

1. 预警

脚本的注释我基本都加了，想改随便改，愿意的话可以标一下原作者**洛书**，不想标也问题不大，只要别当面吐槽写得不行就好

1.1. 亿点缺点

有任何一条接受不了就不要用了：

- 需要搭建 Python 环境，如果有任何 Python 相关的问题，请自己百度
- 这个最开始是因为摸了一礼拜鱼没事干，ES 又刚开 8 弹拍立得上头才写的，非摸鱼时间写这个是真的非常缺乏性价比，而且我自己并不当管理，外加上次提问之后到这两天才有新的摸鱼时间，所以后续九成九是不会再改了，有 bug 可以提，但也要等下次有上班摸鱼的空才会改
- 因为我没当过管理也没真的打过表，所以可能有操作上反人类的地方
- 每次执行的时候需要在网页上打开 QQ 群相册，用浏览器的开发者模式查看相关参数，再手动填进配置文件（所以查都查了建议把群里在开的所有谷子都更新一遍）
- 鉴于相册评论和群昵称真的非常自由，基本不可能实现无脑执行，还是需要自己核对结果是否准确的，具体要求参考 6.2
- 如果不是 ES，需要自己准备一份包含所有角色/队伍缩写的角色清单
- 每次开新谷需要准备一份新的模板文件（这个其实手动打表也需要）
- 只能区分单个角色，如果某种谷子只以队伍的形式出，还是会被识别成队伍里的每一个角色
- 只支持一类谷子一张图的开谷形式，不支持识别谷子类型，例如“A 拍立得 5 B 吧唧 10”这样的评论（除非把拍立得和吧唧当做柄类型填进去）
- 柄类型不支持缩写，也不支持多种写法（例：食玩卡的 sp 和特只能选一个）
- 这个文档真的很长我其实觉得没写废话但有 3000 多字

1.2. 功能

抓 QQ 相册评论，生成排表和详细的点数/金额表，可以再用数据透视表生成肾表，分类汇总也行，我排过序了

优点大概是前期准备完成之后，查一次参数就能把群里在开的所有谷都更新完，同时开非常多谷且多数人评论方式正常的情况比较适用，而且手动核对比手动打表的出错率低一点

*另：其实我觉得可以用来抓肾表下面敲 1 的人的名单，但好像没什么意义，对账还是要手动

2. 环境搭建

2.1. Python 安装

参考 <https://www.liaoxuefeng.com/wiki/1016959663602400/1016959856222624>

只要装的是 python3 就行

2.2. 第三方包安装

win+r，输入 cmd，回车，打开命令行界面，输入下面这行，回车

```
pip install openpyxl pandas requests
```

网络正常等一段时间就安装完了，可以在命令行界面输入 python，然后执行下面这行测试，没报错就没问题

3. 参数

修改 config.py 文件

3.1. 只要改一次

path_goods: 角色清单的路径, 把单引号里面的路径替换成自己电脑上的文件路径

myname: 查看相册的时候, 是不会显示自己的群备注的, 只会显示自己 qq 的昵称, 如果两者不一致又想显示备注, 改成{'name': 'QQ 昵称', 'nickname': '想显示在排表里的名字'}

album: QQ 号和群号, 改成{'uin': 'QQ 号', 'hostUin': 'QQ 群号'}

3.2. 谷子类型不同时需要改

*改到 3.2 模板里了

3.3. 每次打开群相册需要改 (时间太久了刷新也要)

脚本包含 3 个部分, 爬虫, 解析评论, 生成排表, 如果不需要用到爬虫的话, 这部分和参数里的 topics 都是可以填的

1) 在浏览器中打开 QQ 群相册

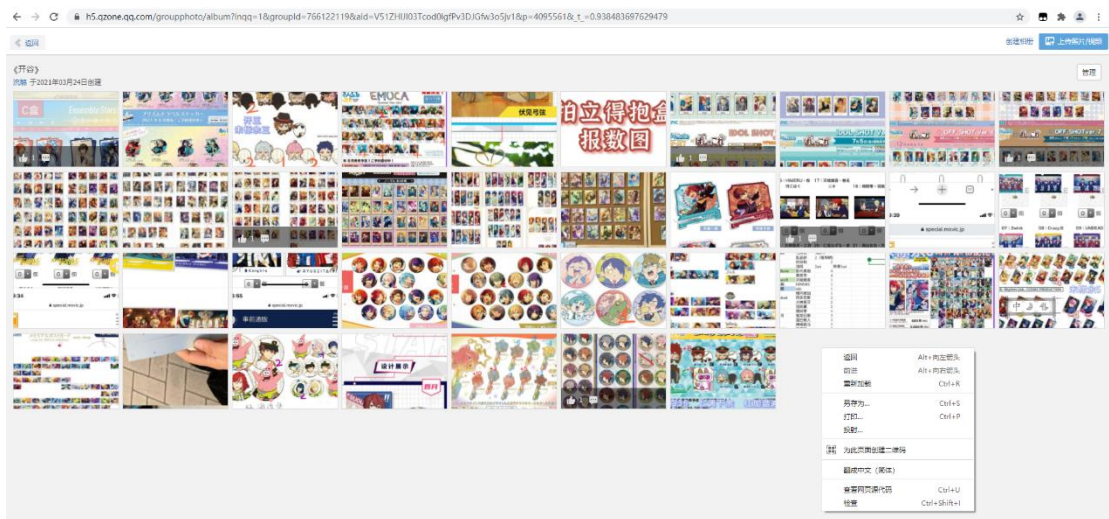


有人在手机评论相册的动态, 电脑端点击可以跳转到浏览器里打开 QQ 群相册 (无论是不是目标相册都可以, 只要是这个群的动态就行)

*注: 这一步不一定能实现, 如果评论了没加载出来, 建议翻翻别人的评论, 如果别人的评论也不行就看历史记录里有没有, 如果历史记录也不行就换个日子跑

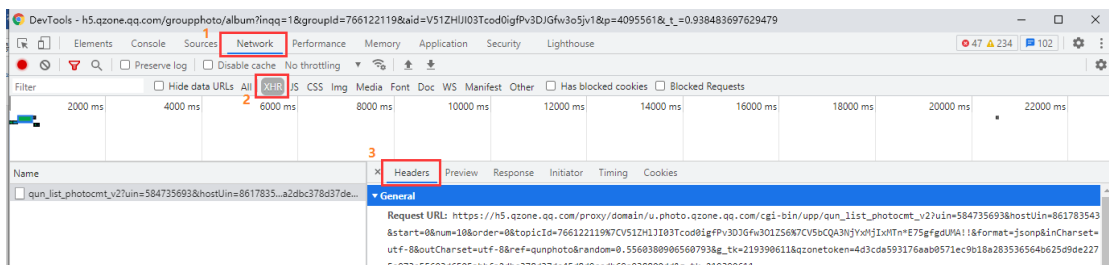
2) 切到对应图片

切到相册, 随便在哪里右键, 点检查, 如果不是 chrome, 百度查一下浏览器的开发者工具怎么打开



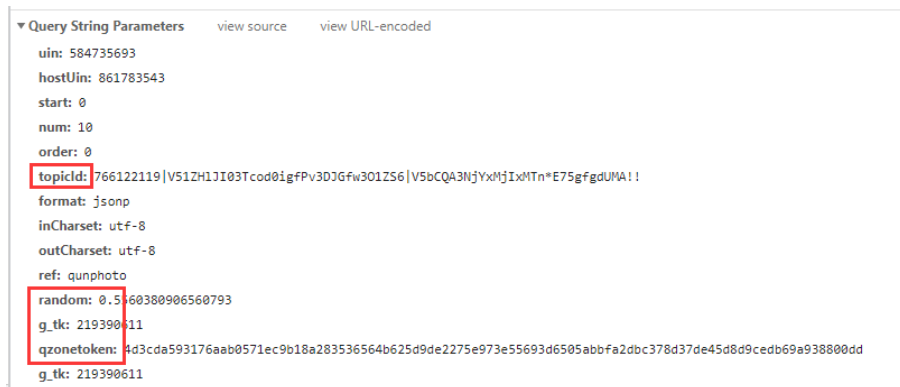
点开相册里的一张图片，然后如下图，在刚刚弹出的页面上切到 Network，再切到 XHR，会有一个开头是 qun_list_photocmt_v2 的请求，查看 Headers 里的内容

*注：如果没有 qun_list_photocmt_v2 的话，点一下弹出的照片评论区的“查看全部 xx 条评论”



3) 查看参数

把上一步 Headers 里的参数，复制到 config.py 里面



```

▼ Request Headers
:authority: h5.qzone.qq.com
:method: GET
:path: /proxy/domain/u.photo.qzone.qq.com/cgi-bin/upp/qun_list_photocmt_v2?uin=584735693&hostUin=861783543&start=0&num=10&order=0&topicId=766122119%7CV51ZHI03Tcod0igfPv3DJGfw301Z56%7CV5bCQA3NjYxMjIxMTn*E75gfgdUMA!!&format=jsonp&inCharset=utf-8&outCharset=utf-8&ref=qunphoto&random=0.5560380906560793&g_tk=219390611&qzonetoken=4d3cda593176aab0571ec9b18a283536564b625d9de2275e973e55693d6505abbfa2dbc378d37de45d8d9cedb69a93880dd&g_tk=219390611
:scheme: https
:accept: */*
:accept-encoding: gzip, deflate, br
:accept-language: zh-CN,zh;q=0.9
:cookie: pgv_pvid=1900968440; RK=1DbEKqXEz; ptcz=3673bf3b210cc0fa15c5bb07d44020d51ffe3ab87d76aedc5b4abf32b86a05cd; tvfe_boss_uuid=c9dcc2104191e416; mobileUV=1_175e5a66015_16288; pgv_pvi=7925114880; eas_sid=K1965153R5I6Q2KH0z0u7u895; ied_qq=o0584735693; o_cookie=584735693; __Q_w_s_hat_seed=1; pac_uid=1_584735693; uin=o0584735693; p_uin=o0584735693; pgv_info=ssid=s3780185346; _qpsvr_localtk=0.721300030540434; skey=@nlojAibFD; pt4_token=33p1XkUUXpL32H3ZSH9xM340d5fmQx5X-o1pvbEAHfs_; p_skey=-1-GwScfTweTifMEFwRHZuL9nX9PgbX4nARtO3qofA8
:referer: https://h5.qzone.qq.com/groupphoto/album?inqq=1&groupId=766122119&aid=V51ZHI03Tcod0igfPv3DJGfw305jv1&p=40955618_t_=0.938483697629479

```

- topics (4.2 里面的参数)
- login = {
 - 'qzonetoken': 'Headers 里 qzonetoken 的值',
 - 'random': 'Headers 里 random 的值',
 - 'g_tk': 'Headers 里 g_tk 的值',
 - 'cookie': 'Headers 里 cookie 的值'

cookie 比较长，替换的时候别把最后面的单引号删了

4. 准备文件

4.1. 角色清单（ES 直接用就行）

需要确认角色清单.xlsx 里面列出了所有此次会提到的角色，sheet 名和列名都不能动。
提供的角色清单应该挺全了，如果有漏的参考已有内容手动补一下

- 名称缩写 sheet 里填角色和队伍的所有缩写
全称别名就是字面意思，脚本执行的时候只用别名来进行匹配，所以别名一定要包含全称的至少一个字，以防有人打了全名识别不了
此外，匹配是按顺序的，如果缩写重复，需要把优先匹配的（字多的）放在最上面
例：游木真如果放在衣更真绪的前面，那么如果有人用“真绪”，只会匹配到“真”；
同理，“兔团”也需要放到仁兔前面
类型就是用来标识角色(role)或者队伍(team)
*愿天下没有相同的缩写，如果有就换个缩写
- 队伍 sheet 里填队伍全称和具体有谁，如果没有队伍，列名和 sheet 保持不变，什么都不填就行

4.2. 模板

每次执行都需要有一个模板文件，用来在模板的基础上生成最后的排表，可以参考已有的三个文件夹里的模板

要求：

- 在脚本所在的文件夹里，创建一个文件夹，里面放模板文件，命名为模板.xlsx，需要

包含排表和参数两个 sheet

排表 sheet 要求：

A	B	C	D	E	F
队伍	角色	柄类型	单价	余量	1
fine	天祥院英智	花前	5.69		
	日日树涉	花前	5.69		
	姬宫桃李	花前	5.69		
	伏见弓弦	花前	5.69		
rabbits	仁兔成鸣	花前	5.69		
	真白友也	花前	5.69		
	紫之创	花前	5.69		
	天满光	花前	5.69		

- 必须包含角色、柄类型、单价、余量这 4 列，生成的排表会直接写到余量的后面一列
*注：单价可以不填或者瞎填，但是单价这一列必须保留；模板里的[队伍]这列只是为了方便找人，不是必需的
- 必须包含[柄类型]这列，如果不区分柄就全空着；**如果区分柄就要全填**，无论是不是只有一个特典
- **柄类型里的每一个字都不能包含角色缩写，且不能包含阿拉伯数字
例：ES 花札，如果用柄名区分，“忍忍者 1”会被识别成[忍忍][者][1]，无效；
如果用月份区分，“三毛缟斑 1”会被识别成[三][毛缟斑][1]，其实还是能识别的但是最好不要这样填

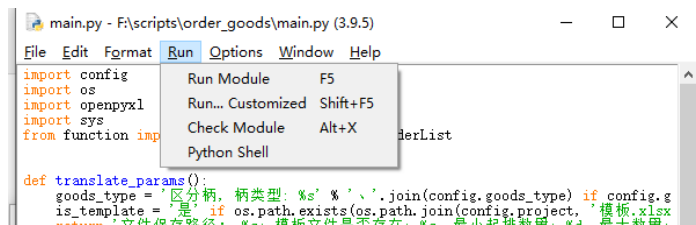
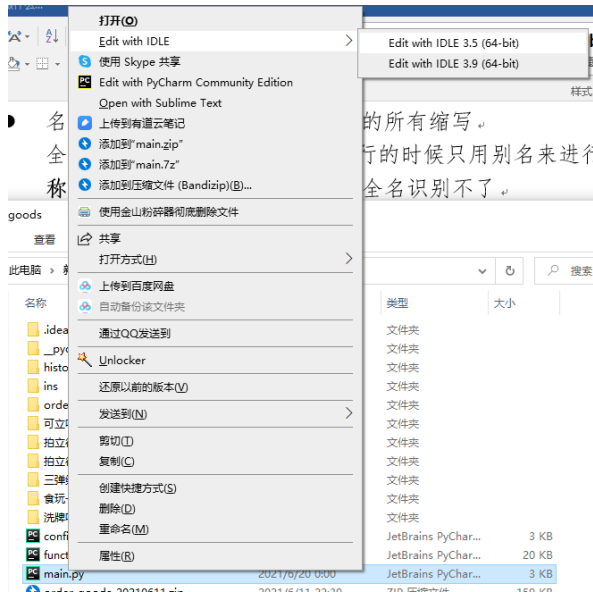
参数 sheet 要求：

参数名	参数值
最小起排数量	1
最大配比数	10
topicId	766122119 V51ZH1JI03Tcod0igfPv3DJGfw301ZS6 V5bCQA3NjYxMjIxMTluY7BgQ1aNFg!!

- 最小起排数量：字面意思
- 最大配比数：字面意思
- topicId: QQ 相册用来标识不同的照片，是不会变的，一次改完再也不用动了，怎么查见 [3.3](#)

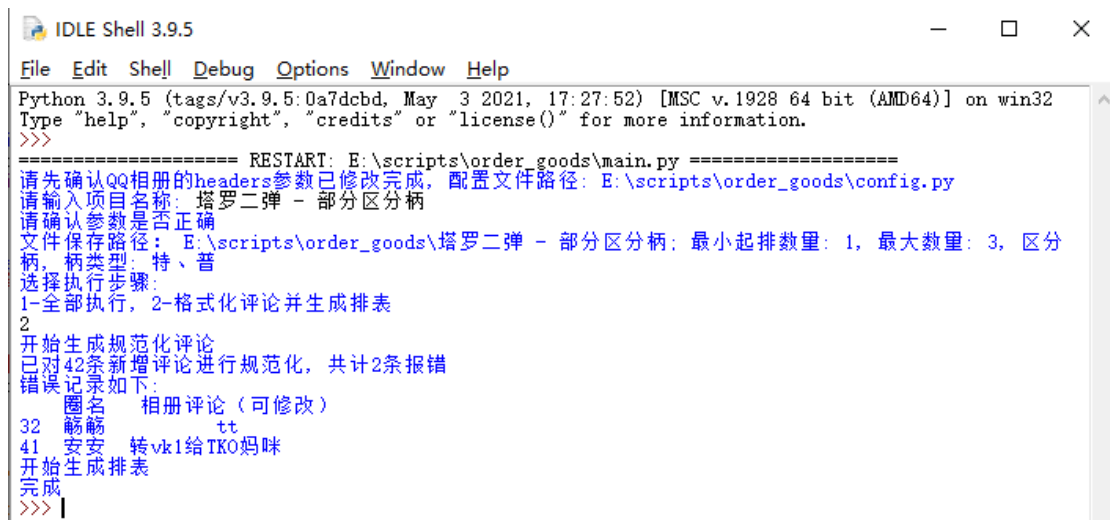
5. 执行

右键 main.py，如下，然后按 F5 执行



参数正确的话就会像这样，需要输入两个参数

- 1 项目名称：填 4.2 里创建的文件夹名称
 - 2 执行步骤：正常的流程是 a. 抓取 QQ 相册的评论，b. 格式化评论，c. 生成排表
- 输入 1 会全部执行，2 执行 b, c, 3 只执行 c



*注：会输出无法解析的评论，建议不要太相信只有这点错，手动检查一下。

《相册评论.xlsx》里有两列[相册评论（请勿修改）]和[相册评论（可修改）]，字面意思，第一个不能改，第二个用来手动修改一些格式无法支持的评论

[评论规范化]这列，在 Excel 里点自动换行，可以核对一下评论和转换的结果是不是对得

上，对不上就手动改一下[相册评论（可修改）]，然后重跑脚本，选 2
具体需要手动修改哪些，可以看 6.2

6. 逻辑

正常的流程是 a. 抓取 QQ 相册的评论，b. 格式化评论，c. 生成排表
a 和 c 都是完全不用管的，a 报错的话最大的可能是参数填错了，类似于页面挂太久了，cookie 失效了之类的，刷新一下页面替换参数看看；c 假如报错了八成是脚本有 bug

6.1. 抓取评论（这块可以不看）

假如文件夹下面没有《相册评论.xlsx》这个文件，就抓取目标相册里的所有评论，生成相册评论.xlsx 文件；
如果存在，就在相册评论.xlsx 的最后一行下，添加新增的评论（用来匹配的是最后一行，如果没有匹配到的话，会把所有新评论写进去之后再删除重复记录）

6.2. 格式化评论

逻辑是以数字为分隔，再以柄做分隔，如果区分柄但未匹配到柄，则记为该角色所有柄，写到《相册评论.xlsx》的[评论规范化]列里面
[评论规范化]列里有换行符，直接点 excel 的自动换行就行
例：

相册评论（请勿修改）	相册评论（可修改）	评论规范化
2vk菓对走1 2千凹	vk菓对走1 千凹	期间凛月,a11,1; valkyrie,a11,1 守泽千秋,a11,10
2英花前10花后4 2纯花前凹	英花前10英花后4 纯花前凹	天祥院英智,花前,10; 天祥院英智,花后,4 涟纯,花前,10
		大神晃牙,花前,1; 乱凪砂,a11,1; 期间零,花后,5; 日日树涉,花后,5; 鸣上岚,花后,1; 朱樱司,花后,1
2花前晃牙1 ngs1 花后零5涉5岚1司1	花前晃牙1 ngs1 零花后5涉花后5岚花后1司花后1	姬宫桃李,花前,1; 期间凛月,花前,1; 月永雷欧,花后,1; 期间凛月,花后,1; 朱樱司,花后,1; 逆先夏目,花后,1
2花前：桃1菓子1 花后：leo1菓子1司1夏目1 2敬人花前1	：桃花前1菓子花前1 ：leo花后1菓子花后1司花后1夏目花后1 敬人花前1	莲巳敬人,花前,1 期间凛月,a11,1; 鸣上岚,a11,1; 衣更真绪,a11,1
2毛栗岚各1	毛栗岚各1	

*手动修改内容（如果愿意可以点进附赠的三份文件夹里的相册评论看看）：

- 把中文数字手动改成阿拉伯数字，因为有人名字带数字所以没加替换功能
- 尽量保证数字在角色后面，在前面也不一定有问题，但是放后面一定没问题
例：加排 1 薰，2 零，会被拆成[加排 1][薰，2][零]
- 如果区分柄，尽量保证柄类型在角色后面，在前面也不一定有问题，但是放后面一定没问题，同上；
- 区分柄的情况，先以数字分隔，再分隔柄和角色
例：薰花前 2 花后 1，这种是无法识别的，需要手动改成薰花前 2 薰花后 1
- 区分柄的情况，柄类型一定要填完整
例：薰花前后各 2，会被识别成薰花前 2，请手动把花后的花字补上去，或者干脆改成薰 2

- 撒排需要以“撒”字开头，而且必须写数量
如果不清楚自己排了多少可以直接写“撒某人 all/凹”，只评论“撒某人”会无法识别
- 转单必须以“转”字开头，仅支持用转 xx 给 xx 的格式，且接转单人必须以“接 xx”开头评论
例：A 评论“转薰 2 给 B”，B 评论“接 A”，就能被识别
- 具体角色和数量只以 A 评论内容为准，不会判断 AB 评论的数量是否匹配
- 转和接的评论一一匹配，请勿转 2 次接 1 次谢谢合作

***如果圈名识别有问题：**

- 如果群昵称没有分隔符，让他改
- 如果有分隔符，建议改个通用的识别符，实在不行再考虑改代码，第 261 行，把新的分隔符加进去

```

257 def _normalize_user(self, string):
258     # 规范化用户名称，把自己的id替换成圈名
259     if string == self._myname['name']:
260         return self._myname['nickname']
261     return re.split('[- ( (?![-_.,!]', string)[0]

```

6.3. 生成排表

有两个 sheet，排表和个人

- 排表

正常的生成排表，假如有最小起排数量，比如 10 起排，那一个人排了 10 就只占一格；排表是从[余量]右侧第一列开始写

队伍	角色	柄类型	单价	余量	1	2	3
fine	天祥院英智	普	10.5	0	雪雪	雪雪	雪雪
	日日树涉	普	12.5	0	雪雪	雪雪	雪雪
	姬宫桃李	普	9.5	0	雪雪	雪雪	雪雪
	伏见弓弦	普	6.5	0	雪雪	涂涂	涂涂
rabbits	仁兔成鸣	普	9.5	2	尝尝		
	真白友也	普	6.5	3			
	紫之创	普	9.5	1	安安	临临	
	天满光	普	6.5	2	安安		
switch	逆先夏目	普	9.5	0	冷冷	冷冷	菲菲
	春川宙	普	5.5	2	维维		

- 个人

根据圈名、角色、柄类型进行汇总，如果想要肾表，用 Excel 自带的数据透视表即可

	A	B	C
1	行标签 ▾	求和项:数量	求和项:总价
2	allall	1	9.5
3	LLL	2	21
4	安安	8	85
5	白简	1	11.5
6	尝尝	1	9.5
7	澈澈	5	120
8	菲菲	4	44
9	临临	1	9.5
10	冷冷	2	19
11	面面	8	69
12	暮暮	4	44
13	太太	2	19
14	桃桃	7	175
15	涂涂	4	28
16	维维	4	47