

第五周学习笔记

查看哪些同学没有提交用户名, 用命令 `openpyxl`, 代码可以从官网文档复制粘贴。
`pandas.read_excel` 和 `polars.read_excel` (三种引擎) 可以读取 excel 文件, 读取为表格形式的数据, `pandas` 也可以读写 stata 的 `dta` 文件; 如果 excel 文件不规范, 需要用底层软件包, 如 `openpyxl`。

```
! environment.yml
1  name: names
2  channels:
3    - conda-forge
4  dependencies:
5    - python=3.12
6    - ruff
7    - openpyxl
8    - pandas
9    - polars-lts-cpu
```

调用函数前可以先看 `example`, 举一反三。

`io` 参数必须给出, 否则报错, 支持字符串和字节串等

```
import pandas as pd

df = pd.read_excel("选课名单-24专硕")

from openpyxl import load_workbook

wb = load_workbook(filename="empty_book.xlsx")
sheet_ranges = wb["range names"]
print(sheet_ranges["D18"].value)
```

接下来，调试代码.....

最后检查 columns

新增知识： 浏览器搜索快捷键 Ctrl+F

修改代码

```
import pandas as pd

df = pd.read_excel("选课名单-24专硕.xlsx", skiprows = 3)

from openpyxl import load_workbook

wb = load_workbook(filename="empty_book.xlsx")
sheet_ranges = wb["range names"]
print(sheet_ranges["D18"].value)
```

```
import pandas as pd

df = pd.read_excel("选课名单-24专硕.xlsx", skiprows=3, usecols="A:D")
breakpoint()

from openpyxl import load_workbook

wb = load_workbook(filename="empty_book.xlsx")
sheet_ranges = wb["range names"]
print(sheet_ranges["D18"].value)
```