构	78
7.1.4	三元运算符替代品——条件表达式	79
7.2	循环语句	79
7.2.1	while 语句	79
7.2.2	for 语句	80
7.3	跳转语句	81
7.3.1	break 语句	82
7.3.2	continue 语句	82
7.3.3	while 和 for 中的 else 语句	83
7.4	使用范围	84
7.5	本章小结	84
7.6	同步练习	84
<b>第8章 数据结构</b>		<b>86</b>
8.1	元组	86
8.1.1	序列	86
8.1.2	创建元组	89
8.1.3	访问元组	90
8.1.4	遍历元组	91
8.2	列表	91
8.2.1	创建列表	91
8.2.2	追加元素	92
8.2.3	插入元素	93
8.2.4	替换元素	93
8.2.5	删除元素	93
8.2.6	其他常用方法	94
8.2.7	列表推导式	95
8.3	集合	96

28	8.3.1	创建可变集合	96
30	8.3.2	修改可变集合	97
30	8.3.3	遍历集合	98
33	8.3.4	不可变集合	98
33	8.3.5	集合推导式	99
37	8.4	字典	99
37	8.4.1	创建字典	100
37	8.4.2	修改字典	101
37	8.4.3	访问字典	102
37	8.4.4	遍历字典	103
37	8.4.5	字典推导式	104
37	8.5	本章小结	104
37	8.6	同步练习	104
	<b>第9章</b>	<b>函数</b>	<b>106</b>
37	9.1	定义函数	106
37	9.2	函数参数	107
37	9.2.1	使用关键字参数调用函数	107
37	9.2.2	参数默认值	107
37	9.2.3	可变参数	108
37	9.3	函数返回值	109
38	9.3.1	无返回值函数	109
38	9.3.2	多返回值函数	110
38	9.4	函数变量作用域	111
38	9.5	生成器	112
38	9.6	嵌套函数	114
38	9.7	函数式编程基础	115
38	9.7.1	函数类型	115
38	9.7.2	Lambda 表达式	116
38	9.7.3	三大基础函数	117
38	9.8	本章小结	119
38	9.9	同步练习	119
	<b>第10章</b>	<b>面向对象编程</b>	<b>121</b>
38	10.1	面向对象概述	121
38	10.2	面向对象三个基本特性	121
38	10.2.1	封装性	121
38	10.2.2	继承性	122
38	10.2.3	多态性	122
38	10.3	类和对象	122
38	10.3.1	定义类	122
38	10.3.2	创建和使用对象	122
38	10.3.3	实例变量	123
38	10.3.4	类变量	124
38	10.3.5	构造方法	126



10.3.6	实例方法	126
10.3.7	类方法	127
10.3.8	静态方法	128
10.4	封装性	129
10.4.1	私有变量	129
10.4.2	私有方法	130
10.4.3	定义属性	130
10.5	继承性	132
10.5.1	继承概念	132
10.5.2	重写方法	133
10.5.3	多继承	134
10.6	多态性	135
10.6.1	多态概念	135
10.6.2	类型检查	136
10.6.3	鸭子类型	137
10.7	Python 根类——object	138
10.7.1	__str__()方法	138
10.7.2	对象比较方法	139
10.8	枚举类	140
10.8.1	定义枚举类	140
10.8.2	限制枚举类	141
10.8.3	使用枚举类	141
10.9	本章小结	142
10.10	同步练习	142
<b>第 11 章 异常处理</b>		<b>143</b>
11.1	异常问题举例	143
11.2	异常类继承层次	144
11.3	常见异常	145
11.3.1	AttributeError 异常	145
11.3.2	OSError 异常	146
11.3.3	IndexError 异常	146
11.3.4	KeyError 异常	147
11.3.5	NameError 异常	147
11.3.6	TypeError 异常	147
11.3.7	ValueError 异常	148
11.4	捕获异常	148
11.4.1	try-except 语句	148
11.4.2	多 except 代码块	150
11.4.3	try-except 语句嵌套	151
11.4.4	多重异常捕获	152
11.5	异常堆栈跟踪	153
11.6	释放资源	154
11.6.1	finally 代码块	154



11.6.2	else 代码块	155
11.6.3	with as 代码块自动管理资源	156
11.7	自定义异常类	157
11.8	显式抛出异常	157
11.9	本章小结	158
11.10	同步练习	158
<b>第 12 章</b>	<b>常用模块</b>	<b>160</b>
12.1	math 模块	160
12.1.1	舍入函数	160
12.1.2	幂和对数函数	161
12.1.3	三角函数	161
12.2	random 模块	162
12.3	datetime 模块	164
12.3.1	datetime、date 和 time 类	164
12.3.2	日期和时间计算	167
12.3.3	日期和时间格式化和解析	168
12.3.4	时区	169
12.4	logging 日志模块	170
12.4.1	日志级别	170
12.4.2	日志信息格式化	172
12.4.3	日志重定位	173
12.4.4	配置文件的使用	174
12.5	本章小结	176
12.6	同步练习	176
<b>第 13 章</b>	<b>正则表达式</b>	<b>177</b>
13.1	正则表达式字符串	177
13.1.1	元字符	177
13.1.2	字符转义	178
13.1.3	开始与结束字符	178
13.2	字符类	179
13.2.1	定义字符类	179
13.2.2	字符类取反	180
13.2.3	区间	180
13.2.4	预定义字符类	181
13.3	量词	182
13.3.1	量词的使用	182
13.3.2	贪婪量词和懒惰量词	184
13.4	分组	184
13.4.1	分组的使用	184
13.4.2	分组的命名	185
13.4.3	反向引用分组	186
13.4.4	非捕获分组	187
13.5	re 模块	188

13.5.1	search()和 match()函数	188
13.5.2	findall()和 finditer()函数	189
13.5.3	字符串分割	190
13.5.4	字符串替换	191
13.6	编译正则表达式	191
13.6.1	已编译正则表达式对象	192
13.6.2	编译标志	193
13.7	本章小结	197
13.8	同步练习	197
<b>第 14 章</b>	<b>文件操作与管理</b>	198
14.1	文件操作	198
14.1.1	文件打开	198
14.1.2	文件关闭	200
14.1.3	文本文件读写	201
14.1.4	二进制文件读写	202
14.2	os 模块	202
14.3	os.path 模块	204
14.4	本章小结	205
14.5	同步练习	205
<b>第 15 章</b>	<b>数据库编程</b>	206
15.1	数据持久化技术概述	206
15.2	MySQL 数据库管理系统	207
15.2.1	数据库安装与配置	207
15.2.2	连接 MySQL 服务器	212
15.2.3	常见的管理命令	214
15.3	Python DB-API	218
15.3.1	建立数据连接	218
15.3.2	创建游标	219
15.4	案例: MySQL 数据库 CRUD 操作	220
15.4.1	安装 PyMySQL 模块	220
15.4.2	数据库编程的一般过程	221
15.4.3	数据查询操作	222
15.4.4	数据修改操作	224
15.5	NoSQL 数据存储	227
15.5.1	dbm 数据库的打开和关闭	227
15.5.2	dbm 数据存储	228
15.6	本章小结	229
15.7	同步练习	229
<b>第 16 章</b>	<b>网络编程</b>	230
16.1	网络基础	230
16.1.1	网络结构	230
16.1.2	TCP/IP 协议	231

16.1.3	IP 地址	231
16.1.4	端口	232
16.2	TCP Socket 低层次网络编程	232
16.2.1	TCP Socket 通信概述	232
16.2.2	TCP Socket 通信过程	232
16.2.3	TCP Socket 编程 API	233
16.2.4	案例：简单聊天工具	234
16.2.5	案例：文件上传工具	236
16.3	UDP Socket 低层次网络编程	238
16.3.1	UDP Socket 编程 API	238
16.3.2	案例：简单聊天工具	238
16.3.3	案例：文件上传工具	239
16.4	访问互联网资源	241
16.4.1	URL 概念	241
16.4.2	HTTP/HTTPS 协议	241
16.4.3	使用 urllib 库	242
16.4.4	发送 GET 请求	243
16.4.5	发送 POST 请求	244
16.4.6	实例：Downloader	245
16.5	本章小结	245
16.6	同步练习	245
<b>第 17 章</b>	<b>wxPython 图形用户界面编程</b>	<b>246</b>
17.1	Python 图形用户界面开发工具包	246
17.2	wxPython 安装	247
17.3	wxPython 基础	247
17.3.1	wxPython 类层次结构	248
17.3.2	第一个 wxPython 程序	248
17.3.3	wxPython 界面构建层次结构	251
17.4	事件处理	252
17.4.1	一对一事件处理	252
17.4.2	一对多事件处理	254
17.4.3	示例：鼠标事件处理	255
17.5	布局管理	256
17.5.1	Box 布局器	257
17.5.2	StaticBox 布局	259
17.5.3	Grid 布局	261
17.5.4	FlexGrid 布局	262
17.6	wxPython 控件	264
17.6.1	静态文本和按钮	264
17.6.2	文本输入	265
17.6.3	复选框和单选按钮	267
17.6.4	下拉列表	269
17.6.5	列表	270



17.6.6	静态图片	272
17.7	高级窗口	273
17.7.1	分隔窗口	273
17.7.2	使用树	275
17.7.3	使用网格	278
17.8	使用菜单	282
17.9	使用工具栏	284
17.10	本章小结	286
17.11	同步练习	286
<b>第 18 章</b>	<b>Python 多线程编程</b>	<b>288</b>
18.1	基础知识	288
18.1.1	进程	288
18.1.2	线程	289
18.2	threading 模块	289
18.3	创建线程	290
18.3.1	自定义函数作为线程体	290
18.3.2	继承 Thread 线程类实现线程体	292
18.4	线程管理	293
18.4.1	等待线程结束	293
18.4.2	线程停止	295
18.5	线程安全	296
18.5.1	临界资源问题	296
18.5.2	多线程同步	298
18.6	线程间通信	300
18.6.1	使用 Condition 实现线程间通信	301
18.6.2	使用 Event 实现线程间通信	304
18.7	本章小结	305
18.8	同步练习	305
<b>第 19 章</b>	<b>项目实战 1: 网络爬虫与爬取股票数据</b>	<b>307</b>
19.1	网络爬虫技术概述	307
19.1.1	网络通信技术	307
19.1.2	多线程技术	307
19.1.3	数据交换技术	308
19.1.4	Web 前端技术	308
19.1.5	数据分析技术	308
19.1.6	数据存储技术	310
19.2	爬取数据	310
19.2.1	网页中静态和动态数据	311
19.2.2	使用 urllib 爬取数据	311
19.2.3	使用 Selenium 爬取数据	316
19.3	分析数据	320
19.3.1	使用正则表达式	320
19.3.2	使用 BeautifulSoup 库	322