

Section 14

Excel, Word, and PDF Documents

Mohammed Asir Shahid

2021-08-05

Contents

1	Reading Excel Spreadsheets	1
----------	-----------------------------------	----------

1 Reading Excel Spreadsheets

The `openpyxl` module lets us modify Excel files using Python. It is a third party module that we'll need to install ourselves.

```
pip install openpyxl
```

The Excel document is called a workbook that is saved by `.xlsx` file extension. Each workbook contains sheets/worksheets. Inside each sheet there are columns (letters) and rows (numbers). The intersection of a column and row is called a cell.

```
import openpyxl,os

workbook=openpyxl.load_workbook("example.xlsx")

print(type(workbook))

print(workbook.get_sheet_names())

sheet=workbook.get_sheet_by_name("Sheet1")
print(type(sheet))
```

```

cell=sheet["A1"]

print(type(cell))
print(cell.value)

cell=sheet["B1"]

print(type(cell))
print(cell.value)

cell=sheet["C1"]

print(type(cell))
print(cell.value)

print(type(cell))
print(cell.value)

for i in range(1,8):
    print(i, sheet.cell(row=i, column=2).value)

<class 'openpyxl.workbook.workbook.Workbook'>
['Sheet1', 'Sheet2', 'Sheet3']
<class 'openpyxl.worksheet.worksheet.Worksheet'>
<class 'openpyxl.cell.cell.Cell'>
2015-04-05 13:34:02
<class 'openpyxl.cell.cell.Cell'>
Apples
<class 'openpyxl.cell.cell.Cell'>
73
<class 'openpyxl.cell.cell.Cell'>
73
1 Apples
2 Cherries
3 Pears
4 Oranges
5 Apples
6 Bananas

```

7 Strawberries