



清华大学
Tsinghua University

小伙伴计划暑期学习营——零基础Python入门

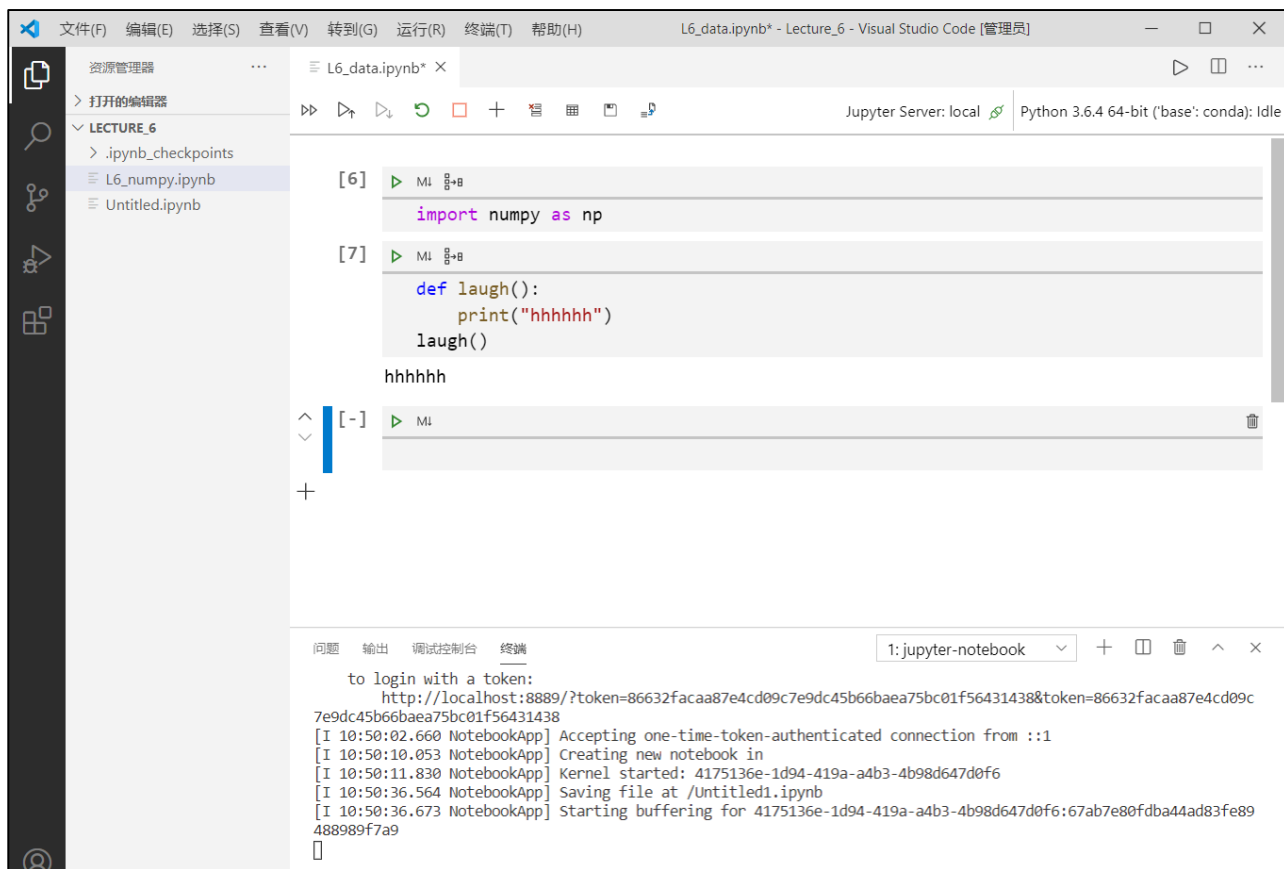
第四讲：玩转文件——预习

张智帅 电子系

清华大学学生学业与发展指导中心
2019-2020学年夏季学期

Jupyter notebook

- ❑ Vscode支持原生notebook❤️❤️❤️
- ❑ 把文件保存为.ipynb格式即可自动识别



python-docx

- ❑ 安装: `pip install python-docx`
- ❑ 导入: `import docx`
- 读取文档:

```
import docx

file = docx.Document("背影.docx") # 获取文档对象
print("段落数:" + str(len(file.paragraphs))) # 读取文档的段落数

# 输出段落编号及段落内容的前三十个字
for i, paragraph in enumerate(file.paragraphs):
    print("第" + str(i) + "段: " + paragraph.text[:30])
```

段落数:12

第0段: 背影

第1段: 我与父亲不相见已二年余了,我最不能忘记的是他的背影。那年冬天

第2段: 回家变卖典质,父亲还了亏空;又借钱办了丧事。这些日子,家中光

第3段: 到南京时,有朋友约去游逛,勾留了一日;第二日上午便须渡江到浦

第4段: 我们过了江,进了车站。我买票,他忙着照看行李。行李太多了,得

第5段: 我说道,“爸爸,你走吧。”他望车外看了看,说,“我买几个橘子