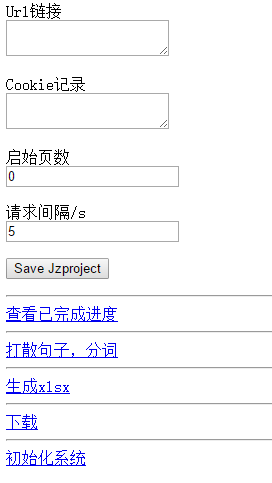
Tips1：使用chrome浏览器

Tips2：最好别用自己的微博账号。。。。不过估计风险不大

1.首先 访问：<http://23.106.131.137:3000/jzproject> （服务器在美国 可能有点慢。。



2. 然后 寻找微博 点击日期 获得url

http://weibo.com/3229125510/EgKzYEtyk?ref=home&rid=2\_0\_8\_3063646978791068240&type=comment#\_rnd1478676471329



3. 将url开头修改成 m.weibo.cn 后面格式无所谓

Before:

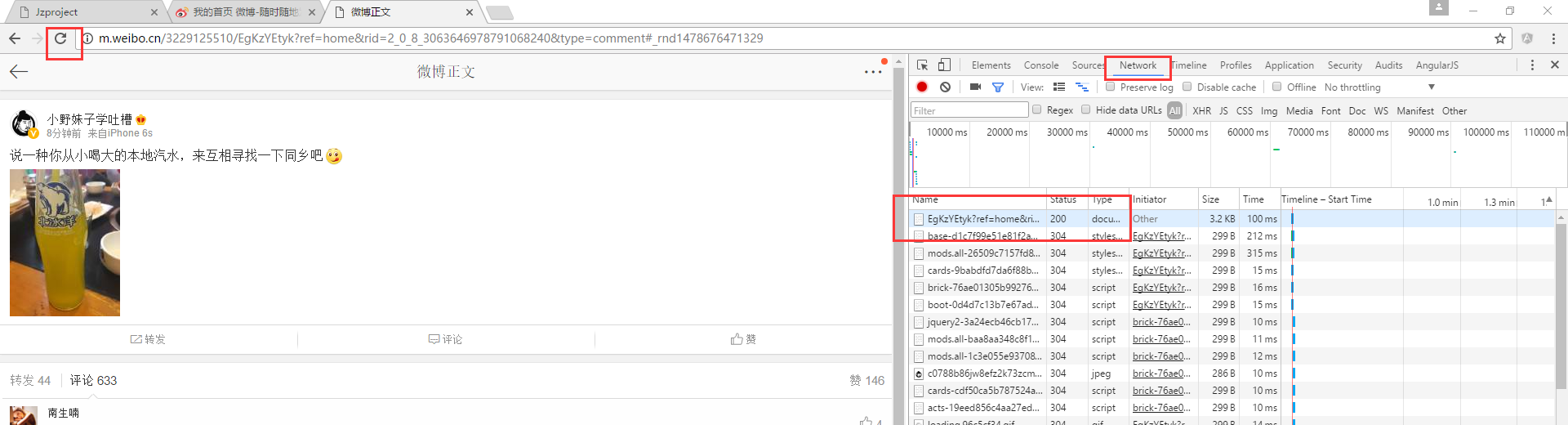


After:

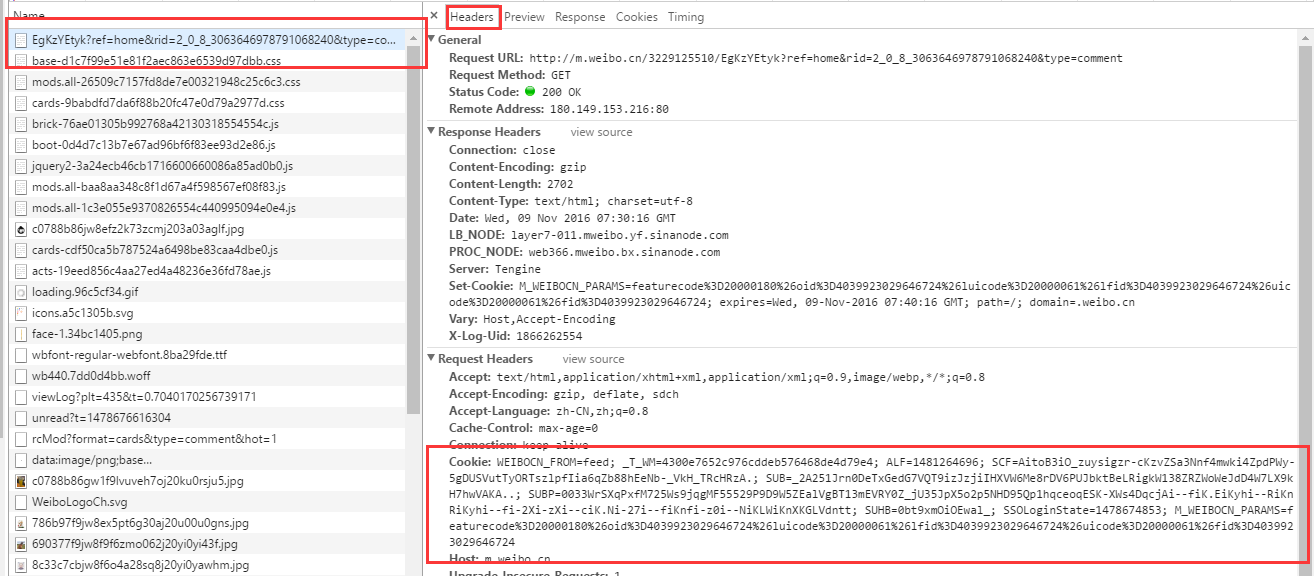


按回车

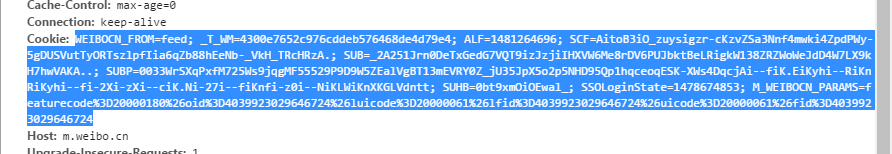
4.按f12 点这个network 然后刷新几下



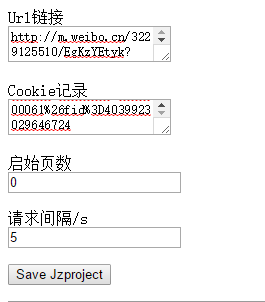
点击network下面第一个选项



把cookie全部内容复制



5.粘贴



这里 url就是之前修改后的url m.weibo.cn这样

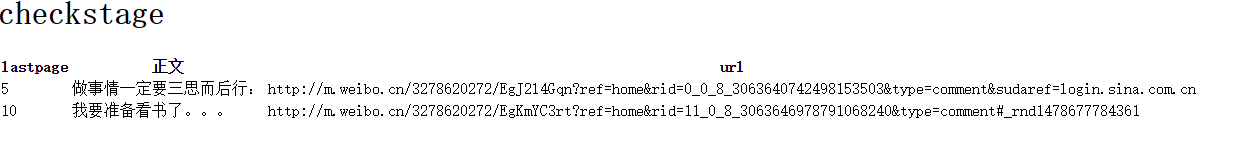
点 save jzproject

点了以后过一会就可以关掉网页了 后台它会自己跑 不过如果不关掉的话完成了就会跳转到完成也没能告诉你完成了。。



请求间隔是每两个评论之间的间隔 5有点长。。不过我之前测试的时候设置的2 偶尔会只获取一半 可以继续

6. 可以点这个查看已完成进度



Lastpage是说翻到了哪一页 哪个中间断掉了的话想继续 改一下那个起始页数就行了~ 改成lastpage的数字（lastpage的编码的那一页未被获取

7. 点



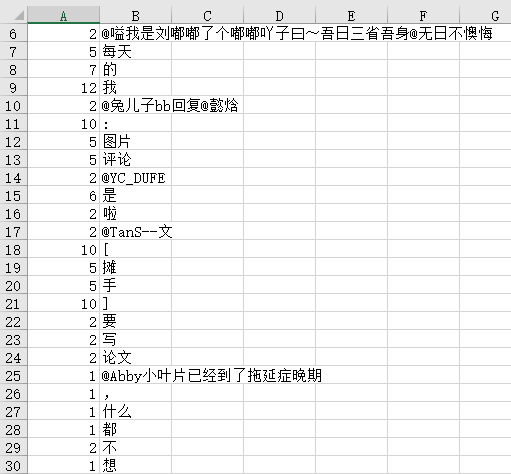
8. 点



9. 点



10. ok啦~



一共有三张表 第一张是完成了哪些评论 评论的正文和url是什么

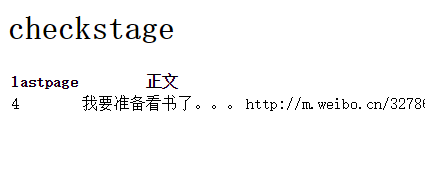
第二张是每条评论 第二张的第一列是点赞数

第三张是词条频率（我没排序。。不同微博的相同词条也没合并计数 0 0

Ps 第三张词频统计有乘以点赞数

# 然后使用步骤可以这样来

填这个 点save jzproject 重复

等到多个checkstage都显示出来了以后 打散句子 生成xlsx

点击初始化系统会清空前面的统计结果

就这样。。。出了bug找我！或者要实现什么找我！

Ps 新加一个初始化分词器的选项 这样你用过一次分词之后 点初始化分词器 就可以继续抓评论 然后再分词 生成xlsx而不用清空之前的抓取结果了

就这样。。。出了bug找我！或者要实现什么找我！