



福建省高等学校计算机规划教材

【第三版】

Python

程序设计与应用教程

福建省高等学校计算机教材编写委员会 组织编写

主编◎鄂大伟

福建省高等学校计算机规划教材

Python

程序设计与应用教程

福建省高等学校计算机教材编写委员会 组织编写

【第三版】

主 编：鄂大伟

副主编：陈志德 吴运兵 李丽珊

贾红伟 叶 双 陈 琼

傅 为 崔建峰 杨 岚



厦门大学出版社 国家一级出版社
XIAMEN UNIVERSITY PRESS 全国百佳图书出版单位

图书在版编目(CIP)数据

Python 程序设计与应用教程 / 鄂大伟主编. -- 3 版
. -- 厦门 : 厦门大学出版社, 2023.12
ISBN 978-7-5615-9268-7

I. ①P… II. ①鄂… III. ①软件工具-程序设计-
高等学校-教材 IV. ①TP311.561

中国版本图书馆CIP数据核字(2024)第013860号

策划编辑 宋文艳
责任编辑 眭蔚
美术编辑 李夏凌
技术编辑 许克华

出版发行 **厦门大学出版社**
社 址 厦门市软件园二期望海路 39 号
邮政编码 361008
总 机 0592-2181111 0592-2181406(传真)
营销中心 0592-2184458 0592-2181365
网 址 <http://www.xmupress.com>
邮 箱 xmup@xmupress.com
印 刷 厦门市明亮彩印有限公司

开本 787 mm × 1 092 mm 1/16
印张 24.25
字数 606 千字
版次 2019 年 8 月第 1 版 2023 年 12 月第 3 版
印次 2023 年 12 月第 1 次印刷
定价 56.00 元

本书如有印装质量问题请直接寄承印厂调换



厦门大学出版社
微信二维码



厦门大学出版社
微博二维码

序言：大道至简，众妙之门

“Life is short, you need Python.”这句话被 Python 语言的发明者吉多(Guido)印在自己的 T 恤上,目前已经成为众多 Python 学习者的时尚 Logo。

其实,早在 2010 年,Python 就被 TIOBE(权威的编程语言流行指数)评为年度最佳语言。特别是近年来,随着数据科学(大数据)、人工智能、机器学习等领域的兴起,Python 语言迅速成为最受欢迎的编程语言之一。在 IEEE Spectrum 发布的 2017 年和 2018 年编程语言排行榜中,Python 都居首位。

“大道至简,众妙之门”,道德经上的这句话用来形容 Python 的特点十分贴切。Python 语言的设计思想是“优雅、明确、简单”,更有著名的“Python 之禅”的哲学思想作为指导。Python 语言易于学习掌握,编写代码具有很高的效率,非常适合用来解决各专业领域的问题,是打开一切奥妙之门的钥匙。

Python 语言是开源代码,具有先天的开放基因,并且容易拓展。用 Python 语言或其他语言编写的模块可以轻松地扩展为 Python 程序可以调用的类库,这种可扩展性使得 Python 迅速拥有丰富的、开源的第三方库或模块的支持,构建了良好的计算生态。这些库是整个 Python 社区(Python's community)的贡献,大多数用户需要的功能都能通过已有的库完成。因此,编程者可以把精力集中在自己的专业领域问题上,而不必过多地考虑计算机底层实现的细节。

随着 Python 的日益普及,为推动福建省高校计算机基础教学改革,2018 年和 2019 年,中国高校计算机慕课联盟(CMOOC)福建区工作委员会、福建省计算机学会教育工作委员会联合福建省主要高校,共同主办了两场福建省高校计算机教学论坛。论坛研讨的主题主要是 Python 教学,福建省各个高校的代表就 Python 教学的内容和方法进行了深入交流,还推出了“1+X”的 Python 教学方案。另外,根据我们对省内主要高校所做的调查,目前许多高校已经开始调整计算机基础课程的教学方案,将 Python 语言作为非计算机类专业的首选程序设计语言,并积累了宝贵的教学经验。

在这个背景下,2018 年初,厦门大学出版社联合省内几所主要高校成立了 Python 教材编写组。为编写好本书,编写组教师们数次集中开会讨论,几易其稿。

经过近一年半的努力,如今本书终于与读者见面了。虽然说有些姗姗来迟,但也是生逢其时,因为这是福建省高校第一本 Python 教材,也是汇聚各高校 Python 教学经验与典型案例的经典之作。

本书的结构与内容按照“1+X”教学方案组织。这里的“1”是指 Python 语言程序设计基础部分,是通往 Python 奇妙世界的入门之道;“X”是指 Python 的高级应用部分,是漫游 Python 奇境的进阶之路。“X”部分的内容相当广泛,如 Python 科学计算、pandas 数据分析、数据可视化、网络爬虫技术、Python 与机器学习、构建与部署 Web 云主机等。“1+X”是一个灵活的教学方案,为不同专业和领域的教学提供了可选择的内容。

Python 的计算生态给我们以启迪,打造 Python 语言的教学生态圈是我们的追求目标。这里所说的教学生态圈,应该包括这样几个要素:教材+课程团队+MOOC+在线课程资源+程序在线评测平台。本书由多所高校合作编写,是课程团队合作的结果,同时也欢迎各学校更多的老师参与本课程建设(包括 MOOC 建设)。在教学资源方面,本书的所有行文代码、各个案例的完整代码以及相关数据文件,都发布在开源代码托管平台 GitHub 上,并以项目形式分享,以方便老师和同学们获得本课程资源。

本书由鄂大伟教授担任主编并负责统稿。在编写过程中,首先要感谢编写组的各位老师,他们是福建师范大学陈志德教授、福州大学吴运兵副教授、华侨大学叶双老师、福建农林大学李丽珊老师和陈琼副教授、集美大学贾红伟副教授、闽南师范大学傅为副教授、厦门理工学院崔建峰副教授。他们精心设计了一个个程序案例,竭尽全力地为读者奉献一道道 Beautiful_soup(借用 Python 的库名)。感谢福建省考试中心刘雄恩教授,他数十年如一日,为福建省计算机水平考试各语种提供机考系统的支持和服务。最后,要特别感谢厦门大学出版社宋文艳总编辑,没有她的执着努力与协调,也许本书的出版还停留在策划之中。

作为序言的结束语,真诚期望各高校在 Python 语言的教学过程中,对本教材提出宝贵的改进意见,以期冀收博见,嘉惠来学。

作 者

2019 年 7 月

本教材(课程)资源 GitHub 项目链接:

<https://github.com/David-Github-Project/Python-Curriculum-Resources>

第三版前言

时值 2023 年岁末,刚完成本教材第三版的修订,再次感叹一句“流水光阴急”。借此机会,回顾一下近三年来在教材和课程建设方面所做的一些工作,也是对本教材的建设做一个小结。

党的二十大报告指出,要“深化教育领域综合改革,加强教材建设与管理”。本教材第一版于 2019 年出版,先后被十余所高校采用并开展不同层次的课程教学;第二版于 2022 年出版,主要做了内容结构上的调整:将章节分为上篇和下篇两个部分,上篇主题是 Python 程序设计基础;下篇主题是 Python 高级应用,涉及多个可进一步拓展的应用方向。

在以教材为基础的课程建设方面,也取得了优异的成绩。“Python 程序设计与高级应用”2020 年获得福建省一流本科课程(闽教高[2020]23 号);“Python 高级应用”2023 年获得国家一流本科课程(教高函[2023]7 号)。这些成绩的取得,离不开教学团队的努力和广大读者的支持。

教学团队特别重视在线教学资源建设,“Python 程序设计与高级应用”的在线课程于 2020 年 2 月上线(学银在线平台),目前已经开设了 8 期教学,该 MOOC 被评为福建省 2023 年一流本科课程。

关于本教材的特色,在前两版本前言中已经做了较详细的介绍,这里仅就第三版教材的主要改动做个说明:

修订内容 1:第 12 章 Python 数据可视化

PyEchart 一节内容因版本升级,原有代码已经不能正常运行,所以进行了重新修订。

修订内容 2:第 13 章 网络爬虫

本章案例中有一个全球各国新冠疫情数据的爬虫案例,这个案例的原有 URLs 资源已经不能访问,且新冠疫情数据分析与当前形势已不相符,所以设计一个新案例替换原来案例。

修订内容 3:第 14 章 Python 与机器学习

新增加一个基于 MNIST 数据集的生成式对抗网络案例。

泉州信息工程学院杨岚副教授和许戈静老师参与了本次的修订工作。

优雅、清晰和简约是 Python 语言的一贯风格,也是作者努力追求的境界。“义典则弘,文约为美”,希望第三版教材以及在线课程资源能给读者带来新的收获和体验。

作者

2023 年 12 月 31 日于集美学村



学银在线课程资源:

<https://www.xueyinonline.com/detail/222545786>



码云 gitee 课程资源:

https://gitee.com/Course_Resources/Python-Programming



注:包括本教材各章行代码、完整案例代码和相关数据集

第二版前言

本书第一版 2019 年出版,感叹“流水光阴急”,如今第二版又与大家见面了。借此机会,回顾一下近三年来在教材和课程建设方面所做的一些工作,也算是一个小结。

第一版出版后得到同行的认可,共四次重印,先后有十余所高校采用本教材开展不同层次的课程教学。另外,以教材编委会成员为主体的教学团队,重视与福建省内各本科院校开展广泛的教学交流与合作。在近几年举办的福建省高校计算机教学论坛上,都有以 Python 教学为主题的经验交流,共同打造 Python 教学生态。

为响应教育部实施的一流课程“双万计划”的建设,教学团队以此为契机,按照一流课程建设标准进行混合式教学改革。Python 语言程序设计和高级应用两门课程获得福建省 2020 年一流本科课程。其中,Python 高级应用被推荐参加第二批国家一流本科课程评选。

2021 年 5 月,福建省高校第四届计算机教学论坛在泉州信息工程学院举办期间,教材编委会和部分高校代表召开了第 6 次 Python 教材建设研讨会,确定了第二版的修改内容和时间表。较之第一版,第二版主要的改动有:

(1)教材结构调整。根据教材内容,将章节分为上篇和下篇两个部分,上篇是 Python 程序设计基础,包括第 1 章至第 8 章内容;下篇是 Python 高级应用,包括第 9 章至第 14 章内容,涉及多个可进一步拓展的应用方向。

(2)章节顺序有所调整。将第一版的“第 5 章 程序控制结构”提前到第 3 章讲授,有利于后续案例和数据类型的引用。

(3)考虑到“新工科”背景下工科教育要求学生具备工程实践能力,能够利用计算机程序解决复杂的专业问题,加强了 Python 高级应用的内容介绍,概念性原理和应用案例都进行了重新设计。

教学团队特别重视在线教学资源建设,与集美大学郭晓曦副教授等老师的 Python 教学团队合作,设计开发了“Python 程序设计与高级应用”在线课程资源,并于 2020 年 2 月上线(超星的学银在线平台)。经过持续建设,在线课程提供的

授课视频总数量 95 个,视频总时长 838 分钟,还提供了丰富的课程学习资料,目前已开展五轮教学。在线课程的建设促进了信息技术与 Python 教学的融合,特别是为应对新冠肺炎疫情期间实施的在线教学做出了贡献。如今,结合 MOOC 平台开展线上线下混合式教学,已经成为新常态。

教学实践证明,在学习 Python 程序设计基础(上篇)之后,作为本教材的进阶内容,开设 Python 高级应用(下篇)课程,或引导学生结合各专业问题,学习本课程的在线 MOOC 资源,如科学计算、数据处理、数据可视化和机器学习等,对于提高学生解决本专业领域复杂问题的能力,及专业综合能力的提升具有较明显的作用。本课程的在线课程资源列于文末,欢迎各位老师和同学们参考使用。

优雅、清晰和简约是 Python 语言的一贯风格,也是作者努力追求的境界。“义典则弘,文约为美”,希望教材第二版以及在线课程资源能给读者带来新的收获和体验。

作者


2022 年 5 月于集美学村



学银在线课程资源:

<https://www.xueyinonline.com/detail/222545786>

The graphic is a dark grey vertical rectangle. At the top, it features the Python logo (two interlocking snakes) on the left and the text 'Python程序设计' and '与高级应用' on the right. Below this, the title 'Python程序设计与高级应用' is prominently displayed. Underneath the title, the instructor's name '郭晓曦 副教授 | 集美大学' is listed. A paragraph of text describes Python as a scripting language with simple syntax and powerful functionality, suitable for various professional learning. At the bottom left, there is a small '学银在线' logo and the text '扫描二维码, 开始学习课程'. At the bottom right, there is a QR code.

 码云 gitee 课程资源：

https://gitee.com/Course_Resources



The screenshot shows the Gitee profile page for the user 'My_Course_Resources'. On the left is a QR code. The profile header includes the user's name, a bio, and statistics: 1 Star, 2 Watches, 0 Followers, and 0 Following. The '热门项目' (Popular Projects) section lists two repositories: 'Python程序设计与高级应用第2版课程资源' and 'Python程序设计与应用第1版课程资源'. The '贡献度' (Contribution) section shows a calendar grid for the year 2022, with activity bars indicating contributions on specific days.

注：包括本教材各章行文代码、完整案例代码和相关数据集。

目 录

上篇 Python 程序设计基础

第 1 章 程序设计与 Python	1
1.1 程序设计语言	1
1.1.1 软件定义的时代	1
1.1.2 程序设计语言	2
1.1.3 编译程序与解释程序	3
1.1.4 计算思维与程序设计	4
1.2 Python 语言与计算生态	5
1.2.1 Python 语言的历史	5
1.2.2 Python 语言的特征	5
1.2.3 Python 的计算生态	7
1.3 Python 的安装与运行方式	9
1.3.1 Python 的下载与安装	9
1.3.2 Python 的运行方式	12
1.4 Python 开发者的编辑器和 IDE	13
1.4.1 Jupyter Notebook	14
1.4.2 PyCharm	15
本章习题	16
第 2 章 Python 基础	17
2.1 Python 程序基本构成与风格	18
2.1.1 认识第一个 Python 程序	18
2.1.2 Python 程序的基本结构	19
2.1.3 Python 程序规范	20
2.1.4 Python 之禅——大道至简	21
2.2 Python 语法基础	22
2.2.1 标识符	22
2.2.2 变 量	23
2.2.3 关键字	23
2.2.4 运算符与表达式	25

2.2.5 语句	28
2.3 Python 数据类型	28
2.3.1 数据类型	29
2.3.2 序列类型	31
2.3.3 字典类型	32
2.4 Python 的模块与函数	33
2.4.1 库与模块	33
2.4.2 第三方库的安装	34
2.4.3 导入模块	35
2.4.4 函数	36
2.4.5 包	37
本章习题	37
第3章 程序控制结构	38
3.1 顺序结构	38
3.1.1 赋值语句	38
3.1.2 基本输入和输出	42
3.1.3 顺序结构	44
3.2 选择结构	46
3.2.1 单分支 if 语句	46
3.2.2 双分支 if-else 语句	47
3.2.3 多分支 if-elif-else 语句	49
3.2.4 嵌套的 if 语句	52
3.3 循环结构	54
3.3.1 while 语句	54
3.3.2 for 语句	57
3.3.3 循环嵌套	59
3.3.4 break 语句与 continue 语句	61
3.3.5 循环结构中的 else 子句	63
3.3.6 实例	65
本章习题	67
第4章 序列	70
4.1 序列概述	71
4.1.1 索引	71
4.1.2 分片	71
4.1.3 通用序列类型操作	74
4.1.4 通用序列类型常用的内置函数	76
4.2 字符串	78

4.2.1	字符串的表示	78
4.2.2	字符串的格式化	80
4.2.3	字符串的方法	85
4.2.4	字符串实例	86
4.3	列表	87
4.3.1	列表的表示	87
4.3.2	列表的方法	88
4.3.3	列表实例	90
4.4	元组	93
4.4.1	元组的表示	93
4.4.2	元组的其他特性和作用	94
4.4.3	元组实例	96
4.5	range 对象	98
4.5.1	range 函数	98
4.5.2	range 实例	98
	本章习题	99
第 5 章	字典与集合	101
5.1	字典	102
5.1.1	创建字典	102
5.1.2	字典的基本操作和方法	102
5.1.3	字典实例	105
5.2	集合	108
5.2.1	创建集合	108
5.2.2	集合的基本操作和方法	109
5.2.3	集合实例	112
	本章习题	113
第 6 章	函数	115
6.1	函数的概念	116
6.1.1	函数的定义	116
6.1.2	函数的返回	117
6.1.3	匿名函数	117
6.2	函数的参数与调用	118
6.2.1	参数传递	118
6.2.2	参数的分类	119
6.2.3	函数改造示例	121
6.3	递归函数	122
6.3.1	递归模式	122

6.3.2 递归应用案例	123
6.4 内置模块	124
6.4.1 os 模块中的函数	124
6.4.2 random 模块中的函数	126
6.4.3 datetime 模块中的函数	128
6.5 变量的作用域	131
6.5.1 变量的作用域	131
6.5.2 作用域的类型	131
本章习题	133
第7章 异常捕捉	135
7.1 Python 中的异常	135
7.2 异常捕捉与处理	137
7.2.1 try-except 语句	137
7.2.2 多个 except 子句和一个 except 子句捕捉多个异常	138
7.2.3 else 子句与 finally 子句	143
本章习题	147
第8章 文 件	149
8.1 文件基本概念	150
8.1.1 Python 文件系统	150
8.1.2 文件系统的使用过程	151
8.2 文件的基本操作	153
8.2.1 文件的打开和关闭	153
8.2.2 文件的读写	154
8.2.3 文件的定位	155
8.2.4 文件的其他操作	157
8.3 各种类型文档的文件操作	157
8.3.1 txt 文件	158
8.3.2 csv 文件	158
8.3.3 JSON 文件	162
* 8.4 Python 图像库 PIL	164
8.4.1 PIL 案例	165
8.4.2 PIL 常用类	166
8.5 综合实例	169
8.5.1 文本分词	169
8.5.2 分词结果清洗	170
8.5.3 TopN 词汇输出	170
本章习题	173

下篇 Python 高级应用

第 9 章 面向对象程序设计	175
9.1 面向对象程序设计基本概念	175
9.1.1 面向对象程序设计	175
9.1.2 面向对象程序设计的基本特征	176
9.2 类与对象	176
9.2.1 类的定义	177
9.2.2 对象的定义	177
9.2.3 构造函数和析构函数	179
9.2.4 类属性与实例属性	180
9.3 继 承	182
9.3.1 子类的创建	182
9.3.2 访问控制	183
9.4 常用类和实例相关内置函数	185
本章习题.....	186
第 10 章 Python 科学计算	187
10.1 什么是科学计算.....	187
10.2 NumPy 简介	188
10.2.1 ndarray 的基本属性	188
10.2.2 创建 ndarray	189
10.2.3 ndarray 的操作和运算	192
10.3 SciPy 简介	196
10.3.1 常数模块(constants)	196
10.3.2 特殊函数模块(special)	196
10.3.3 线性代数运算模块(linalg)	197
10.3.4 优化和拟合模块(optimize)	199
10.3.5 统计与随机数模块(stats)	201
10.3.6 积分计算模块(integrate)	204
10.3.7 图像处理模块(ndimage)	206
本章习题.....	208
第 11 章 pandas 数据分析	210
11.1 pandas 简介与安装	210
11.1.1 pandas 简介	211
11.1.2 pandas 的安装与导入	211
11.2 pandas 数据结构——Series 与 DataFrame	211

11.2.1	pandas 数据结构	211
11.2.2	Series	212
11.2.3	DataFrame	217
11.3	pandas 数据清洗	226
11.3.1	删除无效列与行	227
11.3.2	重复值处理	227
11.3.3	空值与缺失值处理	228
11.3.4	空格处理	229
11.3.5	异常值处理	229
11.4	pandas 数据分析	230
11.4.1	导入所需库	230
11.4.2	数据加载与数据集统计信息查看	230
11.4.3	数据清洗	231
11.4.4	数据分析	232
11.5	pandas 绘图	233
11.5.1	Series 绘图	233
11.5.2	DataFrame 绘图	234
	本章习题	235
第 12 章	Python 数据可视化	237
12.1	Python 数据可视化简介	237
12.2	Matplotlib	238
12.2.1	Matplotlib 简介与安装	238
12.2.2	Matplotlib 基本绘图	238
12.2.3	Matplotlib 图像类型及其绘制	247
12.2.4	Matplotlib 绘图展示	255
12.3	seaborn	257
12.3.1	seaborn 简介与安装	257
12.3.2	seaborn 绘图	257
12.3.3	seaborn 绘图展示	268
12.4	PyEcharts	269
12.4.1	PyEcharts 简介与安装	269
12.4.2	PyEcharts 基础	269
12.4.3	PyEcharts 绘图	271
12.4.4	PyEcharts 其他图例展示	280
	本章习题	282
第 13 章	网络爬虫	283
13.1	预备知识	284

13.1.1	HTML 基础	284
13.1.2	Chrome 浏览器开发者工具	287
13.1.3	Network 面板的 XHR 页面	288
13.2	网络爬虫概述	289
13.2.1	网络爬虫系统的结构	289
13.2.2	网络爬虫的类型	289
13.2.3	网络爬虫的工作流程	290
13.3	Requests 爬虫库	291
13.3.1	requests 库的使用	291
13.3.2	requests 的异常处理	294
13.4	BeautifulSoup4 库的使用	295
13.4.1	BeautifulSoup4 库概述	296
13.4.2	BeautifulSoup 的对象类型	296
13.4.3	创建 BeautifulSoup 对象	297
13.4.4	BeautifulSoup 应用示例	299
13.4.5	BeautifulSoup:CSS	302
13.4.6	遍历文档树	303
13.5	案例 1:爬取中国天气网数据	304
13.5.1	HTML 页面分析	304
13.5.2	程序实现	306
13.6	案例 2:爬取国家统计局历年 GDP 数据	309
13.6.1	主题页面的结构特征	309
13.6.2	网页数据接口分析	309
13.6.3	JSON 格式解析	311
13.6.4	网络爬虫实现代码	312
13.7	案例 3:爬取古诗文网的名句	314
13.7.1	古诗文名句数据源的确定	314
13.7.2	网页结构分析	315
13.7.3	程序实现	316
	本章习题	320
第 14 章	Python 与机器学习	322
14.1	人工智能技术概述	323
14.1.1	人工智能发展简史	323
14.1.2	人工智能的研究领域和应用场景	325
14.1.3	人工智能、机器学习与深度学习	326
14.1.4	机器学习的分类	326
14.1.5	机器学习的流程	328