

普通高等教育通识类课程新形态教材

Python 办公自动化——玩转 Excel

郝春吉 刘智杨 周永福 黄 詮 编著



中国水利水电出版社
www.waterpub.com.cn

· 北京 ·

内 容 提 要

本教材为 Python 数据分析立体教材，以任务为导向，结合大量数据分析案例及教学经验，深入浅出地介绍 Python 语言在处理 Excel 格式数据时的重要方法及内容，详尽地介绍了 Python 第三方库 xlrd 库、xlwt 库、xlwings 库、pandas 库、matplotlib 库、numpy 库和 pyplot 库的基本操作方法。为读者明确如何利用所学知识解决问题，设置了多个案例，每个案例都配有程序运行源代码，读者可通过二维码链接到网站上下载实际代码，查看运行结果，同时配以作者精心制作的视频讲解，轻松愉快地学会用 Python 玩转各领域数据，能够真正地学以致用。

本教材适用于高校计算机公共基础程序设计课程和数据分析课程，也可供需要处理大量 Excel 数据、工作重复度较高的工作人员使用。

策划编辑：石永峰 责任编辑：周春元 封面设计：李 佳

书 名	普通高等教育通识类课程新形态教材 Python 办公自动化——玩转 Excel Python BANGONG ZHIDONGHUA——WANZHUAN Excel
作 者	郝春吉 刘智杨 周永福 黄 诠 编著
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	三河市德贤弘印务有限公司
规 格	184mm×260mm 16 开本 12.75 印张 318 千字
版 次	2022 年 1 月第 1 版 2022 年 1 月第 1 次印刷
印 数	0001—3000 册
定 价	39.00 元

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

前 言

大学计算机公共基础课已经开设了二十余年了，不可否认，其为计算机知识的普及做出了不可磨灭的贡献。然而时代是发展的，唯一不变的就是变化。随着时间的推移，原有的大学计算机公共基础课程内容已经不能适应当前社会发展的需要，亟需更新内容跟上时代发展的步伐。基于此，编写了这本教材，用以开展并推广新一轮的大学计算机基础课教学改革。

大数据时代已经到来，无论我们接受与否都必须面对这样的一个事实，数据处理是最直接的体现。在数据量巨大的情况下，原有的手工操作已经不能满足人们日常生活和工作所需，办公自动化迎来了新一轮的革新。本书以 Python 语言代替之前 Excel 中的手工操作，实现一种全新的“办公自动化”方式。使用 Python 语言，只需要几行代码就可以轻松解决问题，特别是当工作重复度很高的时候，只要略微改动代码即可，大大地节省了时间，提高了工作效率。

Python 是一种跨平台的计算机程序设计语言，是一个高层次的结合了解释性、编译性、互动性和面向对象的脚本语言。其最初被设计用于编写自动化脚本(shell)，随着版本的不断更新和语言新功能的添加，现更多地被用于独立的、大型项目的开发。

在实际应用中，利用 Python 对 Excel 电子表格进行处理与可视化表格非常快速。因此，如何将 Excel 与 Python 语言相结合进行数据的处理，是数据分析从业者需要掌握的重点内容。

虽然 Excel 也是进行数据处理的专业软件，使用 Excel 制作报表进行数据管理也是最常见的数据管理方式，但其在进行自动化操作等方面不如 Python 灵活。利用 Python 可以很容易地读取、计算和编辑 Excel 文档中的数据，提高数据分析工作的效率。

本书主要讲述如何利用 Python 处理 Excel 文件，进而进行数据分析以及可视化等操作。

Python 与 Excel 之所以能够结合应用，主要还是因为其各自的特点：

- Python 语言编写程序非常方便，统一语言带来记录方法的规范统一。当需要修改或者复制重要功能时，只需要调整好设定的参数就可以。
- Excel 电子表格：繁琐的每一步都来自鼠标点击，中间有一处错误，很多步骤都需要重新调整，浪费大量时间，完全体现不出自动化的功效。而用来提升 Excel 电子表格功能的 VBA（Visual Basic for Applications）和宏，又过于复杂，不如 Python 简单和容易使用。

Python 语言中，仅仅对于可视化基本库 matplotlib 而言，官方 Gallery 就有 26 个大类 527 个样式，无论在数量上还是在质量上都不是其他软件可比拟的，其他软件只能用软件自带的工具，拓展模块复杂耗时。无论是对从事财务、审计、行政、HR（Human Resource，人力资

源)、运营、管理、销售等工作的工作人员，还是在校大学生而言，Python 程序设计语言都是他们的工作利器，其作用远远不止是提高工作效率，同时也是工作生活中不可或缺的一项技能。

通过学习本书，希望读者能够实现自动化办公，拒绝重复工作；可以处理各种海量数据，生成可视化图表；提高效率，给工作和学习带来更多的惊喜。

由于作者水平有限，加之编写时间仓促、内容较多，书中难免存在不当之处，恳请广大读者批评指正。

编者
2021 年 9 月

目 录

前言

第 1 章 Python 基础与主要第三方库..... 1	2.2.9 修改源工作表中的数据方式 (修改标题)21
1.1 Python 语言介绍1	2.2.10 在源工作簿中增加新的工作表21
1.2 快速搭建 Python 开发环境1	2.2.11 复制工作簿22
1.2.1 Python 软件的下载与安装 (Windows 系统)1	2.2.12 激活活动表格22
1.2.2 Python 软件的使用4	2.2.13 获取工作表中有效范围内的 有效数据23
1.3 Python 语言的核心5	2.3 格式控制24
1.3.1 循环语句5	2.3.1 设置工作表的行高和列宽24
1.3.2 条件语句9	2.3.2 设置工作表文字格式25
1.3.3 列表11	2.3.3 字体 (Font) 属性设置26
1.3.4 字典12	2.3.4 边框 (Borders) 属性设置27
1.4 常用第三方库简介12	2.3.5 对齐 (Allignment) 属性设置28
1.5 本章总结13	2.3.6 背景 (Pattern) 属性设置30
第 2 章 Python 办公自动化之 xlrd、xlwt 和 xlwings 库.....14	2.3.7 设置字体颜色31
2.1 用 Python 创建及读取 Excel 文件14	2.3.8 设置表格边框31
2.1.1 用 Python 创建 Excel 文件14	2.3.9 设置行高, 列宽, 内容位置33
2.1.2 读取 Excel 文件14	2.3.10 合并单元格34
2.1.3 读取 Excel 工作表 (以工作表名称打开)15	2.3.11 拆分单元格35
2.1.4 读取 Excel 工作表 (以工作表序号打开)15	2.3.12 设置表格背景颜色36
2.2 写入数据及数据计算16	2.3.13 获取指定范围背景颜色36
2.2.1 写入数据16	2.3.14 清除表格背景颜色37
2.2.2 获取工作表总行数 (rows)16	2.3.15 清除指定范围的数据内容38
2.2.3 获取工作表总列数 (ncols)17	2.3.16 清除指定范围的内容和格式38
2.2.4 row (索引) 获取对应的行17	2.3.17 测试指定范围的行标39
2.2.5 col (索引) 获取对应的列17	2.3.18 测试指定范围的列标40
2.2.6 使用字典向工作表中写入数据18	2.3.19 测试指定范围的“范围”, 列的 个数, 列的内容40
2.2.7 利用公式计算并填充数据19	2.4 其他应用41
2.2.8 修改源工作表中的数据方式 (修改内容)20	2.4.1 自动创建表格41
	2.4.2 在工作表中筛选数据46
	2.4.3 Python 文件打包输出48
	2.4.4 设置文件的图标49

2.4.5 为目标设置超链接	51	3.4.7 删除整行数据（按指定范围删除） ..	79
2.5 本章总结	51	3.4.8 删除整行数据（按切片方式删除） ..	80
第3章 Python 办公自动化之 pandas 库	54	3.4.9 有条件地删除整行数据	80
3.1 基本操作	54	3.5 有关工作表中列的操作	81
3.1.1 创建文件（无数据工作簿）	55	3.5.1 以列的方式合并两个工作表	81
3.1.2 创建文件（有数据工作簿）	55	3.5.2 追加列数据（追加空列）	82
3.1.3 创建文件（带索引的工作簿）	56	3.5.3 追加列数据（追加并填充数据）	83
3.1.4 判断工作表中数据内容的行数 和列数	56	3.5.4 删除列数据	83
3.1.5 显示工作表中数据的部分内容	57	3.5.5 插入列数据	84
3.1.6 在工作表中添加数据	58	3.5.6 修改列标题	85
3.1.7 工作表中数据的计算	59	3.5.7 删除空值	86
3.1.8 填充日期序列	59	3.6 本章总结	87
3.1.9 填充年份序列	60	第4章 Python 数据可视化之 matplotlib 库和 pyplot 库	88
3.1.10 填充月份序列	61	4.1 条形图与散点图	88
3.1.11 函数填充（求和）	62	4.1.1 绘制条形图	88
3.1.12 函数填充（平均值）	63	4.1.2 绘制散点图	90
3.2 数据分析统计	64	4.2 柱状图	91
3.2.1 排序（升序排列）	64	4.2.1 绘制普通柱状图	91
3.2.2 排序（降序排列）	65	4.2.2 分组柱状图及优化	92
3.2.3 多重排序	66	4.2.3 绘制叠加柱状图	93
3.2.4 数据筛选	66	4.3 饼图	95
3.2.5 提取工作表数据到列表和字典中	67	4.3.1 绘制普通饼图	95
3.2.6 数据分类汇总（按字符型汇总）	68	4.3.2 饼图优化	96
3.2.7 数据分类汇总（按数值型汇总）	68	4.3.3 绘制环形图	97
3.2.8 创建数据透视表	69	4.4 直方图与密度图	98
3.2.9 数据透视表分组	70	4.4.1 绘制直方图	98
3.3 csv、tsv、txt 文件与 Excel 文件的区别 与联系	71	4.4.2 直方图优化	99
3.3.1 读取 csv 文件到 Excel	71	4.4.3 绘制密度图	100
3.3.2 读取 tsv 文件到 Excel	72	4.5 折线图	101
3.3.3 读取 txt 文件到 Excel	73	4.5.1 绘制折线图	101
3.4 有关工作表中行的操作	74	4.5.2 折线图优化	102
3.4.1 合并两个工作表	74	4.5.3 折线图叠加	103
3.4.2 追加工作表中的数据	74	4.5.4 绘制多折线图	104
3.4.3 修改工作表中原有的值	75	4.6 其他常用图表	105
3.4.4 替换整行数据	76	4.6.1 绘制气泡图	105
3.4.5 插入整行数据	77	4.6.2 绘制雷达图	106
3.4.6 删除整行数据（按指定行删除）	78	4.6.3 绘制面积图	108
		4.6.4 绘制叠加区域图	109

4.6.5	绘制股票趋势图	110	5.4.9	边框设置	136
4.6.6	制作数据透视表	111	5.4.10	边框线型设置	136
4.7	本章总结	111	5.4.11	边框宽度设置	136
第 5 章	Python 第三方库简介	113	5.4.12	区域单元格内部边框设置	137
5.1	xlrd 库——快速读取 Excel 文件包	113	5.4.13	Alignment 位置设置	137
5.1.1	获取所有工作表名称	113	5.4.14	merge 拆分单元格	137
5.1.2	根据下标获取工作表名称	113	5.4.15	超链接设置	138
5.1.3	输出工作表名称、行数和列数	114	5.4.16	r 颜色设置	138
5.1.4	根据工作表名称获取整行和 整列的值	114	5.4.17	常用设置	139
5.1.5	获取指定单元格的内容	114	5.4.18	综合实例	139
5.1.6	获取单元格内容的数据类型	115	5.5	pandas 库——高效数据分析师	141
5.1.7	xlrd 命令汇总	115	5.5.1	创建 Series	141
5.2	xlwt 库——快速写入 Excel 文件包	118	5.5.2	创建 DataFrame（以数据方式 创建）	142
5.2.1	写入数据	118	5.5.3	创建 DataFrame（以字典的方式 创建）	142
5.2.2	设置单元格宽度	119	5.5.4	查看 DataFrame 数据信息	143
5.2.3	单元格中输入日期	119	5.5.5	对 DataFrame 标签排序	144
5.2.4	向单元格添加一个公式	120	5.5.6	对 DataFrame 数据排序	144
5.2.5	单元格中添加超链接	120	5.5.7	提取部分数据（单列数据）	145
5.2.6	工作表中合并单元格	121	5.5.8	提取部分数据（多行数据）	145
5.2.7	设置单元格内容的对齐方式	121	5.5.9	提取部分数据（条件数据）	146
5.2.8	单元格添加边框	122	5.6	openpyxl 库	146
5.2.9	单元格设置背景色	123	5.6.1	创建文件	147
5.3	xlutils 库——保留原有格式读取 Excel 文件包	123	5.6.2	打开已有文件	147
5.3.1	xlutils 对源文件备份	124	5.6.3	创建工作表（最后位置）	147
5.3.2	获取工作簿信息	124	5.6.4	创建工作表（开始位置）	148
5.3.3	复制后修改文件内容	124	5.6.5	选择工作表	149
5.3.4	获取所有单元格索引坐标	125	5.6.6	查看工作表名称	149
5.3.5	单元格内容的修改	126	5.6.7	访问单元格	149
5.4	xlwings 库——让 Excel 飞起来	127	5.6.8	访问行	150
5.4.1	App 常用 API: 启动操作	128	5.6.9	访问列	150
5.4.2	Book 常用 API: 工作簿操作	128	5.6.10	输入数据	151
5.4.3	Sheets 常用 API: 工作表操作	128	5.6.11	获取工作表总行数和总列数	151
5.4.4	Range 常用 API: 单元格操作	130	5.6.12	输出工作表中的数据	152
5.4.5	列所表示的意义	134	5.6.13	设置单元格风格	152
5.4.6	行所表示的意义	134	5.6.14	合并和拆分单元格	153
5.4.7	自动调整行高列宽	135	5.6.15	修改 sheet 标签按钮颜色	154
5.4.8	字体设置	135	5.6.16	删除工作表	155

5.7	matplotlib.pyplot 库	155	5.7.28	subplots()函数返回值	180
5.7.1	figure 函数	156	5.7.29	twinx()函数或 twiny()函数	181
5.7.2	subplot 函数 (建立单个子图)	157	5.8	numpy 库——科学计算软件包	182
5.7.3	axes 函数 (建立轴线)	158	5.8.1	数组的创建 (array 创建)	182
5.7.4	subplot 函数 (建立多个子图)	159	5.8.2	数组的创建 (arange 创建)	183
5.7.5	subplot 函数 (建立有轴线的子图)	159	5.8.3	特殊数组	184
5.7.6	绘制一条直线	160	5.8.4	数组索引	184
5.7.7	绘制多条直线	161	5.8.5	数组运算	185
5.7.8	绘制曲线	162	5.8.6	数组的复制 (浅复制)	186
5.7.9	创建 artist 对象	163	5.8.7	数组的复制 (深复制)	186
5.7.10	通过关键字参数绘图	164	5.8.8	创建矩阵	187
5.7.11	绘制散点图	164	5.8.9	矩阵运算	188
5.7.12	绘制条形图	165	5.8.10	numpy.dot()函数 (计算点积)	188
5.7.13	修改条形图	166	5.8.11	numpy.vdot()函数 (计算点积)	189
5.7.14	绘制直方图	167	5.8.12	numpy.inner()函数 (计算内积)	190
5.7.15	绘制饼图	168	5.8.13	numpy.matmul()函数 (计算 矩阵乘积)	190
5.7.16	绘制气泡图	169	5.8.14	numpy.linalg.det()函数 (计算 数组乘积)	191
5.7.17	绘制轮廓图	170	5.8.15	numpy.linalg.solve()函数 (求方程解)	192
5.7.18	设置区间上下限	171	5.8.16	numpy.linalg.solve()函数 (求逆矩阵)	193
5.7.19	添加图例说明	172	5.8.17	ndarray (创建一维数据)	193
5.7.20	设置区间分段	173	5.8.18	ndarray (创建多维数据)	194
5.7.21	设置布局 (具有不同的轴线)	174	5.9	本章总结	195
5.7.22	设置布局 (具有相同的轴线)	175		参考文献	196
5.7.23	设置轴相关问题	176			
5.7.24	正弦曲线	177			
5.7.25	subplot()函数	177			
5.7.26	subplot_adjust()函数	178			
5.7.27	subplots()函数	179			