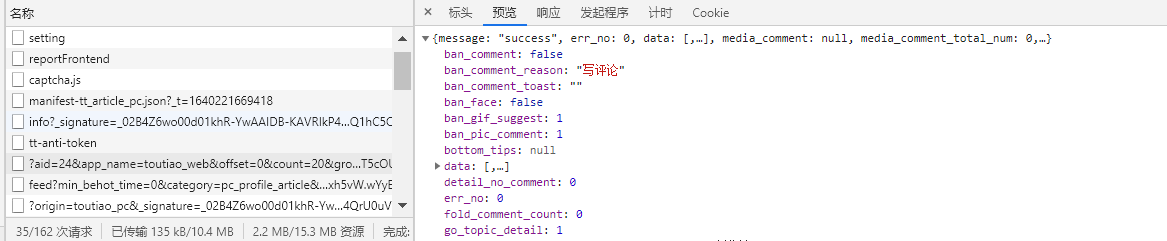
1. 数据爬取

爬取今日头条，url = <https://www.toutiao.com/> 的新闻评论信息。

通过分析url请求的情况，可以看到：



评论的url在一个异步http请求中，返回来的是一个json结构。因此，可以直接对该url进行直接请求，并使用json库对返回的数据进行解析。

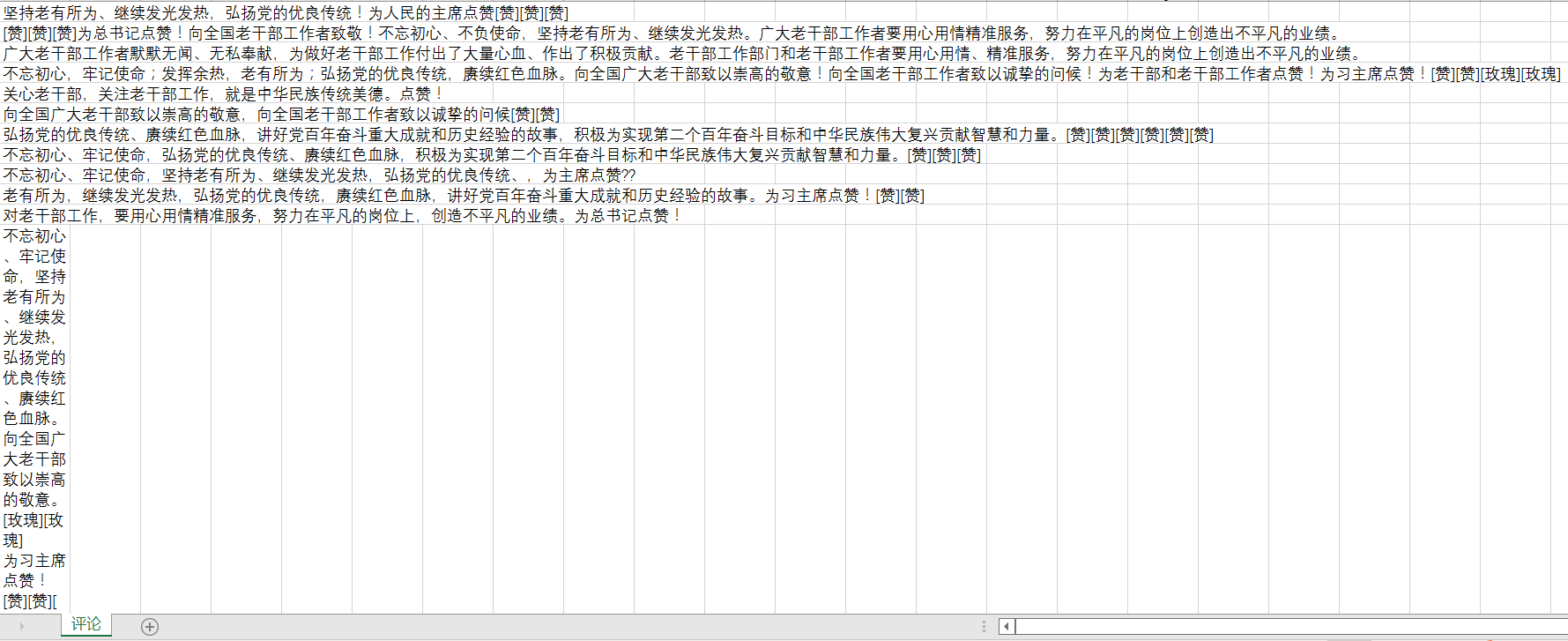
爬虫代码如下：



首先调入所需要的第三方库，函数get\_comment()接受一个json结果的参数作为输入，对数据进行解析后直接保存在本地文件中。主函数中，对url进行请求，并调用get\_comment()函数对数据进行解析。

1. 数据分析

爬取某一篇红色主题的时事新闻的评论信息，内容如下：



可以看到，所有的评论都是非常正向且坚定的，希望中国的未来越来越好。