

题目：

在百度贴吧中寻找一个自己喜欢的贴吧，将其中一篇热门帖子的每层楼的发帖人、发帖内容和发帖时间爬取下来。

代码：

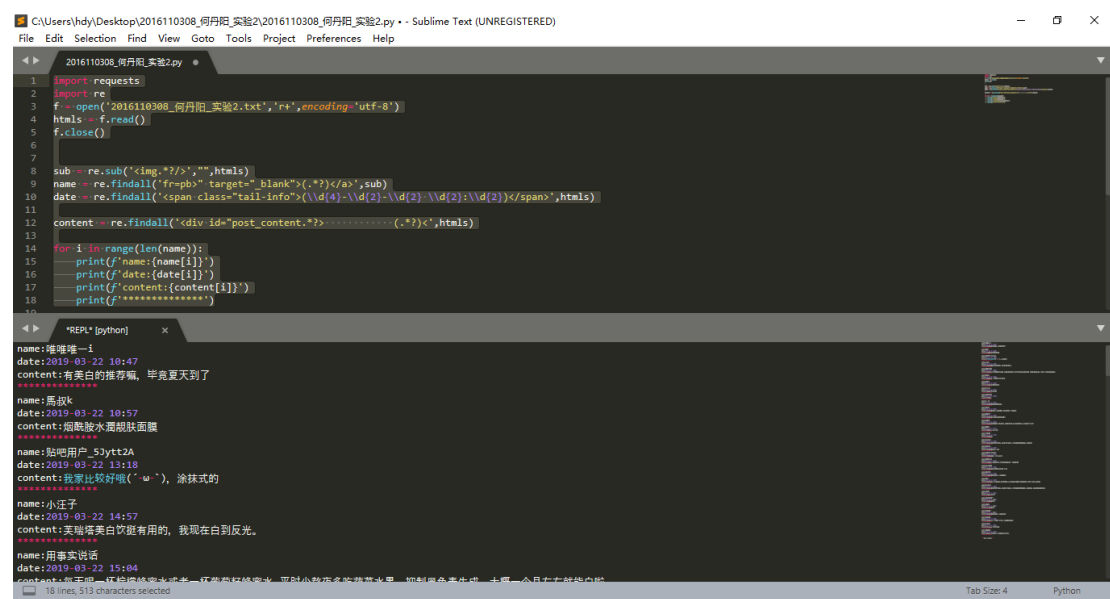
```
import requests
import re
f = open('2016110308_何丹阳_实验 2.txt','r+',encoding='utf-8')
htmls = f.read()
f.close()

sub = re.sub('<img.*?/>','',htmls)
name = re.findall('fr=pb>" target="_blank">(.*)</a>',sub)
date = re.findall('<span class="tail-info">(\d{4}-\d{2}-\d{2})\d{2}:\d{2})</span>',htmls)

content = re.findall('<div id="post_content.*?>(.*)</div>',htmls)

for i in range(len(name)):
    print(f'name: {name[i]}')
    print(f'date: {date[i]}')
    print(f'content: {content[i]}')
    print(f'*****')
```

截图：



The screenshot shows a Sublime Text editor window with a Python script for web scraping. The script imports requests and re, opens a file named '2016110308_何丹阳_实验 2.txt', reads its content, and uses regular expressions to extract post names, dates, and content from HTML. The script then iterates over the extracted data and prints it. Below the editor, the REPL console shows the output of the script, displaying the name, date, and content of several posts from a Baidu Tieba forum.

```
name: 喵喵喵-1
date: 2019-03-22 10:47
content: 喵喵喵的推荐啊，毕竟夏天到了
*****
name: 喵喵喵
date: 2019-03-22 10:57
content: 喵喵喵的水蜜桃面膜
*****
name: 喵喵喵用户_53ytt2A
date: 2019-03-22 13:18
content: 我家比较好看(“u”)，涂沫式的
*****
name: 小丸子
date: 2019-03-22 14:57
content: 喵喵喵美白面膜有用的，我现在白到反光。
*****
name: 喵喵喵说话
date: 2019-03-22 15:04
content: 喵喵喵说话
```

```

*REPL* [python]
date:2019-03-22 23:52
content:吃美白丸
*****
name:Denegeal
date:2019-03-22 23:59
content:楼主私聊我详细了解呀，这些天七天体验，七天就能看到明显效果，效果说话
*****
name:白到发光
date:2019-03-23 07:33
content:美瑞塔美白饮，私我
*****
name:贴吧用户_5KQQRbC
date:2019-03-23 11:16
content:用美白精油 niluxgnahz
*****
name:略略略luel
date:2019-03-23 11:21
content:我用过一款美白产品 平时注意防晒的话 不容易反弹
*****
name:浪人的青春
date:2019-03-23 15:52
content:法国艾璐卡和纽西之谜了解一下呀
*****
name:微梨梦雨
date:2019-03-23 19:05
content:内服胶原蛋白肽片，外注意防晒
*****
name:晚生i
date:2019-03-23 22:17
content:有亮白冻。少女秘密的。我之前用过一盒。最开始不相信。后来真的白了。那个小哥哥人还很好
*****
name:Denegeal
date:2019-03-23 23:49
content:楼主私聊我详细了解呀，这些天七天体验，七天就能看到明显效果，效果说话，私聊给我微信我加你
*****

```

贴吧链接:

<https://tieba.baidu.com/p/6075766796?pn=1>