

Excel文档与python

Excel 是 Windows 环境下流行的、强大的电子表格应用。openpyxl 模块让 Python 程序能读取和修改 Excel 电子表格文件。例如,可能有一个无聊的任务,需要从一个电子表格拷贝一些数据,粘贴到另一个电子表格中。或者可能需要从几千行中挑选几行,根据某种条件稍作修改。或者需要查看几百份部门预算电子表格,寻找其中的赤字。正是这种无聊无脑的电子表格任务,可以通过 Python 来完成。LibreOffice Calc, WPS 和 OpenOffice Calc 都能处理 Excel 的电子表格文件格式,这意味着 openpyxl 模块也能处理来自这些应用程序的电子表格。你可以从 <https://www.libreoffice.org/>和 <http://www.openoffice.org/>下载这些软件。即使你的计算机上已经安装了 Excel,可能也会发现这些程序更容易使用。

1. Excel文档基本定义

- 工作簿(workbook)：一个 Excel 电子表格文档；
- 工作表(sheet)：每个工作簿可以包含多个表, 如：sheet1，sheet2等；
- 活动表(active sheet)：用户当前查看的表；
- 列(column): 列地址是从 A 开始的；
- 行(row): 行地址是从 1 开始的；
- 单元格(cell)：特定行和列的方格；

2. 安装 openpyxl 模块

Python 没有自带 openpyxl,所以必须安装。

```
pip3 install openpyxl
```

3. 读取 Excel 文档

将使用一个电子表格 `example.xlsx`, 你可以自己创建这个电子文档, 内容如下:



- 在导入 openpyxl 模块后,就可以使用 openpyxl.load_workbook()函数。打开 Excel 文档

```
# 导入工作簿
```

```
1. wb = openpyxl.load_workbook('excelDemo/example.xlsx')    # 加载工作簿
2. wb.sheetnames      # 获取当前所有工作表的名称， 返回一个列表
3. wb.active          # 获取当前活跃的工作表
```

- 从工作簿中取得工作表

```
4. sheet = wb['Sheet1']           # 工作表
5. sheet.title                     # 获取当前活动表的名称
```

```
6. sheet.cell(row=1, column=2)          # 获取单元格指定行和指定列的内容
```

- 表中取得单元格

Cell对象有一个 value 属性,不出意外,它包含这个单元格中保存的值。Cell对象也有 row、column 和 coordinate 属性,提供该单元格的位置信息。

row属性给出的是整数 1,column 属性给出的是'B',coordinate 属性给出的是'B1'。

```
6. cell = sheet['A1']
7. cell_value = sheet['A1'].value
8. cell.row, cell.column cell.coordinate
```

4. 工作簿、工作表、单元格

下面是从电子表格文件中读取单元格涉及的所有函数、方法和数据类型。一个完整的操作过程如下:

- 导入 openpyxl 模块。
- 调用 openpyxl.load_workbook()函数。
- 取得 Workbook 对象。
- 调用 wb.sheetnames和 wb.active 获取工作簿详细信息。
- 取得 Worksheet 对象。
- 使用索引或工作表的 cell()方法,带上 row 和 column 关键字参数。
- 取得 Cell 对象。
- 读取 Cell对象的 value 属性

一个完整的案例代码如下:

```
import openpyxl

# 1. 读取excel文档
wb = openpyxl.load_workbook('excelDemo/example.xlsx')

# # 返回一个workbook对象， 有点类似于文件对象；
# print(wb, type(wb))

# 2. 在工作簿中取得工作表
# print(wb.get_sheet_names())
# 返回一个列表， 存储excel表中所有的sheet工作表；
print(wb.sheetnames)

# 返回一个worksheet对象， 返回当前的活动表；
# print(wb.get_active_sheet())
# print(wb.active)

# 3. 获取工作表中， 单元格的信息
# wb.get_sheet_by_name('Sheet1')
sheet = wb['example']
print(sheet['A1'])
```

```
print(sheet['B1'].value)

cell = sheet['B1']
print(cell.row, cell.column)

print(sheet.cell(row=3, column=2))
print(sheet.cell(row=3, column=2).value)
print(sheet.cell(row=3, column=2, value='www'))

# sheet的属性

print(sheet.max_column)
print(sheet.max_row)
print(sheet.title)
sheet.title = 'example'
print(sheet.title)

for row in sheet.rows:
    for cell in row:
        print(cell.value, end='\t')
    print('\n')

wb.save(filename="excelDemo/example.xlsx")
```