**大数据财税与python应用**

**2019-2020年第3学期**

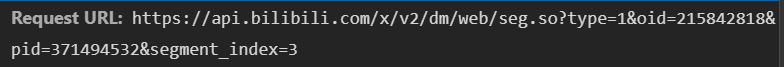
学号：2019312256

姓名：陈祥锡

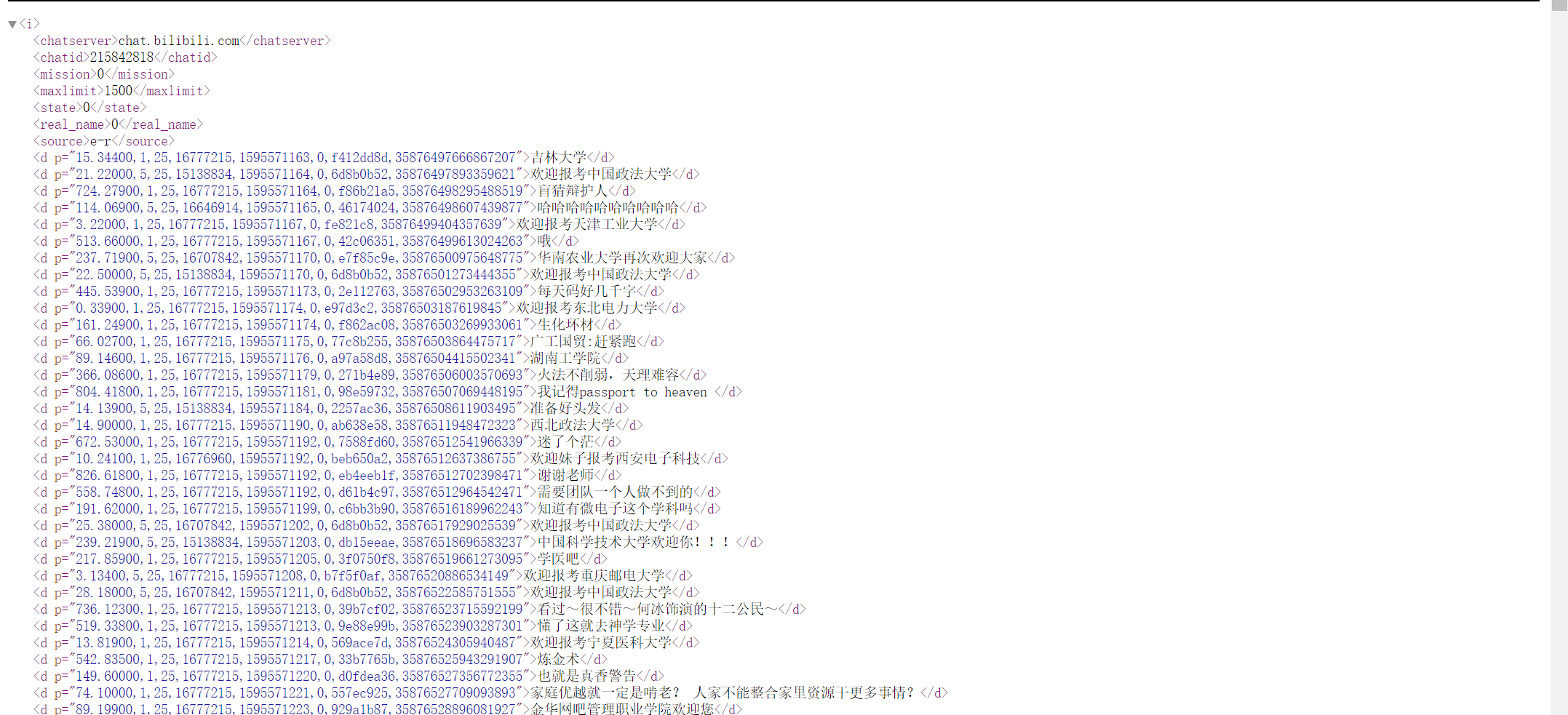
专业：电子商务类19-1

此python程序的目的是爬取罗翔老师7月24日发布在bilibili上的视频——“【罗翔】大学报志愿，选法学怎么样？”弹幕并制作词云图。

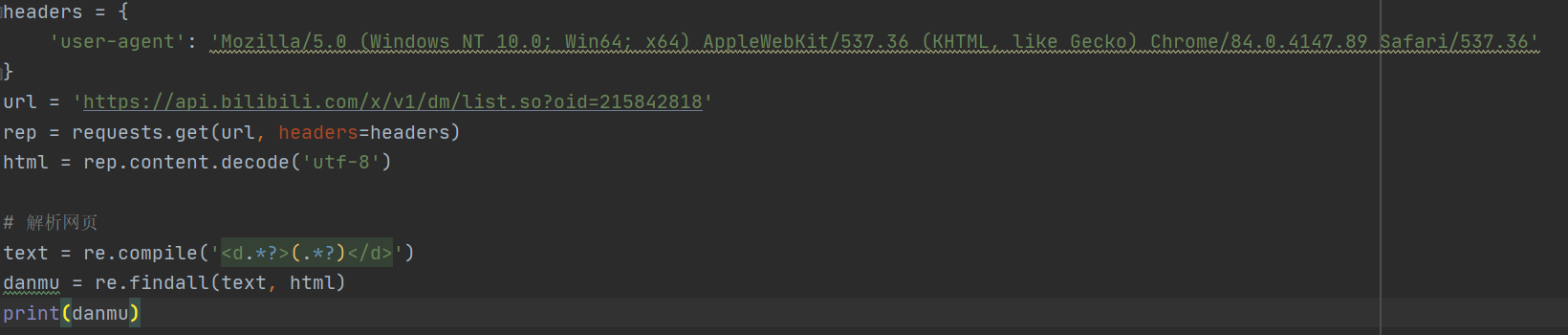
1、使用检查工具找到视频的oid:



爬取的目标网站即为<https://api.bilibili.com/x/v1/dm/list.so?oid=215842818>（弹幕列表网站）

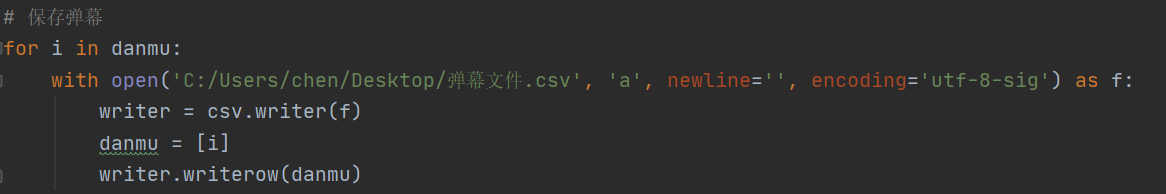


2、爬取以及解析网站内容



发现弹幕内容都是在<d>中，即可通过这个规律获得弹幕

3、保存弹幕



将弹幕写入csv文件

4、制作词云图



.read()获取每条弹幕；jieba.lcut()将每条弹幕分割成词语；wordcould.WordCould()设置相关参数；.generate()将信息传入照片