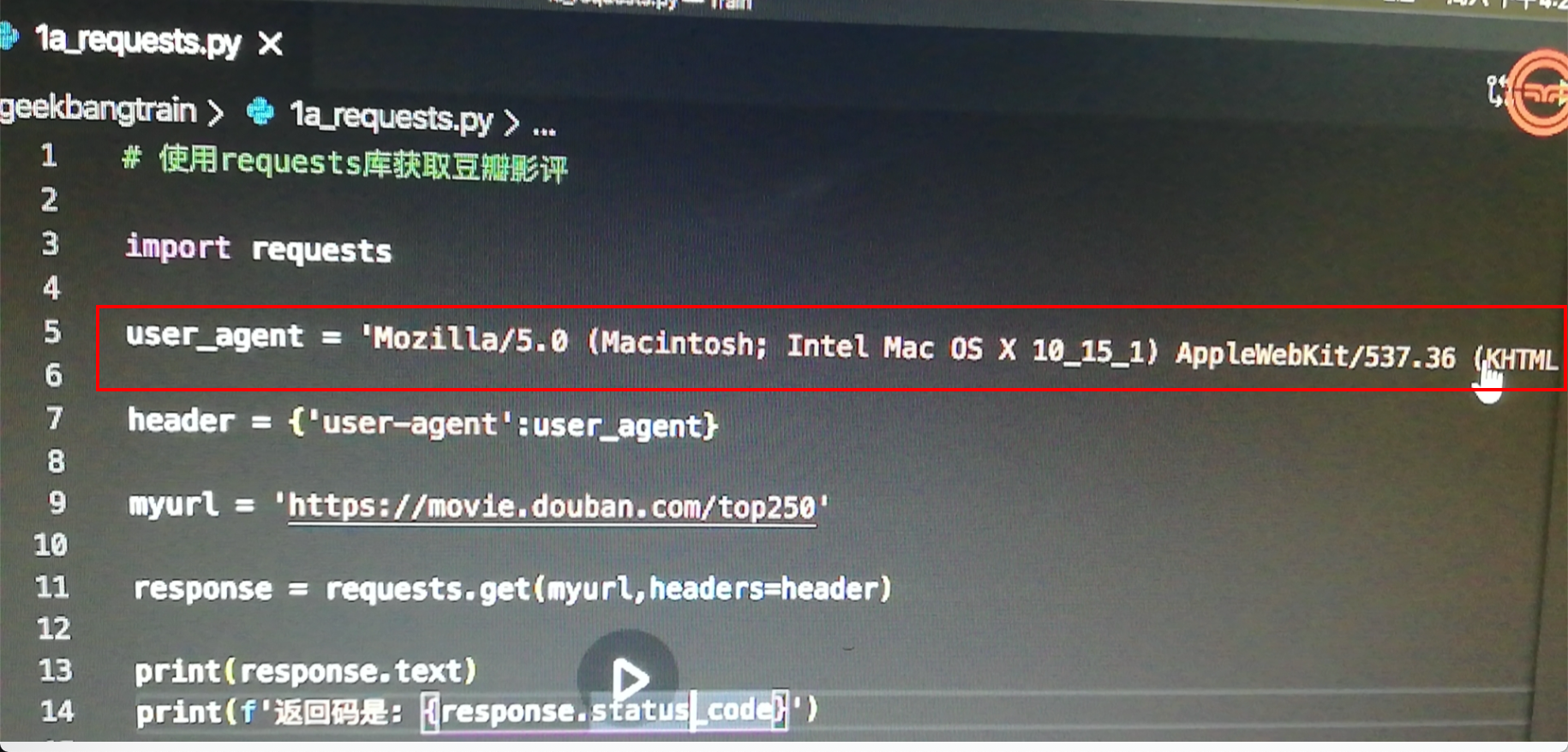
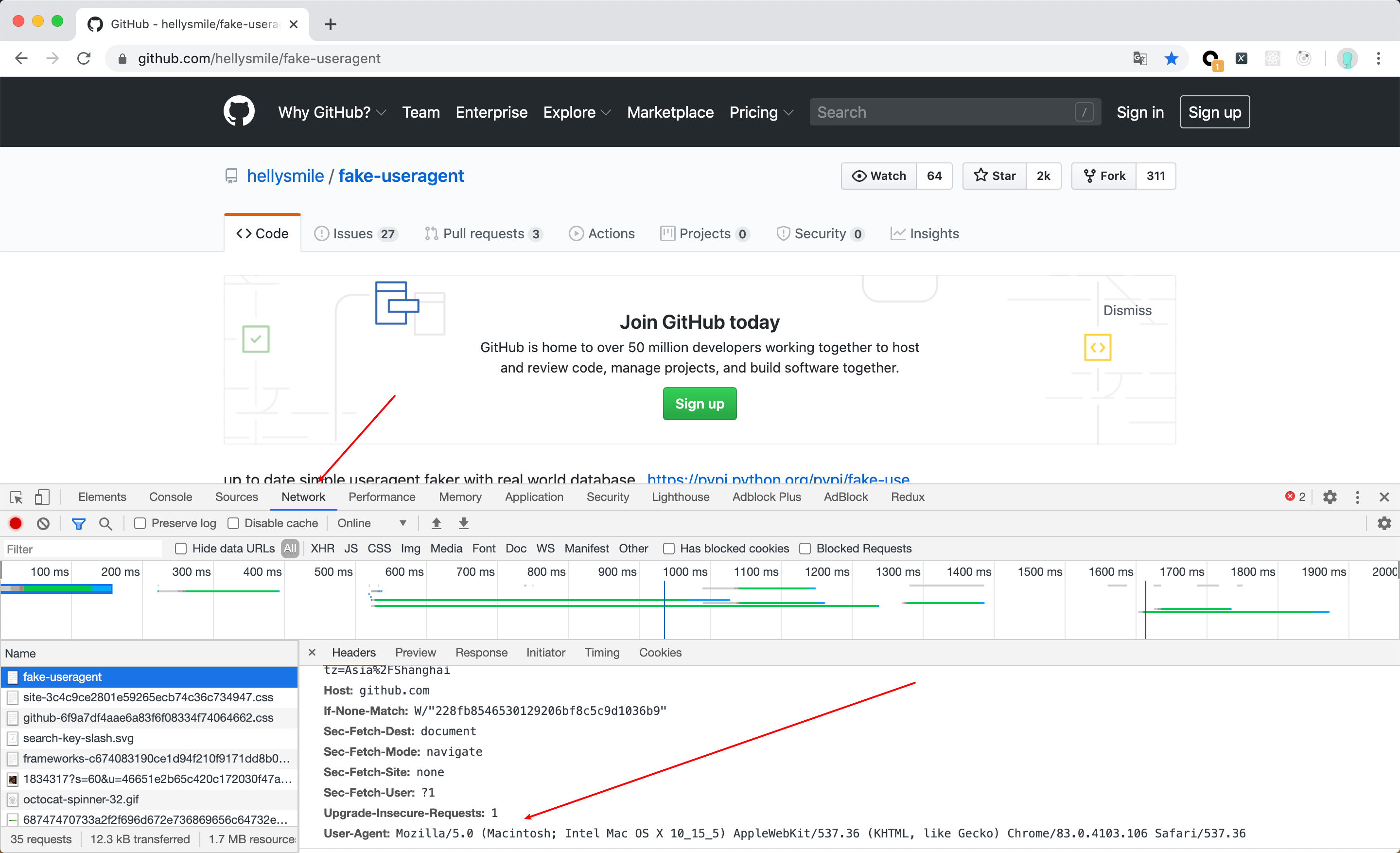
**Python 进阶训练营常见问题汇总**

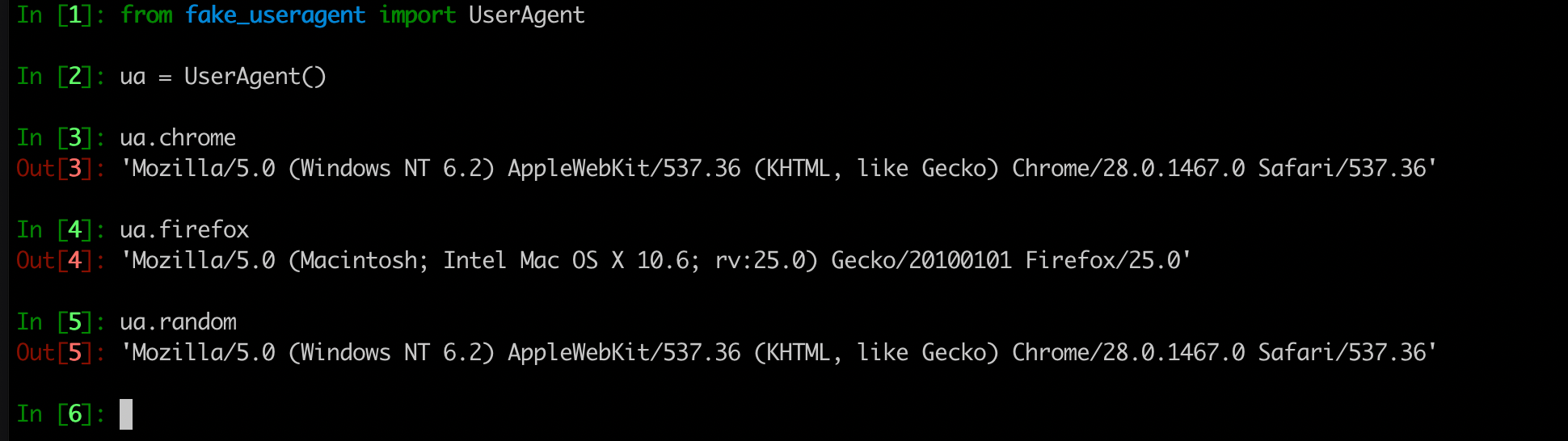
1. **关于 User-Agent：**



User-Agent 不必跟老师写的一样，只要在任意浏览器打开开发者工具把 User-Agent 复制出来就行。

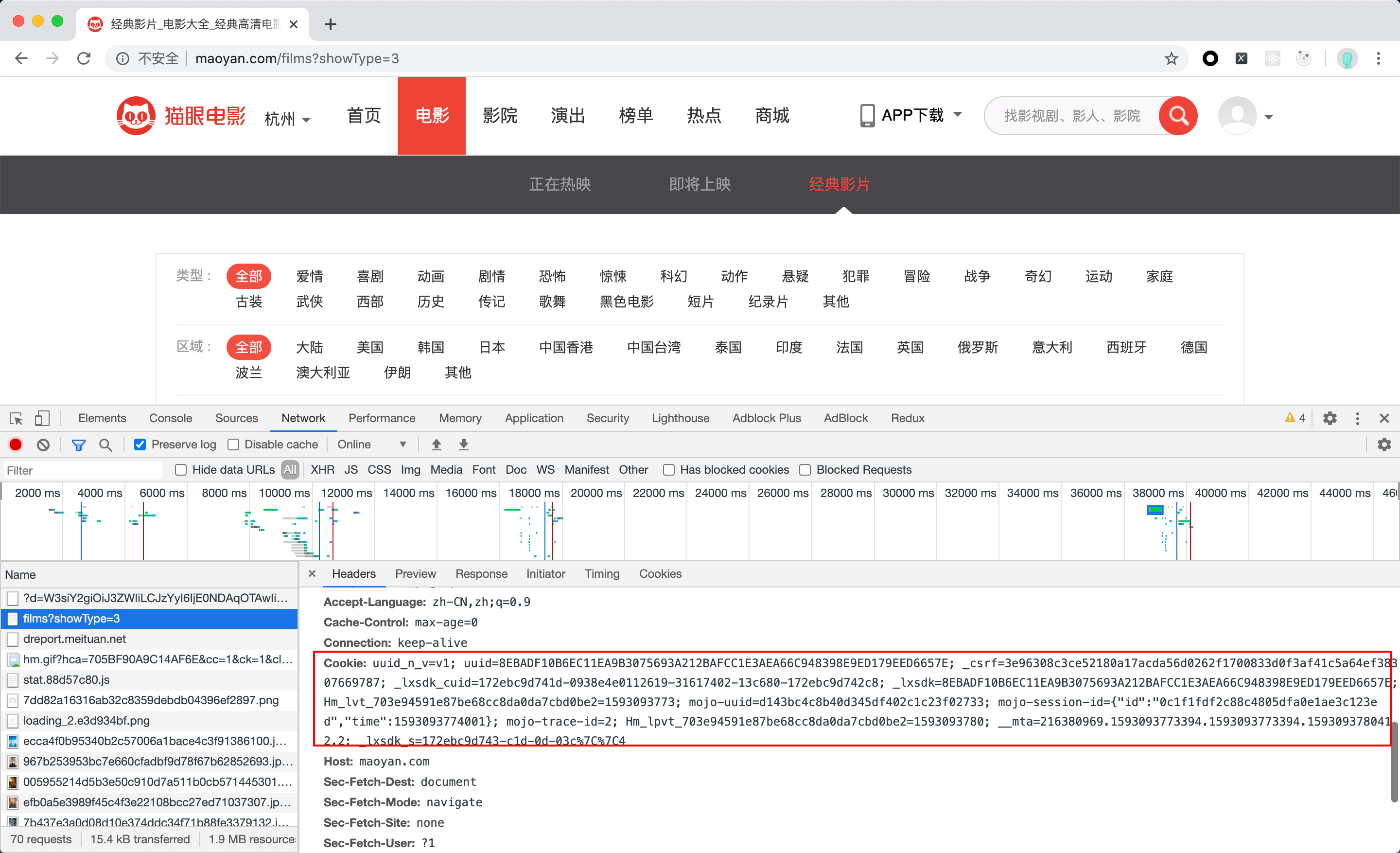


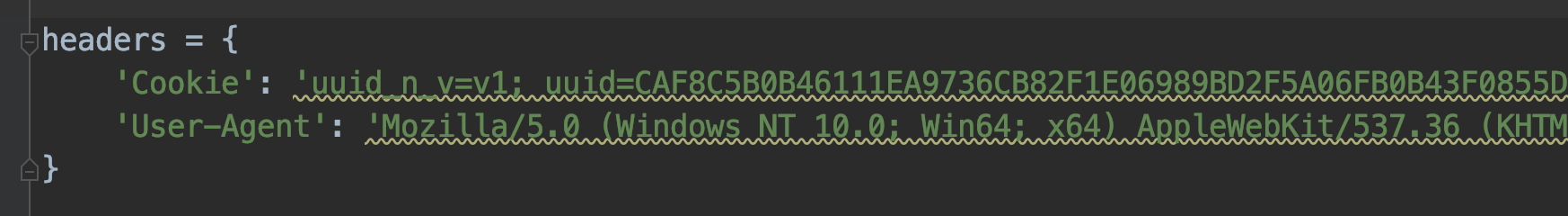
也可以使用 https://github.com/hellysmile/fake-useragent 这个库随机生成，简单示例如下，更多使用方法可以参考 github 文档。



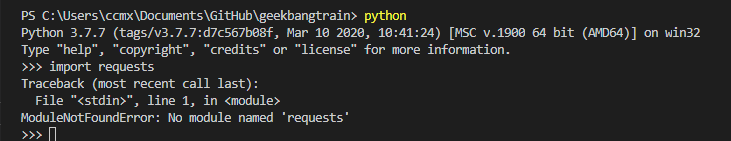
1. **关于爬取猫眼电影出现美团验证码的问题**

是被猫眼网站反爬了，解决办法就是把浏览器的 cookie 复制出来，添加到请求头中即可

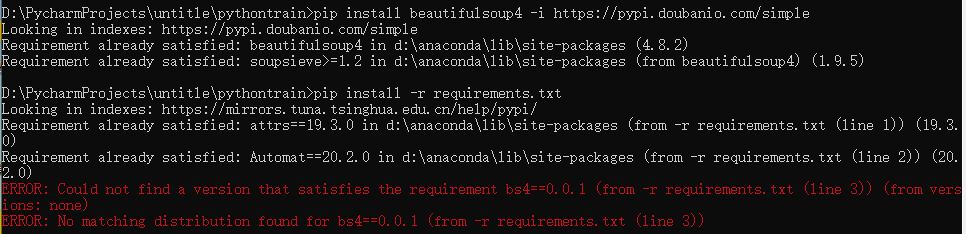




1. **关于导包失败问题**



1. 可能是库没有安装，使用 pip install requests 类似命令即可安装。
2. 如果你改成了国内源，有些源的包不全，提示找不到包之类的问题



请使用豆瓣源，pip install xxx -i <https://pypi.douban.com/simple>

1. 你还可能遇到使用 vscode 时，vscode 配置的 python 不是你安装了包的 python，所以你可以给 vscode 配置你需要的那个 python

在项目根目录下新建一个 .vscode/ 目录，然后在 .vscode/ 目录下新建 settings.json 文件，

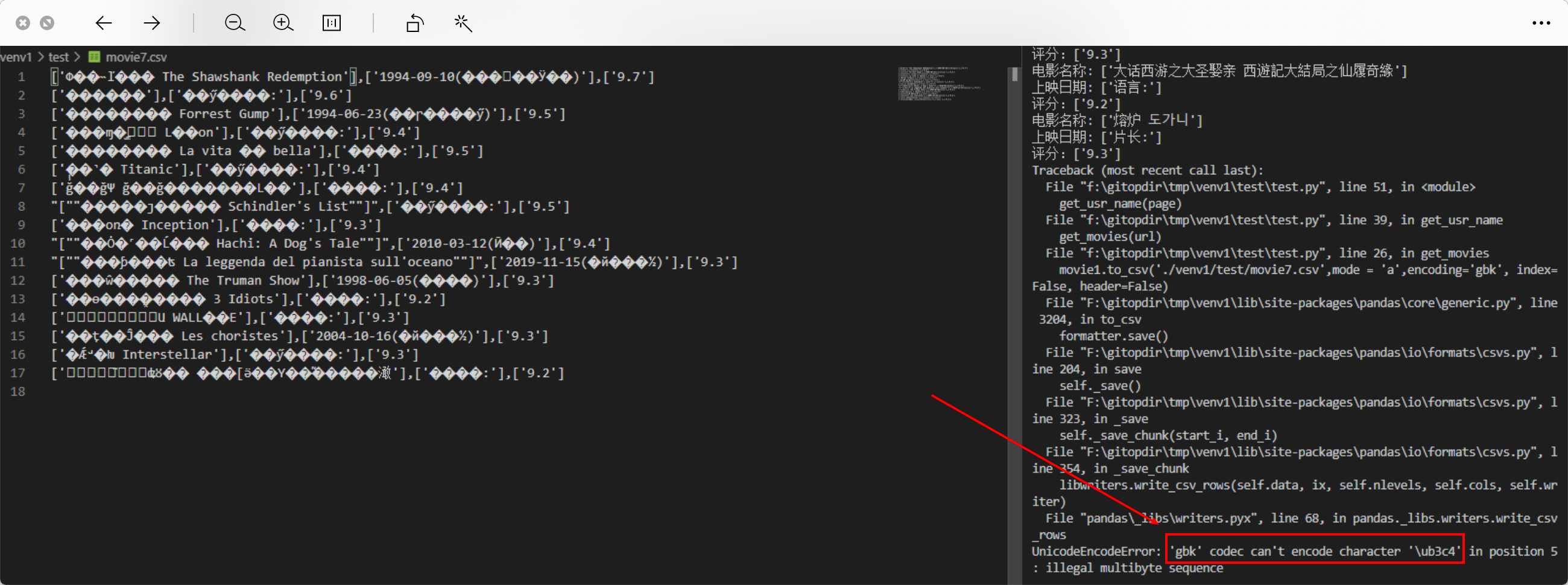
再把你安装了 requests 的 python 绝对路径写上去试试

{

"python.pythonPath": "XXX\\python.exe"

}

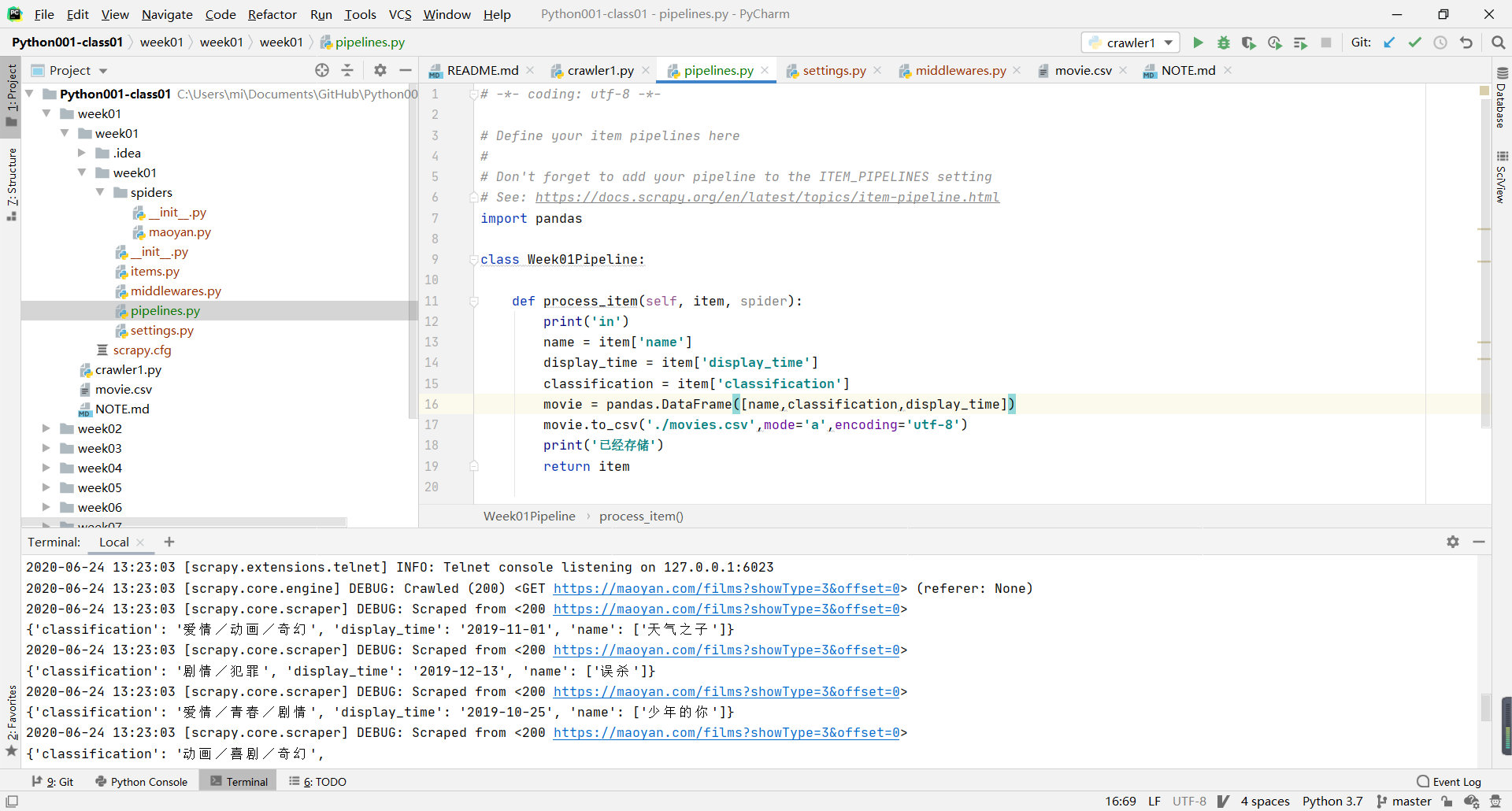
1. **使用 pandas 的 to\_csv 方法保存 gbk 格式报错**

****

这是因为有些字符不在 gbk 编码，所以不能保存，请使用 utf8 格式，如果使用 utf8 格式保存的 .csv 文件，使用 excel 打开时遇到中文乱码，请保存成 utf\_8\_sig 格式

movie1.to\_csv('movie11.csv', mode='a', encoding=' utf\_8\_sig ', index=False, header=False)

1. **代码不走 pipelines.py 文件**



在 settings.py 中把 ITEM\_PIPELINES 加进去

1. **GitHub 慢**

请使用 gitee：<https://gitee.com/wilsonyin/pythontrain.git>

1. **关于 find\_all 和 find 方法中那个逗号的问题**



这两个逗号可以不写，如下，写不写在 python 中语法都是可以通过的。



其实你可以简单的理解为这个语法是为了你扩展代码方便，比如一个函数可以传俩参数，第一个必传第二个选传，你第一次写的时候只传一个，下次如果新的需求来了，需要传第二个，你就得加一个逗号，然后写上第二个参数。但是如果你以前就把逗号写上了，它也不错，你传第二个参数的时候就不用在写逗号了。

1. **关于猫眼的反爬**

一种解决办法是你把 cookie 加上，见问题2

另一种办法是可以换个思路，可以参考 @凌金隆-珠海 同学爬取手机端的思路来爬取数据

1. **关于一些类似 BeautifulSoup 获取不到数据等问题**

这种问题没有统一解，只能靠你自己多实践才能提升自己编程能力

比如你获取的标签不对，那么你就多测试几个就好了

如果你语法不熟悉，想要获取一个数据不知道如何操作，那么请多看官方文档

1. **关于Git**

Git 的相关问题建议大家回头去看下预习课程哈，我们的课程讲解的 git 足够大家交作业用了，如果学习了课程还是解决不了，再到群里来问哈