

## 第 6 关、csv & excel

### 1、存储数据的方式

常用的存储数据的方式有两种——存储成 csv 格式文件、存储成 Excel 文件。

#### 1-1、csv 和 Excel 的区别

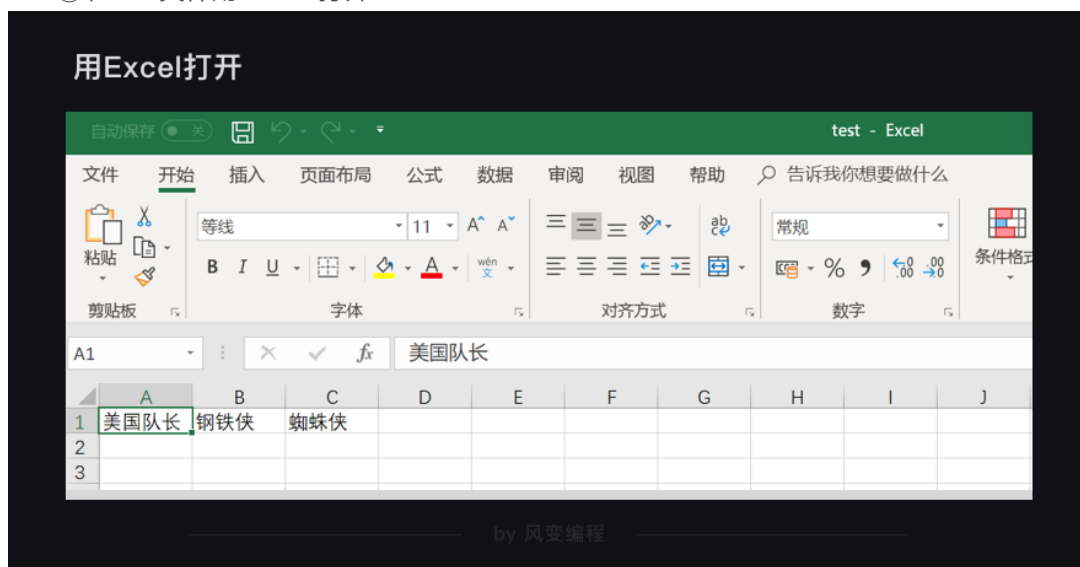
csv 也是一种字符串文件的格式，它组织数据的语法就是在字符串之间加分隔符——行与行之间是加换行符，同列之间是加逗号分隔。

它可以用任意的文本编辑器打开（如记事本），也可以用 Excel 打开，还可以通过 Excel 把文件另存为 csv 格式（因为 Excel 支持 csv 格式文件）。

##### ①、csv 文件用记事本打开



##### ②、csv 文件用 Excel 打开



csv 文件里的逗号可以充当分隔同列字符串的作用

## 1-2、csv 和 Excel 的优缺点

csv 格式存储数据，读写比较方便，易于实现，文件也会比 Excel 文件小。但 csv 文件缺少 Excel 文件本身的很多功能，比如不能嵌入图像和图表，不能生成公式。

Excel 文件，不用我多说你也知道就是电子表格。它有专门保存文件的格式，即 xls 和xlsx（Excel2003 版本的文件格式是 xls，Excel2007 及之后的版本的文件格式就是xlsx）



## 2、存储数据的基础知识



### 2-1、csv 写入与读取

#### (1) 模块调用

因为 csv 为内置函数，无需安装，即可使用

```
1 import csv
```

#### (2) 创建 csv 文件

```
1 import csv
2 #引用csv模块。
3 csv_file = open('demo.csv','w',newline='',encoding='utf-8')
```

```
4 #创建csv文件，我们要先调用open()函数，传入参数：文件名“demo.csv”、写入模式“w”、  
   newline=''、encoding='utf-8'。
```

- 'demo.csv' 为我们新创建的一个 csv 文件；
- 'w' 就是 writer，即文件写入模式，它会以覆盖原内容的形式写入新添加的内容；

open()函数				
	b(bytes,字节)		+	b +
r(read,读)	r只读，指针在开头 文件不存在则报错	rb二进制只读， 其余同左	r+读写， 其余同左	rb+二进制读写， 其余同左
w(write,写)	w只写，文件不存在 则新建，存在则覆盖	wb二进制只写， 其余同左	w+读写， 其余同左	wb+二进制读写， 其余同左
a(append,追加)	a追加，文件存在指 针放在末尾，文件 不存在则新建	ab二进制追加， 其余同左	a+追加且可读， 其余同左	ab+二进制追加， 且可读，其余同左

by 风变编程

- newline='' 参数可以避免 csv 文件出现两倍的行距（就是能避免表格的行与行之间出现空白行）；
- encoding='utf-8' 可以避免编码问题导致的报错或乱码。

### (3) 创建 writer 对象

```
1 import csv  
2 #引用csv模块。  
3 csv_file = open('demo.csv','w',newline='',encoding='utf-8')  
4 #调用open()函数打开csv文件，传入参数：文件名“demo.csv”、写入模式“w”、  
   newline=''、encoding='utf-8'。  
5 writer = csv.writer(csv_file)  
6 # 用csv.writer()函数创建一个writer对象。
```

### (4) 往 csv 文件写入新的内容

调用 writer 对象的 `writerow()` 方法（提醒：`writerow()` 函数里，需要放入列表参数，所以我们得把要写入的内容写成列表。就像 ['电影','豆瓣评分']）

```
1 import csv  
2 #引用csv模块。  
3 csv_file = open('demo.csv','w',newline='',encoding='utf-8')  
4 #调用open()函数打开csv文件，传入参数：文件名“demo.csv”、写入模式“w”、  
   newline=''、encoding='utf-8'。  
5 writer = csv.writer(csv_file)  
6 # 用csv.writer()函数创建一个writer对象。  
7 writer.writerow(['电影','豆瓣评分'])  
8 #调用writer对象的writerow()方法，可以在csv文件里写入一行文字 “电影”和“豆瓣评分”。  
9 writer.writerow(['银河护卫队','8.0'])  
10 #在csv文件里写入一行文字 “银河护卫队”和“8.0”。  
11 writer.writerow(['复仇者联盟','8.1'])  
12 #在csv文件里写入一行文字 “复仇者联盟”和“8.1”。  
13 csv_file.close()
```

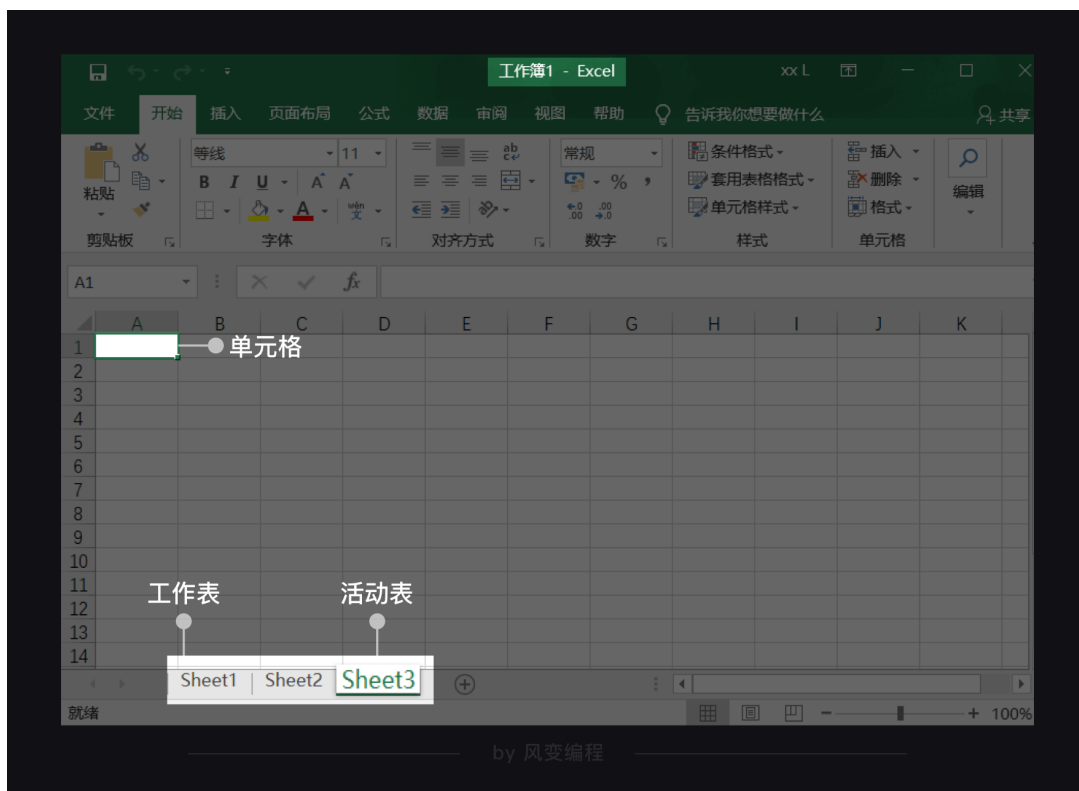


csv模块本身还有很多函数和方法，附上csv模块官方文档链接：[https://yiyibooks.cn/x/python\\_352/library/csv.html#module-csv](https://yiyibooks.cn/x/python_352/library/csv.html#module-csv)



## 2-2、Excel 写入与读取

### 2-2-1、Excel 的基本概念



- 一个Excel文档也称为一个工作簿（workbook），每个工作簿里可以有多个工作表（worksheet），当前打开的工作表又叫活动表；
- 每个工作表里有行和列，特定的行与列相交的方格称为单元格（cell）。比如上图第A列和第1行相交的方格我们可以直接表示为A1单元格。

## 2-2-2、Excel 的写入读写

### （1）模块安装

- ①、window 电脑：在终端输入命令：pip install openpyxl，按下 enter 键；
- ②、mac 电脑：在终端输入命令：pip3 install openpyxl，按下 enter 键。

### （2）模块调用

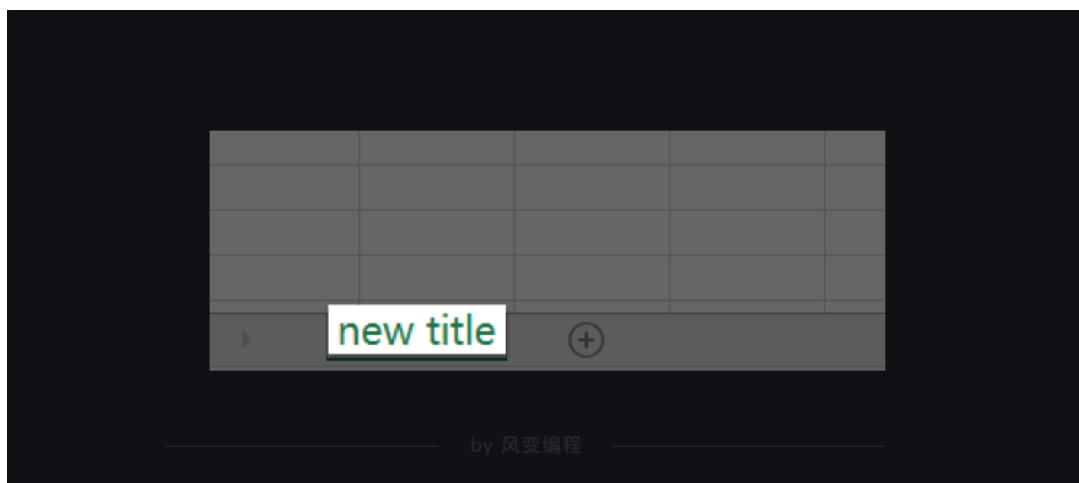
```
1 import openpyxl
```

### （3）创建新的工作簿

```
1 import openpyxl
2 #引用openpyxl。
3 wb = openpyxl.Workbook()
4 #利用openpyxl.Workbook()函数创建新的workbook（工作簿）对象，就是创建新的空的Excel文件。
```

### （4）获取工作表

```
1 sheet = wb.active
2 #wb.active就是获取这个工作簿的活动表，通常就是第一个工作表。
3 sheet.title = 'new title'
4 #可以用.title给工作表重命名。现在第一个工作表的名称就会由原来默认的“sheet1”改为“new title”。
```



#### (5) 单元格内写入内容

```
1 sheet['A1'] = '漫威宇宙'
2 #把'漫威宇宙'赋值给第一个工作表的A1单元格，就是往A1的单元格中写入了'漫威宇宙'。
```

可用 `append` 函数往工作表内写入一行内容

```
1 row = ['美国队长', '钢铁侠', '蜘蛛侠']
2 #把我们想写入的一行内容写成列表，赋值给row。
3 sheet.append(row)
4 #用sheet.append()就能往表格里添加这一行文字。
```

#### (6) 保存 Excel 文件

```
1 wb.save('Marvel.xlsx')
2 #保存新建的Excel文件，并命名为“Marvel.xlsx”
```

#### (7) 完整代码

```
1 import openpyxl
2 #写入的代码：
3 wb = openpyxl.Workbook()
4 sheet = wb.active
5 sheet.title = 'new title'
6 sheet['A1'] = '漫威宇宙'
7 rows = [['美国队长', '钢铁侠', '蜘蛛侠', '雷神'], ['是', '漫威', '宇宙', '经典', '人物']]
8 for i in rows:
9     sheet.append(i)
10 print(rows)
11 wb.save('Marvel.xlsx')
12
13 #读取的代码：
14 wb = openpyxl.load_workbook('Marvel.xlsx')
15 sheet = wb['new title']
16 sheetname = wb.sheetnames
17 print(sheetname)
18 A1_cell = sheet['A1']
19 A1_value = A1_cell.value
20 print(A1_value)
```

## Excel文件写入的步骤

### 0 创建工作簿

利用openpyxl.Workbook()创建workbook对象

### 1 获取工作表

借助workbook对象的active属性

### 2 操作单元格

单元格: sheet['A1']; 一行: append()

### 3 保存工作簿

save()

by 风变编程

## 2-3、项目：存储周杰伦的歌曲信息

```
1 import requests,openpyxl
2 wb=openpyxl.Workbook()
3 #创建工作簿
4 sheet=wb.active
5 #获取工作簿的活动表
6 sheet.title='restaurants'
7 #工作表重命名
8
9 sheet['A1'] = '歌曲名'      #加表头，给A1单元格赋值
10 sheet['B1'] = '所属专辑'    #加表头，给B1单元格赋值
11 sheet['C1'] = '播放时长'    #加表头，给C1单元格赋值
12 sheet['D1'] = '播放链接'    #加表头，给D1单元格赋值
13
14 url = 'https://c.y.qq.com/soso/fcgi-bin/client_search_cp'
15 for x in range(5):
16     params = {
17         'ct': '24',
18         'qqmusic_ver': '1298',
19         'new_json': '1',
20         'remoteplace': 'sizer.yqq.song_next',
21         'searchid': '64405487069162918',
22         't': '0',
23         'aggr': '1',
24         'cr': '1',
25         'catZhida': '1',
26         'lossless': '0',
27         'flag_qc': '0',
28         'p': str(x + 1),
29         'n': '20',
30         'w': '周杰伦',
31         'g_tk': '5381',
32         'loginUin': '0',
33         'hostUin': '0',
34         'format': 'json',
```

```
35     'inCharset': 'utf8',
36     'outCharset': 'utf-8',
37     'notice': '0',
38     'platform': 'yqq.json',
39     'needNewCode': '0'
40 }
41
42 res_music = requests.get(url, params=params)
43 json_music = res_music.json()
44 list_music = json_music['data']['song']['list']
45 for music in list_music:
46     name = music['name']
47     # 以name为键，查找歌曲名，把歌曲名赋值给name
48     album = music['album']['name']
49     # 查找专辑名，把专辑名赋给album
50     time = music['interval']
51     # 查找播放时长，把时长赋值给time
52     link = 'https://y.qq.com/n/yqq/song/' + str(music['file']
53 ['media_mid']) + '.html\n\n'
54     # 查找播放链接，把链接赋值给link
55     sheet.append([name, album, time, link])
56     # 把name、album、time和link写成列表，用append函数多行写入Excel
57     print('歌曲名: ' + name + '\n' + '所属专辑:' + album + '\n' + '播放时
58 长:' + str(time) + '\n' + '播放链接:'+ url)
59
60 wb.save('Jay.xlsx')
61 #最后保存并命名这个Excel文件
```