

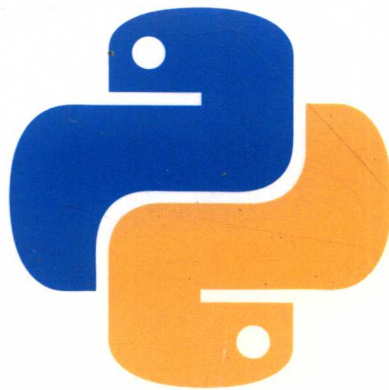
北京课工场教育科技有限公司 **出品**



新技术技能人才培养系列教程

大数据核心技术系列

Python 开发基础



戴歆 罗玉军 / 主编

杨颖 侯勇 王奇志 / 副主编



视频资源



案例素材



交流社区



学习 App

中国工信出版集团

人民邮电出版社
POSTS & TELECOM PRESS

非外借

北京课工场教育科技有限公司 出品



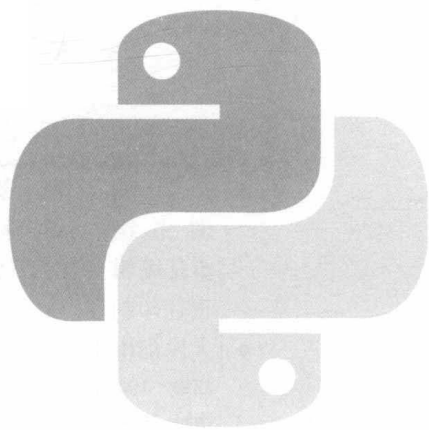
新技术技能人才培养系列教程

大数据核心技术系列

Python 开发基础

戴歆 罗玉军 / 主编

杨颖 侯勇 王奇志 / 副主编



人民邮电出版社

北京

图书在版编目 (C I P) 数据

Python开发基础 / 戴歆, 罗玉军主编. — 北京 :
人民邮电出版社, 2018. 12
新技术技能人才培养系列教程
ISBN 978-7-115-49452-8

I. ①P… II. ①戴… ②罗… III. ①软件工具—程序
设计—教材 IV. ①TP311.561

中国版本图书馆CIP数据核字(2018)第234032号

内 容 提 要

Python 语言是当前最活跃的开发语言之一, 在数据科学领域、网络爬虫领域、Web 开发领域、服务器自动化运维及游戏领域都有着非常广泛的应用。尤其是在数据科学领域, 越来越多的数据科学家开始将 Python 语言作为主要的工具。

本书以 Windows 操作系统为平台, 系统讲解 Python 3 的基础知识。全书共 10 章, 首先介绍了 Python 语言的基础入门, 开发环境搭建, 必备的基础语法, 如变量与数据类型、流程控制语句、常用数据结构、函数与模块、程序调试方法等; 然后介绍了面向对象的思想, 包括封装、继承、多态等, 以及如何使用 Python 完成文件读写功能。这将是一个由浅入深的系统学习过程。

为提升学习效果, 书中结合实际应用提供了大量案例进行说明和训练, 并配以完善的学习资料和支持服务, 包括教学 PPT、案例素材下载、源码下载、学习交流社区、讨论组等, 为读者带来全方位的学习体验。

本书可作为高职高专院校和应用型本科院校的计算机相关专业的教材和参考书。

-
- ◆ 主 编 戴 歆 罗玉军
 - 副 主 编 杨 颖 侯 勇 王奇志
 - 责任编辑 祝智敏
 - 责任印制 马振武
 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
涿州市京南印刷厂印刷
 - ◆ 开本: 787×1092 1/16
印张: 12.5 2018 年 12 月第 1 版
字数: 262 千字 2018 年 12 月河北第 1 次印刷

定价: 39.80 元

读者服务热线: (010) 81055256 印装质量热线: (010) 81055316

反盗版热线: (010) 81055315

广告经营许可证: 京东工商广登字 20170147 号

大数据核心技术系列

编委会

主任：肖睿

副主任：潘贞玉 韩露

委员：李娜 孙苹 张惠军 杨欢

庞国广 王丙晨 刘晶晶 曹紫涵

刘洋 崔建瑞 刘尧 饶毅彬

马志成 李红 尚泽中 杜静华

董海 孙正哲 周嵘 冯娜娜

序 言

丛书设计

大数据已经悄无声息地改变了我们的生活和工作方式，精准广告投放、实时路况拥堵预测已很普遍，在一些领域，人工智能比我们更加聪明、高效，未来的个性化医疗、教育将会真正实现，大数据迎来前所未有的机遇。Google 公司 2003 年开始陆续发表的关于 GFS、MapReduce 和 BigTable 的三篇技术论文，成为大数据发展的重要基石。十几年来大数据技术从概念走向应用，形成了以 Hadoop 为代表的一整套大数据技术。时至今日，大数据技术仍在快速发展，基础框架、分析技术和应用系统都在不断演变和完善，并不断地涌现出大量新技术，成为大数据采集、存储、处理、分析、可视化呈现的有效手段。企业需要利用大数据更加贴近用户、加强业务中的薄弱环节、规范生产架构和策略。对数家企业的调查显示，大数据工程师应该掌握的技能包括：Hadoop、HDFS、MapReduce、Hive、HBase、ZooKeeper、YARN、Sqoop、Spark、Spark Streaming、Scala、Kafka、Confluent、Flume、Redis、ETL、Flink/Streaming、Linux、Shell、Python、Java、MySQL、MongoDB、NoSQL、Cassandra、Spark MLib、Pandas、Numpy、Oozie、ElasticSearch、Storm 等，作为一名大数据领域的初学者，在短时间内很难系统地掌握以上全部技能点。“大数据核心技术系列”丛书根据企业人才实际需求，参考以往学习难度曲线，选取“Hadoop+Spark+Python”技术集作为核心学习路径，旨在为读者提供一站式、实战型大数据开发学习指导，帮助读者踏上由开发入门到实战的大数据开发之旅！

“大数据核心技术系列”以 Hadoop、Spark、Python 三个通用大数据平台为核心，根据它们各自不同的特点，解决大数据中离线批处理和实时计算两种主要场景的应用。以 Hadoop 为核心完成大数据分布式存储与离线计算；使用 Hadoop 生态圈中的日志收集、任务调度、消息队列、数据仓库、可视化 UI 等子系统完成大数据应用系统架构设计；以 Spark Streaming、Storm 替换 Hadoop 的 MapReduce 以实现大数据的实时计算；使用 Python 完成数据采集与分析；使用 Scala 实现交互式查询分析与 Spark 应用开发。书中结合大量项目案例完成大数据处理业务场景的实战。

在夯实大数据领域技术基础的前提下，“大数据核心技术系列”丛书结合当下 Python 语言在数据科学领域的活跃表现以及占有量日益扩大的现状，加强了对 Python 语言基础、Scrapy 爬虫框架、Python 数据分析与展示等相关技术的讲解，为读者将来在大数据科学领域的进一步提升打下坚实的基础。

丛书特点

1. 以企业需求为设计导向

满足企业对人才的技术需求是本系列丛书的核心设计原则，课工场大数据开发教研团队通过对数百位 BAT 一线技术专家进行访谈、对上千家企业人力资源情况进行调研、对上万个企业招聘岗位进行需求分析，实现对技术的准确定位，达到课程与企业需求的高契合度。

2. 以任务驱动为讲解方式

丛书中的知识点和技能点均由任务驱动，读者在学习知识时不仅可以知其然，而且可以知其所以然，帮助读者融会贯通、举一反三。

3. 以实战项目来提升技术

本丛书均设置项目实战环节，以综合运用书中的知识点帮助读者提升项目开发能力。每个实战项目都设有相应的项目思路指导、重难点讲解、实现步骤总结和知识点梳理。

4. 以“互联网+”实现终身学习

本丛书可配合课工场 App 进行二维码扫描，来观看配套视频的理论讲解和案例操作，同时课工场在线开辟教材配套版块，提供案例代码及案例素材下载。此外，课工场还为读者提供了体系化的学习路径、丰富的在线学习资源和活跃的学习社区，方便读者随时学习。

读者对象

1. 大中专院校的学生
2. 编程爱好者
3. 初中级程序开发人员
4. 相关培训机构的老师和学员

读者服务

学习本丛书过程中如遇到疑难问题，读者可以访问课工场在线，也可以发送邮件到 ke@kgc.cn，我们的客服专员将竭诚为您服务。

感谢您阅读本丛书，希望本丛书能成为您大数据开发之旅的好伙伴！

“大数据核心技术系列”丛书编委会

前言

欢迎进入 Python 的世界，本书将带你领略 Python 语言的魅力，感受 Python 语言的简洁，体验它无穷的魅力。各章主要内容如下。

第 1 章：介绍 Python 语言的特点、Python 的版本差异、搭建 Python 开发环境的方法、创建 Python 项目工程的方法、使用 Python 添加注释的方法，掌握这些是学好 Python 语言的基础。

第 2 章：介绍 Python 的基础语法，包括变量、数据类型和运算符、类型转化方法以及操作字符串的方法。只有掌握了各种类型的操作方法，才能灵活地完成数据的处理。学习完本章内容后，读者将能够编写有意义的小程序。

第 3 章：流程控制是编程的基础，本章详细讲解了 Python 中的两种流程控制结构，即循环结构和选择结构。主要介绍 if-elif 语句、多分支 if 语句、while 循环、for 循环。学习完本章内容，读者可开发出能够灵活实现业务控制的 Python 程序。

第 4 章：列表（list）、元组（tuple）、字典（dict）和集合（set）是 Python 中重要的数据存储结构，本章介绍这些数据结构的特点与使用方法。学习完本章内容，读者将掌握不同数据结构的特点、用法与使用场景，能够开发出基于不同数据存取结构的 Python 程序。

第 5 章：函数和模块能够提高代码的复用性，本章详细地讲解了函数和模块的使用方法。学习完本章内容，读者可以使用函数和模块实现代码的封装，提高代码的可读性和可复用性，掌握导入 Python 内置模块和第三方模块，调用函数提高开发效率的能力。

第 6 章：综合前面章节所学知识完成一个实际的项目——在线投票系统。实现添加投票候选人、删除候选人、为候选人投票、按序号投票、删除投票、清空投票、投票统计、退出投票等功能。

第 7 章：讲解如何使用 PyCharm 的 Debug 功能实现对代码的断点调试以及介绍 Python 的异常处理机制。学习完本章内容，读者将掌握开发中重要的断点调试技能，是程序员完成大型复杂项目必不可少的能力。理解 Python 的异常处理机制，将能够提高程序的健壮性、安全性和可维护性。

第 8 章：讲解 Python 面向对象的核心内容，包括对象和封装、继承、多态等，重点培养读者使用面向对象思想进行程序设计的能力。

封装、继承和多态是面向对象的三大特性。封装类的属性可以隐藏类的实现细节，限制不合理操作。继承是软件可重用性的一种表现，新类可以在不增加自身代码的情况下，通过从现有的类中继承其属性和方法来充实自身内容。多态在面向对象编程中无处不在，是解决编程中实际问题的一大利器。

第 9 章：介绍常用的文件类型和特点、读写 txt、csv 文件的方法、导入模块操作电脑中的文件以及 json 格式的特点和解析方法。学习完本章内容后，能够将程序数据保存到文件中或从文件中读取数据，使用 json 格式读写数据可以提高数据的可维护性。

第 10 章：通过项目实训，利用面向对象编程思想来升级在线投票系统，巩固使用面向对象编程开发系统的能力。

学习程序设计语言，要多动手练习，从而深入理解每个知识点，提高编程熟练度，培养分析问题和解决问题的能力，不断积累开发经验。同时，学习中要通过交流消除学习疑惑，分享学习经验，取长补短，共同进步。

本书由课工场大数据开发教研团队组织编写，参与编写的还有戴歆、罗玉军、杨颖、侯勇、王奇志、谢姐姐等院校老师。尽管编者在写作过程中力求准确、完善，但书中不妥或错误之处仍在所难免，殷切希望广大读者批评指正！

智慧教材使用方法

由课工场“大数据、云计算、全栈开发、互联网 UI 设计、互联网营销”等教研团队编写的系列教材，配合课工场 App 及在线平台的技术内容更新快、教学内容丰富、教学服务反馈及时等特点，结合二维码、在线社区、教材平台等多种信息化资源获取方式，形成独特的“互联网+”形态——智慧教材。

智慧教材为读者提供专业的学习路径规划和引导，读者还可体验在线视频学习指导，按如下步骤操作可以获取案例代码、作业素材及答案、项目源码、技术文档等教材配套资源。

1. 下载并安装课工场 App。

(1) 方式一：访问网址 www.ekgc.cn/app，根据手机系统选择对应课工场 App 安装，如图 1 所示。



图1 课工场App

(2) 方式二：在手机应用商店中搜索“课工场”，下载并安装对应 App，如图 2、图 3 所示。

2. 登录课工场 App，注册个人账号，使用课工场 App 扫描书中二维码，获取教材配套资源，依照如图 4 至图 6 所示的步骤操作即可。



图2 iPhone版手机应用下载



图3 Android版手机应用下载

Java 面向对象程序开发及实战

3. 变量

前面讲解了Java中的常量，与常量对应的就是变量。变量是在程序运行中其值可以改变的量，它是Java程序的一个基本存储单元。

变量的基本格式与常量有所不同。

变量的语法格式如下。

[访问修饰符] 变量类型 变量名 [= 初始值];

- “变量类型”可从数据类型中选择。
- “变量名”是定义的名称变量，要遵循标识符命名规则。
- 中括号中的内容为初始值，是可选项。

示例 4

使用变量存储数据，实现个人简历信息的输出。

分析如下。

- (1) 将常量赋给变量后即可使用。
- (2) 变量必须先定义后使用。



图4 定位教材二维码



图5 使用课工场App“扫一扫”扫描二维码



图6 使用课工场App免费观看教材配套视频

3. 获取专属的定制化扩展资源。

(1) 普通读者请访问 <http://www.ekgc.cn/bbs> 的“教材专区”版块，获取教材所需开发工具、教材中示例素材及代码、上机练习素材及源码、作业素材及参考答案、项目素材及参考答案等资源（注：图7所示网站会根据需求有所改版，下图仅供参考）。



图7 从社区获取教材资源

(2) 高校老师请添加高校服务 QQ 群：1934786863（如图8所示），获取教材所需开发工具、教材中示例素材及代码、上机练习素材及源码、作业素材及参考答案、项目素材及参考答案、教材配套及扩展 PPT、PPT 配套素材及代码、教材配套线上视频等资源。



图8 高校服务QQ群

目 录

第1章 初识Python	1
任务1 搭建Python开发环境	2
1.1.1 Python语言的特点	2
1.1.2 Python语言的版本	3
1.1.3 Python的应用场景	4
1.1.4 Python开发环境安装和配置	5
1.1.5 技能实训	9
任务2 在控制台输出“Hello Python”	9
1.2.1 Python程序初体验	9
1.2.2 PyCharm集成开发环境	11
1.2.3 技能实训	20
本章总结	20
本章作业	21
第2章 变量与数据类型	23
任务1 输出学生的信息	24
2.1.1 变量和数据类型	24
2.1.2 运算符	30
2.1.3 技能实训	34
任务2 实现文本处理	34
2.2.1 字符串拼接	35
2.2.2 常用操作字符串的方法	37
2.2.3 技能实训	40
本章总结	40
本章作业	40
第3章 流程控制语句	41
任务1 根据销售业绩输出绩效提成百分比	42
3.1.1 初识流程控制结构	42
3.1.2 Python语句块规范	43
3.1.3 选择结构	43
3.1.4 技能实训	50

任务2 计算销售人员近3个月的平均销售金额	50
3.2.1 循环结构	50
3.2.2 循环跳转语句	57
3.2.3 技能实训	59
本章总结	60
本章作业	60
第4章 常用数据结构	61
任务1 员工薪酬数据管理	62
4.1.1 常用数据结构	62
4.1.2 列表 (list)	63
4.1.3 二维列表	67
4.1.4 元组 (tuple)	68
4.1.5 技能实训	68
任务2 使用工号管理员工信息	70
4.2.1 字典 (dict)	70
4.2.2 集合 (set)	74
4.2.3 技能实训	77
本章总结	77
本章作业	78
第5章 函数与模块	81
任务1 自定义函数计算景区指定条件下的月平均访问量	82
5.1.1 无参函数	82
5.1.2 有参函数	83
5.1.3 函数的返回值	89
5.1.4 技能实训	93
任务2 使用内置模块随机生成双色球中奖号码	95
5.2.1 模块	95
5.2.2 常用模块	99
5.2.3 技能实训	101
本章总结	101
本章作业	101
第6章 项目实训——在线投票系统	103
6.1 项目需求	104
6.2 难点分析	106
6.3 项目实现思路	109
本章总结	114

本章作业	114
第7章 程序调试方法	115
任务1 计算学生平均成绩	116
7.1.1 异常概述	116
7.1.2 异常处理	118
7.1.3 技能实训	124
任务2 对学生成绩进行排名	125
7.2.1 理解崩溃信息	125
7.2.2 利用PyCharm进行debug	126
本章总结	129
本章作业	129
第8章 面向对象编程	131
任务1 使用面向对象的编程思想定义银行员工类	132
8.1.1 面向对象的编程思想	132
8.1.2 类和对象	133
8.1.3 技能实训	142
任务2 使用继承根据职位创建银行员工类的子类	144
8.2.1 继承	144
8.2.2 多态	150
8.2.3 技能实训	152
本章总结	153
本章作业	153
第9章 文件读写	155
任务1 实现新闻编辑功能	156
9.1.1 认识最常用的文件类型	156
9.1.2 txt文件读写操作	157
9.1.3 with语句	162
9.1.4 技能实训	163
任务2 将学生数据整理为csv和json格式	164
9.2.1 csv与json读写	164
9.2.2 路径和文件的操作	168
9.2.3 技能实训	170
本章总结	171
本章作业	171

第10章 项目实训——升级在线投票系统	173
10.1 项目需求	174
10.2 难点分析	176
10.3 项目实现思路	179
本章总结	184
本章作业	184

初识 Python

技能目标

- 了解 Python 语言的特点
- 了解 Python 的版本差异
- 掌握搭建 Python 开发环境的方法
- 掌握使用 PyCharm IDE 编辑 Python 代码的方法
- 掌握使用 Python 添加注释的方法

本章任务

任务 1: 搭建 Python 开发环境

任务 2: 在控制台输出 “Hello Python”



第1章 初识Python

任务1：搭建Python开发环境

1.1.1 Python语言的特点

1.1.2 Python语言的版本

1.1.3 Python的应用场景

1.1.4 Python开发环境安装和配置

1.1.5 技能实训

任务2：在控制台输出“Hello Python”

1.2.1 Python程序初体验

1.2.2 PyCharm集成开发环境

1.2.3 技能实训

本书将通过大量的应用实例来全面介绍 Python，每一章会以任务的形式展开，每章介绍的知识都能完成一个真实的任务，做到学以致用。

通过本章的学习，读者不仅能对 Python 有一个初步的认识，同时也能写下第一行 Python 代码，完成在控制台输出“Hello Python”的任务。

任务 1 搭建 Python 开发环境

【任务描述】

介绍 Python 语言的特点、版本、应用场景并通过 Anaconda 搭建 Python 开发环境。

【关键步骤】

- (1) 了解 Python 语言的特点。
- (2) 了解 Python 的版本差异。
- (3) 了解 Python 的应用场景。
- (4) Python 开发环境的安装和配置。

1.1.1 Python 语言的特点

Python 语言具有以下显著的特点。

1. 简单易学

Python 是一种代表简单思想的语言。Python 的关键字少、结构简单、语法清晰，使学习者可以在相对较短的时间内轻松上手。

2. 易于阅读

Python 代码定义得非常清晰，它没有使用其他语言通常用来访问变量、定义代码块和进行模式匹配的命令式符号，而是采用强制缩进的编码方式，去除了“{}”等语法符号，从而看起来十分规范和优雅，具有极佳的可读性。

3. 免费、开源

Python 是 FLOSS（自由/开放源码软件）之一。使用 Python 是免费的，开发者可以自由地发布这个软件的副本，阅读源代码，甚至对它做改动。