

python3

Quick Start and Actual Combat

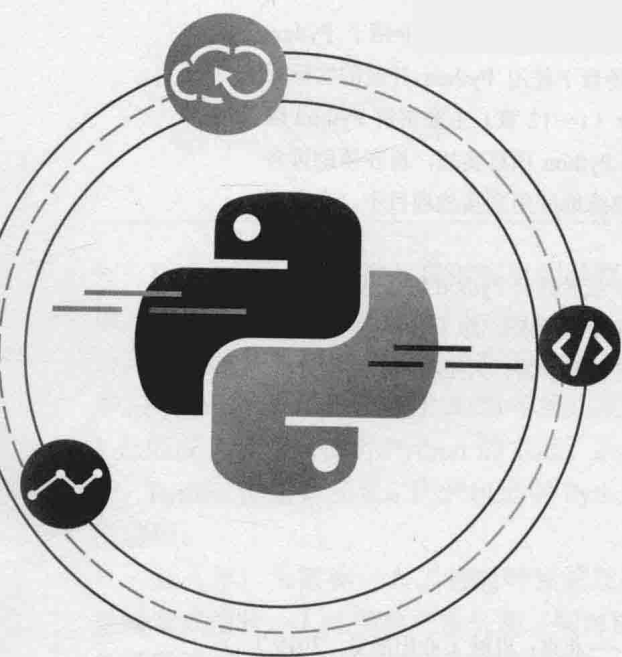
肖冠宇 杨捷 等编著

Python 3 快速入门与实战

系统——系统介绍Python3语法，夯实基础

全面——实例讲解Python数据可视化、爬虫开发、数据分析、机器学习等应用

实用——免费提供本书案例涉及的全部数据集、程序代码



python3

Quick Start and Actual Combat

肖冠宇 杨捷 等编著

快速入门

常州大学图书馆
藏书章

本书通过理论与实战相结合的方式，结合大量案例系统地介绍了 Python 编程涉及的知识点，详细介绍了多个应用场景下使用 Python 开发的实际项目。本书共 16 章，分为两部分，第一部分（1~12 章）主要讲解 Python 编程基础，第二部分（13~16 章）主要讲解 Python 项目实战，旨在帮助读者系统、快速地掌握 Python 语法，并能够熟练地应用到实战项目中。全书代码适用于 Python 3.6 以及更高版本。

本书读者对象为 Python 编程的初学者，或者具有 Python 编程基础想进一步学习 Python 的编程爱好者；具有其他编程语言基础，想了解和学习 Python 的相关技术人员；未来准备从事数据科学、机器学习、人工智能、数据分析、数据采集等方向研究和工作的读者。

图书在版编目（CIP）数据

Python 3 快速入门与实战 / 肖冠宇等编著. —北京：机械工业出版社，2019.7
ISBN 978-7-111-63405-8

I. ①P… II. ①肖… 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 165016 号

机械工业出版社（北京市百万庄大街 22 号 邮政编码 100037）

策划编辑：王 斌 责任编辑：王 斌

责任校对：张艳霞 责任印制：张 博

三河市宏达印刷有限公司印刷

2019 年 9 月第 1 版·第 1 次印刷

184mm×260mm·17.75 印张·2 插页·435 千字

0001—3000 册

标准书号：ISBN 978-7-111-63405-8

定价：79.00 元

电话服务

客服电话：010-88361066

010-88379833

010-68326294

封底无防伪标均为盗版

网络服务

机工官网：www.cmpbook.com

机工官博：weibo.com/cmp1952

金书网：www.golden-book.com

机工教育服务网：www.cmpedu.com

肖冠宇

大数据技术专家，曾就职于小米、人民网等互联网公司的大数据技术团队，拥有多年大数据系统研发经验，曾主导企业级大数据平台建设和多项大数据项目研发，擅长分布式系统架构、数据仓库产品设计与研发、实时计算、机器学习算法等。目前专注于大数据、人工智能领域的技术研究。

著有《企业大数据处理：Spark、Druid、Flume与Kafka应用实践》一书，参与编著《高可用可伸缩微服务架构：基于Dubbo、Spring Cloud和服务Mesh》一书。



作者公众号：DIMPLab

杨捷

英国高等统计学研究型硕士，在校期间荣获最高学术奖学金，英国The Face3D Research Consortium前成员。参与过多项数据分析科研项目，擅长数据统计建模分析与预测。曾担任英国大不列颠奥林匹克数学竞赛高级组指导老师，曾就职于爱丁堡知名咨询公司担任Consultant Assistant。2016年被邀请作为全球科技峰会苏格兰分会场演讲嘉宾。目前专注于机器学习、人工智能等领域的研究。

图书推荐



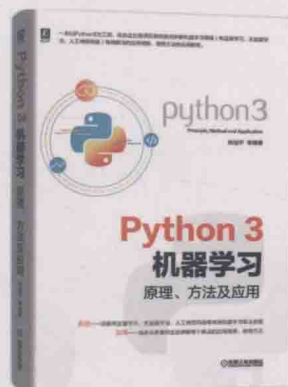
ISBN: 978-7-111-63405-8

定价: 79.00元



ISBN: 978-7-111-62593-3

定价: 79.00元



即将出版

前言

Python 是一种面向对象的解释型编程语言，具有语法简洁、免费开源、跨平台、丰富的第三方库等特点，由被誉为“Python 之父”的 Guido van Rossum 发明，于 1991 年发布第一个公开发行人版，发展到现在已经有近 30 年的历史。Python 从诞生到现在已经发布了几个大版本，Python 核心团队计划在 2020 年停止支持 Python 2，以后不再对其进行维护。目前，Python 3 已经成为学习和使用 Python 的主流。2019 年 7 月 TIOBE 发布的最新热门编程语言排行榜中，Python 位居第三名，因此也证明 Python 是目前最受欢迎的编程语言之一，并且已经被广泛使用。

近几年，大数据、人工智能等领域发展迅速，工程项目不断落地，相关领域的人才需求呈爆发式增长，人才供给严重失衡。同时随着国内人工智能教育的普及，中小學生、大学生、教师需要相关课程的教材和教辅资料。Python 是机器学习、人工智能、数据分析等领域使用最多的编程语言，所谓“工欲善其事，必先利其器”，Python 作为一把利器，可以快速地完成数据处理、数据分析、数据可视化、模型训练等。所以，掌握 Python 编程是从事机器学习、人工智能、数据分析等领域相关工作必备的技能。

为何写作本书

写作这本书的初衷是将自己工作中的编程和项目经验融合到 Python 理论知识中，让抽象、枯燥的编程语言的学习更加地生动、有趣。本书侧重实战，不仅系统介绍了 Python 编程涉及的知识点，同时也会教读者如何使用 Python 做实际的项目开发。希望通过理论与实战相结合的方式，让更多的编程爱好者快速、系统地掌握 Python。也希望读者通过 Python 的学习，掌握学习编程语言的方法，以后学习其他编程语言会更加的从容。

读者将学到什么

读者朋友可以通过本书的学习可以：

- 快速入门 Python 编程，系统掌握 Python 基础语法和进阶知识。
- 具备面向对象编程思想，掌握学习编程语言的方法，学习编程不再愁。
- 了解和学习多种基于 Python 的实战项目和应用场景，深入理解理论知识。
- 初步了解和掌握 Python 在数据采集、数据分析、机器学习等领域的应用。

如何阅读本书

本书共 16 章，分为两部分，第一部分（1~12 章）主要介绍 Python 编程基础，第二部分（13~16 章）主要介绍 Python 项目实战。通过理论与实战相结合的方式介绍 Python 语法和

应用，帮助读者系统、快速地掌握 Python 语法，并能够熟练地应用到实战项目中。全书代码适用于 Python 3.6 以及更高版本。

第 1 章 Python 概述，简要介绍 Python 语言的产生背景及目前的发展情况，Python 3.6、Anaconda 3（集成了 Python 3，可以不单独安装 Python 3.6）在 Windows 系统或者 Mac 系统中的安装配置，目前流行的一款开发工具 PyCharm 在 Windows 系统或者 Mac 系统中的安装配置。

第 2 章 Python 基础，主要介绍 Python 编程的基础语法，包括注释、标识符、变量、数据类型、输入输出函数、运算符、字符串、if 条件判断语句、while 循环与 for 循环。

第 3 章 容器，主要介绍 Python 中常用的四种数据结构，包括列表、元组、字典、集合。

第 4 章 函数，主要介绍函数的定义、参数、返回值、变量作用域、递归函数、匿名函数、闭包、装饰器。

第 5 章 包和模块，主要介绍 Python 中包的结构及作用、模块的用法、内置 `__name__` 变量的使用方法。

第 6 章 面向对象，主要介绍面向对象编程的思想、类和对象、构造方法、访问权限、继承。

第 7 章 异常处理，主要介绍异常捕获的方法、finally 语句的作用、使用 raise 抛出异常的方法。

第 8 章 日期和时间，主要介绍 Python 中内置的日期和时间相关的 time 和 datetime 两个模块。

第 9 章 文件操作，主要介绍读写文件的方法、文件管理、JSON 文件操作、CSV 文件操作。

第 10 章 正则表达式，主要介绍使用 re 模块编写正则表达式的方法、单字符匹配、数量表示、边界表示、转义字符、匹配分组、内置函数、贪婪与非贪婪模式。

第 11 章 Python 网络编程，主要介绍基于 Python 的网络编程基础、扩展库 urllib 和 requests 的用法等。

第 12 章 Python 常用扩展库，主要介绍 Python 中与数据科学相关的两个重要的扩展库：NumPy 和 Pandas 的用法。

第 13 章 Python 数据可视化实战，主要介绍 Matplotlib 绘图、Pandas 绘图、Seaborn 绘图等内容。

第 14 章 Python 爬虫开发实战，主要介绍爬虫开发流程、开发环境搭建、通过爬虫抓取电商网站商品信息的内容。

第 15 章 Python 数据分析实战，主要介绍数据分析基本概述、对房价进行数据分析等内容。

第 16 章 Python 机器学习实战，主要介绍机器学习基础、常用的机器学习库 scikit-learn、使用 k 近邻算法实现红酒质量等级预测等内容。

本书适用读者

对 Python 编程感兴趣的在校大学生，负责计算机相关专业教学的老师。

Python 编程的初学者，或者具有 Python 编程基础想进一步学习 Python 的编程爱好者。

具有其他编程语言基础，想了解和学习 Python 的相关技术人员。

未来准备从事数据科学、机器学习、人工智能、数据分析、数据采集等方向研究和工作的读者。

配套资源

本书配套的源代码可以通过微信公众号“DIMPLab”获取，也可通过扫描关注机械工业出版社计算机分社官方微信订阅号——IT 有得聊，回复 63405 即可获得本书配套资源下载链接。

读者反馈

由于笔者水平有限，书中难免会出现一些错误或者不准确的地方，敬请读者谅解，如果遇到任何问题或者技术交流都可以通过如下联系方式与笔者进行沟通。

- 微信公众号：“DIMPLab”后台留言反馈。
- 电子邮件：dimplab@163.com。

致谢 1

感谢美团点评的王碧琦最初提出写作本书的建议，同时感谢她在本书的写作过程中给予的支持和鼓励。

感谢机械工业出版社王斌老师，本书从 2018 年 6 月开始筹划，王老师从专业角度不断地给予帮助和指导，确定了整本书的写作框架。在写作过程中，王老师一直积极地推动本书的写作进度，使得我们可以顺利完成本书的写作。

感谢我的合作作者——杨捷，我们所有的写作都是在业余时间进行的，但是业余时间又常常被工作挤占，前期的写作进度一直很缓慢，在合作伙伴的鼓励 and 帮助下，克服了写作中的困难，顺利完成了本书。

感谢家人在工作和生活中对我的帮助和照顾。感谢父母，平时因工作原因很少回家看望，但他们一直在背后支持我、鼓励我。感谢我的妻子为家庭无怨无悔默默地付出，家人的陪伴与支持是我不断学习、努力奋斗的强大后盾！

肖冠宇

2019 年 8 月于北京

致谢 2

感谢与我一起合作的肖冠宇老师，感谢他带领我正式成为一名作者，有机会与大家分享我对 Python 的理解和想法。在创作和撰写的过程中，他给予了我很多帮助和鼓励。他总是很耐心地与我交流，分享他的宝贵经验，让我收获良多。他思维敏捷、经验丰富、技术过硬，像一个标杆，是我前进道路上的榜样。他从不吝啬夸赞的言语，常常给我很大的鼓励和支持，

让我信心满满地坚持下去，在此表示我最诚挚的敬意和感谢。

感谢机械工业出版社的王斌老师，他对本书的策划和编辑提出了很多宝贵的建议，并且一直推动本书的撰写进度。同时感谢出版社的其他工作人员，他们的辛勤付出才让本书得以顺利面世。

感谢父母的培育，从小到大他们都鼓励我独立思考，做自己想做的事情，为我提供足够的空间去追寻自己的梦想。每当我做出选择之后，他们都会全力支持，是我最坚强的后盾。他们给予我的爱、理解、关心、包容和支持是我前进的动力。

感谢我的朋友们，在本书的撰写过程中，他们给予了我很大的鼓励和支持。感谢他们让我的生活如此精彩有趣。

杨捷

2019年8月于北京

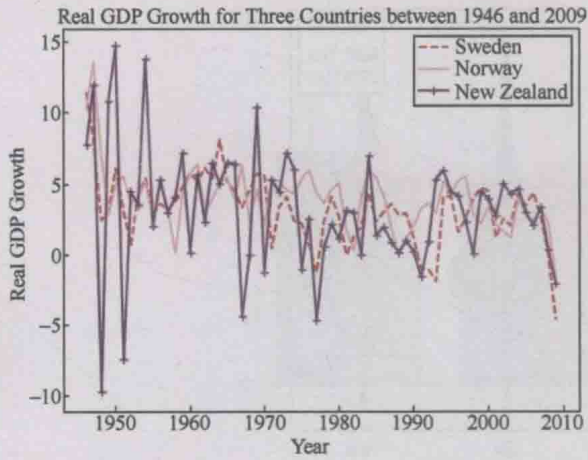


图 13-4 1946 年至 2009 年 3 个国家的 GDP 增长

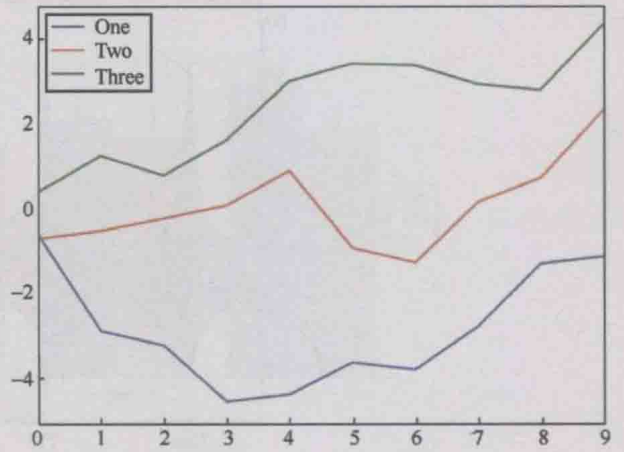


图 13-11 Pandas 绘制 DataFrame 折线图

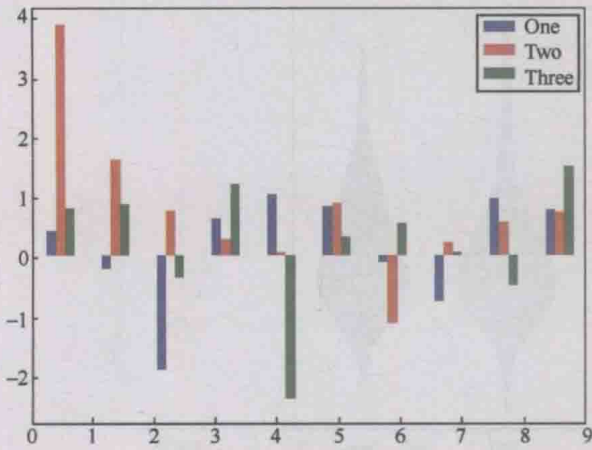


图 13-12 Pandas 绘制 DataFrame 柱状图

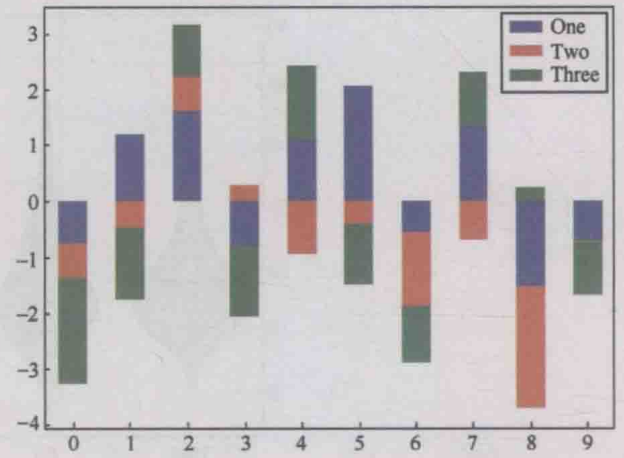


图 13-13 Pandas 绘制 DataFrame 堆叠柱状图

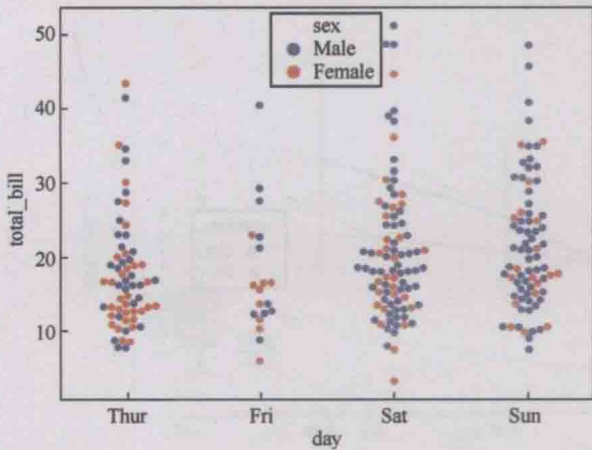


图 13-17 swarmplot 方法绘制双变量条形散点图

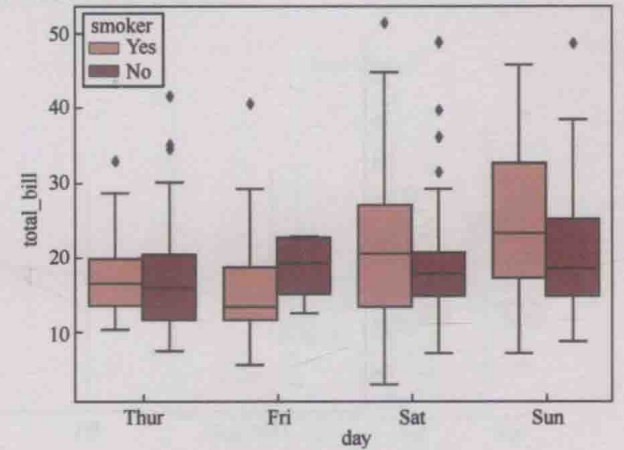


图 13-18 使用 boxplot 绘制箱线图

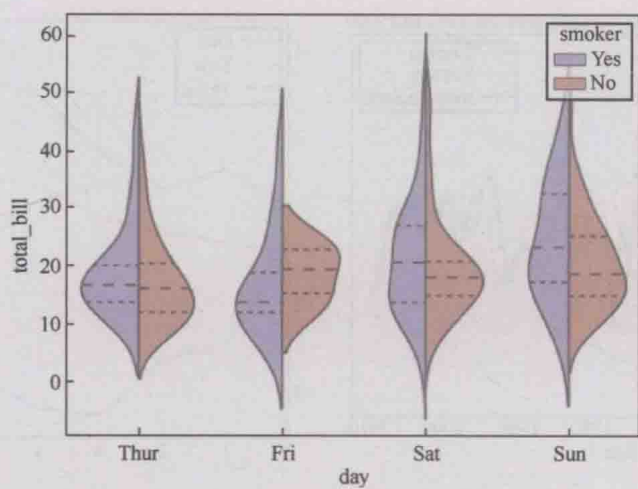


图 13-20 Seaborn 绘制分类琴形图

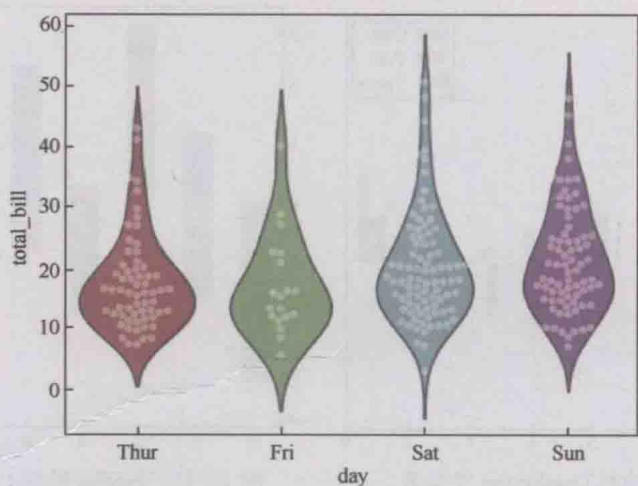


图 13-21 Seaborn 琴形图与散点图结合

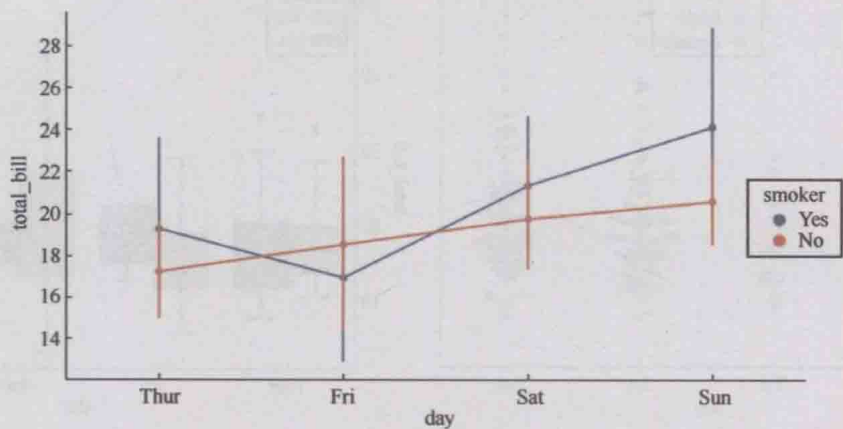


图 13-22 使用 factorplot 绘制点图

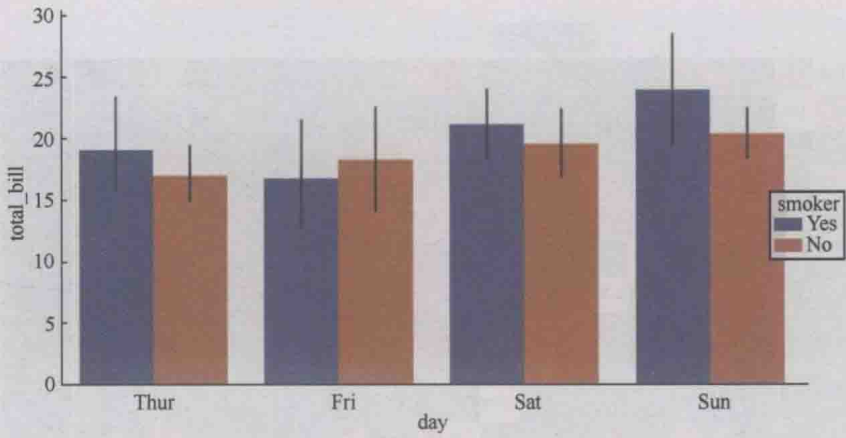


图 13-23 使用 factorplot 绘制柱形图

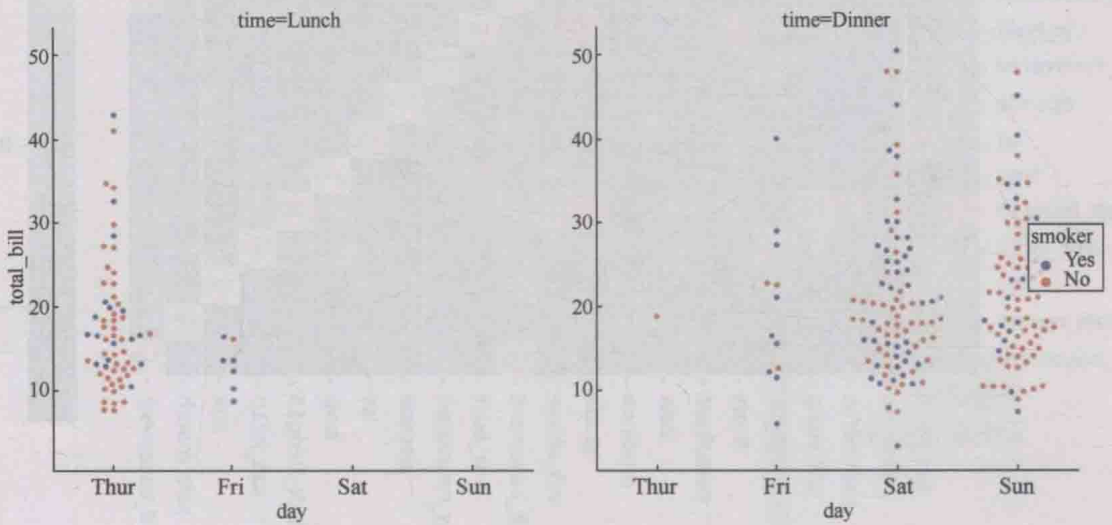


图 13-24 使用 factorplot 绘制条形散点图

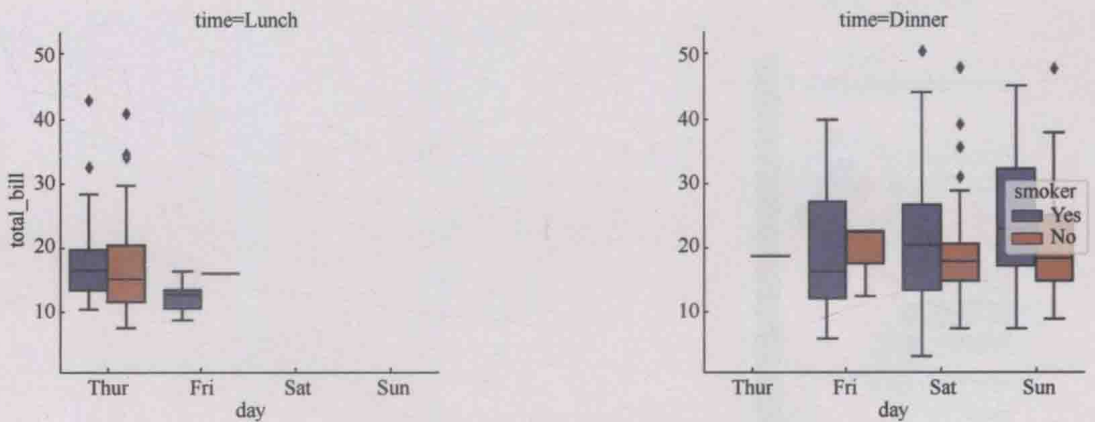


图 13-25 factorplot 绘制箱线图

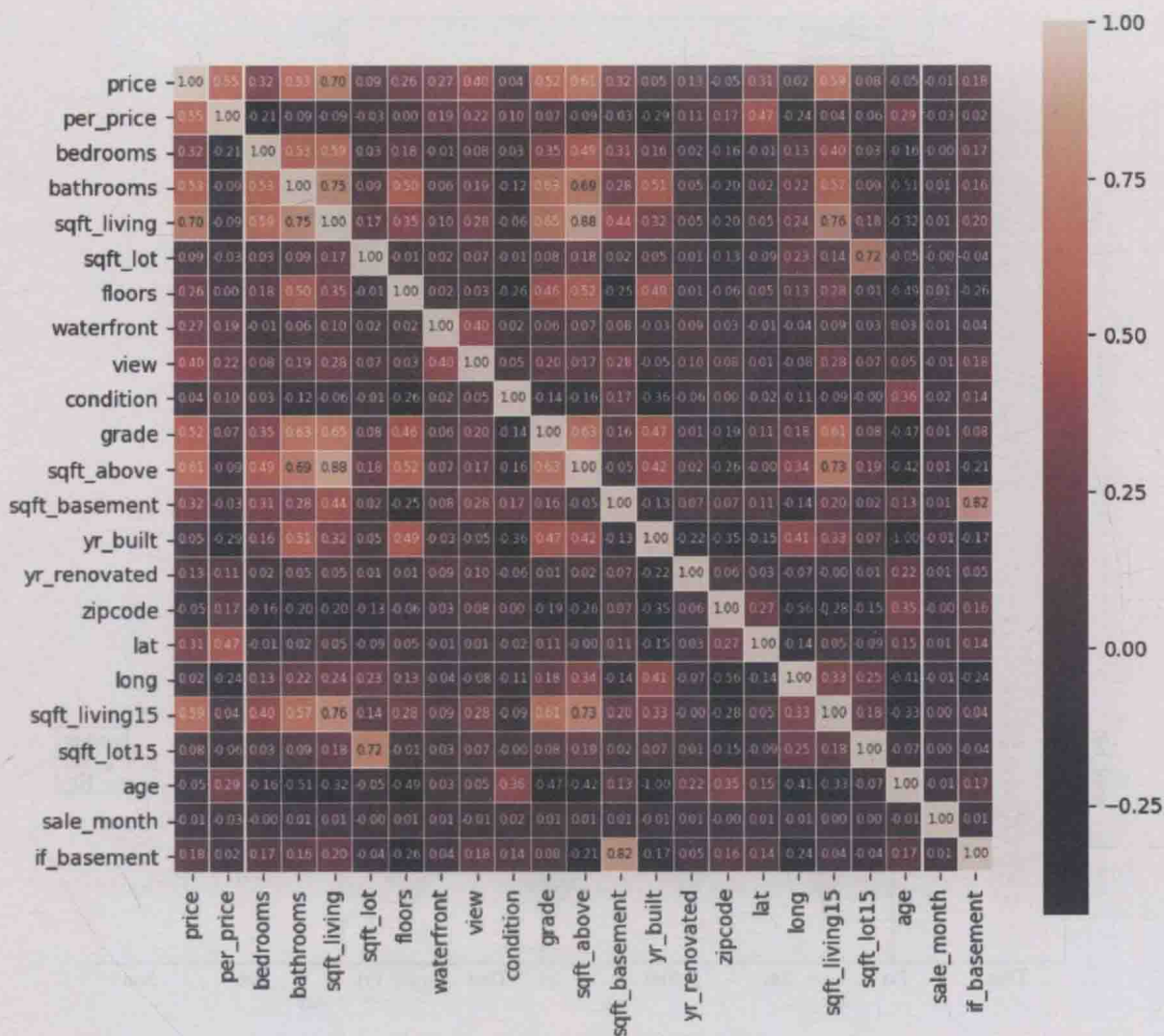


图 15-6 各变量间相关关系热力图

目录

前言

第 1 章 Python 概述	1
1.1 初识 Python	1
1.2 安装配置 Python 3 开发环境	2
1.2.1 Windows 系统下安装 Python 3	3
1.2.2 Mac 系统下安装 Python 3	5
1.3 安装配置 Anaconda	6
1.3.1 Windows 系统下安装 Anaconda	7
1.3.2 Mac 系统下安装 Anaconda	9
1.4 安装 PyCharm 开发工具	10
1.4.1 Windows 系统下安装 PyCharm	11
1.4.2 Mac 系统下安装 PyCharm	13
1.4.3 配置 PyCharm 开发环境	13
第 2 章 Python 基础	15
2.1 注释	15
2.1.1 单行注释	15
2.1.2 多行注释	16
2.2 关键字与标识符	16
2.2.1 关键字	16
2.2.2 标识符	17
2.3 变量	17
2.4 数据类型	18
2.4.1 数字类型	18
2.4.2 类型转换函数	19
2.4.3 布尔类型	20
2.5 输入 (input) 与输出 (print)	21
2.5.1 输入函数 (input)	21
2.5.2 输出函数 (print)	21
2.6 运算符	23
2.6.1 算术运算符	23

2.6.2	比较运算符	24
2.6.3	赋值运算符	24
2.6.4	逻辑运算符	25
2.6.5	运算符优先级	25
2.7	字符串	26
2.7.1	字符串定义	26
2.7.2	字符串格式化	26
2.7.3	字符串内置方法	28
2.8	if 条件判断	32
2.8.1	语法格式	32
2.8.2	条件判断与逻辑运算符	33
2.8.3	if 嵌套	34
2.9	while 循环	35
2.9.1	语法格式	35
2.9.2	while 循环嵌套	36
2.9.3	break 跳出整个 while 循环	37
2.9.4	continue 跳出当次 while 循环	38
2.10	for 循环	38
2.10.1	语法格式	39
2.10.2	break 跳出整个 for 循环	40
2.10.3	continue 跳出当次 for 循环	40
第 3 章	容器	41
3.1	列表	41
3.1.1	列表的定义	41
3.1.2	查询列表中元素	42
3.1.3	嵌套列表	44
3.1.4	向列表中添加元素	44
3.1.5	修改列表中元素值	46
3.1.6	删除列表中元素	46
3.1.7	列表切片	47
3.1.8	列表元素排序	48
3.2	元组	48
3.2.1	元组的定义	48
3.2.2	查询元组中的元素	49
3.3	字典	50
3.3.1	字典的定义	50

3.3.2	查询字典中键值对	50
3.3.3	向字典中添加键值对	51
3.3.4	修改字典中键的值	51
3.3.5	删除字典中的键值对	52
3.3.6	循环遍历字典	52
3.4	集合	54
3.4.1	集合的定义	55
3.4.2	成员运算符在集合中的应用	56
3.4.3	向集合中添加元素	56
3.4.4	删除集合中的元素	58
3.4.5	集合常用操作	59
第4章	函数	61
4.1	函数定义与调用	61
4.2	函数参数	62
4.2.1	带参函数	62
4.2.2	缺省参数	63
4.2.3	命名参数	64
4.2.4	不定长参数	64
4.3	函数返回值	67
4.4	变量作用域	68
4.4.1	局部变量	68
4.4.2	全局变量	69
4.5	递归函数	71
4.6	匿名函数	72
4.7	闭包	73
4.8	装饰器	76
4.8.1	初识装饰器	77
4.8.2	装饰器进阶	78
第5章	包和模块	82
5.1	包	82
5.2	模块	83
5.3	<code>__init__.py</code> 模块	85
5.4	<code>__name__</code> 变量	85
第6章	面向对象	89
6.1	面向对象编程	89
6.2	类和对象	89

6.2.1	类	90
6.2.2	对象	91
6.3	<code>__init__</code> 构造方法	93
6.4	访问权限	95
6.5	继承	97
6.5.1	单继承	97
6.5.2	<code>super</code> 函数	99
6.5.3	重写	101
6.5.4	多继承	102
第 7 章	异常处理	105
7.1	捕获异常	105
7.2	捕获多个异常	106
7.3	捕获全部异常	107
7.4	异常中的 <code>finally</code> 语句	108
7.5	异常传递	109
7.6	<code>raise</code> 抛出异常	110
第 8 章	日期和时间	112
8.1	<code>time</code> 模块	112
8.2	<code>datetime</code> 模块	116
第 9 章	文件操作	119
9.1	读写文件	119
9.1.1	打开文件	119
9.1.2	写文件	120
9.1.3	读文件	123
9.2	文件管理	124
9.3	JSON 文件操作	126
9.4	CSV 文件操作	129
9.4.1	写入 CSV 文件	129
9.4.2	读取 CSV 文件	130
第 10 章	正则表达式	131
10.1	<code>re</code> 模块	131
10.2	单字符匹配	132
10.3	数量表示	134
10.4	边界表示	136
10.5	转义字符	138
10.6	匹配分组	140