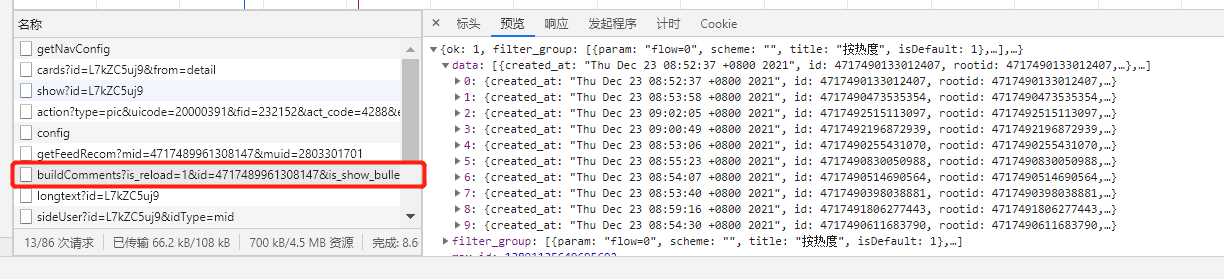
1. 数据爬取

爬取人民日报官方微博的微博评论，人民日报的官方微博url为<https://weibo.com/rmrb>

进入人民日报的官微中，点击某一条微博推送，点击评论，查看全部评论，进入开发者模式中进行查看，可以看到页面的请求如下，通过筛选，可以查找到微博评论的url如下：



该url可以直接请求，返回的是一个json结构的内容。

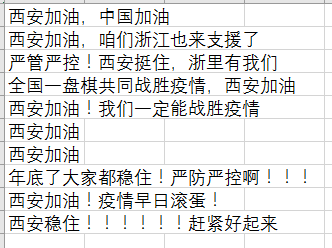
爬虫代码如下：



首先导入需要的第三方库，在主函数中对url进行请求，并得到返回结果。调用get\_data()函数对返回的内容进行解析并保存，在get\_data()函数中，对数据进行解析，由于评论中带有表情符号，因此需要通过分隔符切分进行划分。

1. 数据分析

爬取人民日报2021年12月23日星期四中发布的“【[#12月22日新增本土确诊71例#](https://s.weibo.com/weibo?q=%2312%E6%9C%8822%E6%97%A5%E6%96%B0%E5%A2%9E%E6%9C%AC%E5%9C%9F%E7%A1%AE%E8%AF%8A71%E4%BE%8B%23)，其中[#西安新增本土确诊63例#](https://s.weibo.com/weibo?q=%23%E8%A5%BF%E5%AE%89%E6%96%B0%E5%A2%9E%E6%9C%AC%E5%9C%9F%E7%A1%AE%E8%AF%8A63%E4%BE%8B%23)】”微博评论，获取到的评论如下：



评论全部均为网友对西安疫情的关心，希望西安疫情早日结束。