**Python与数据分析**

**实验报告**

|  |  |  |
| --- | --- | --- |
| **课程名称：**Python与数据分析 | **学期：**2022~2023学年上学期 | **成绩：** |
| **指导教师：**王灿 | **学生姓名：** | **学生学号：** |
| **实验名称：**网络数据的获取与解析 | | |
| **实验编号：**13 | **实验日期：**11月7日 | **实验学时：**NA |
| **学院：**理学院 | **专业：**信息与计算科学 | **班级：**信计xxxx |

**一、实验目的**

1．熟悉网络数据的获取。

2．熟悉网页的解析和内容提取。

**二、实验内容**

1. 获取网络数据。使用urllib库的方法读取网页<http://www.wanglabtools.cn/teach/python3/html/appleleaf/index.html> 的html文档，保存到html\_doc变量里
2. 网页的解析和内容提取。
   1. 使用BeautifulSoup4库解析上一问得到的html\_doc，结果保存到soup变量里。
   2. 使用.find(string=…)或者不断进行.xxx的方法获取文字“展示4张苹果叶照片”
   3. 使用.find\_all()找到所有4个img标签，写循环从4个img标签中把4张图的url提取到一个列表里，并打印这个列表

**三、实验环境**

Windows/Linux/Mac OS操作系统；conda创建的python3.10环境。

**四、实验过程和结果**

**1．实验步骤、结果和结论（包括必要的截图）**

**2．关键代码及其解释**

**3．调试过程（如果和后面的实验总结重复，可以只写实验总结）**

**五、实验总结（不能为空）**

**1．遇到的问题及解决过程**

**2．产生的错误及原因分析**

**3****．体会和收获**

**六、参考文献**

[1]柳毅等著，Python数据分析与实践，北京：清华大学出版社，2019.7.

**七、教师评语**