Twitter zzh-gj-jz association

30 May, 2022

In this analysis I will first extract tweets mentioning "#zhangzhehan", "#gongjun" or "#junzhe" and then check the overlap between accounts tweeting about these topics.

Edit 30/05/22: Added "zhang zhehan" and "gong jun" back to search keywords.

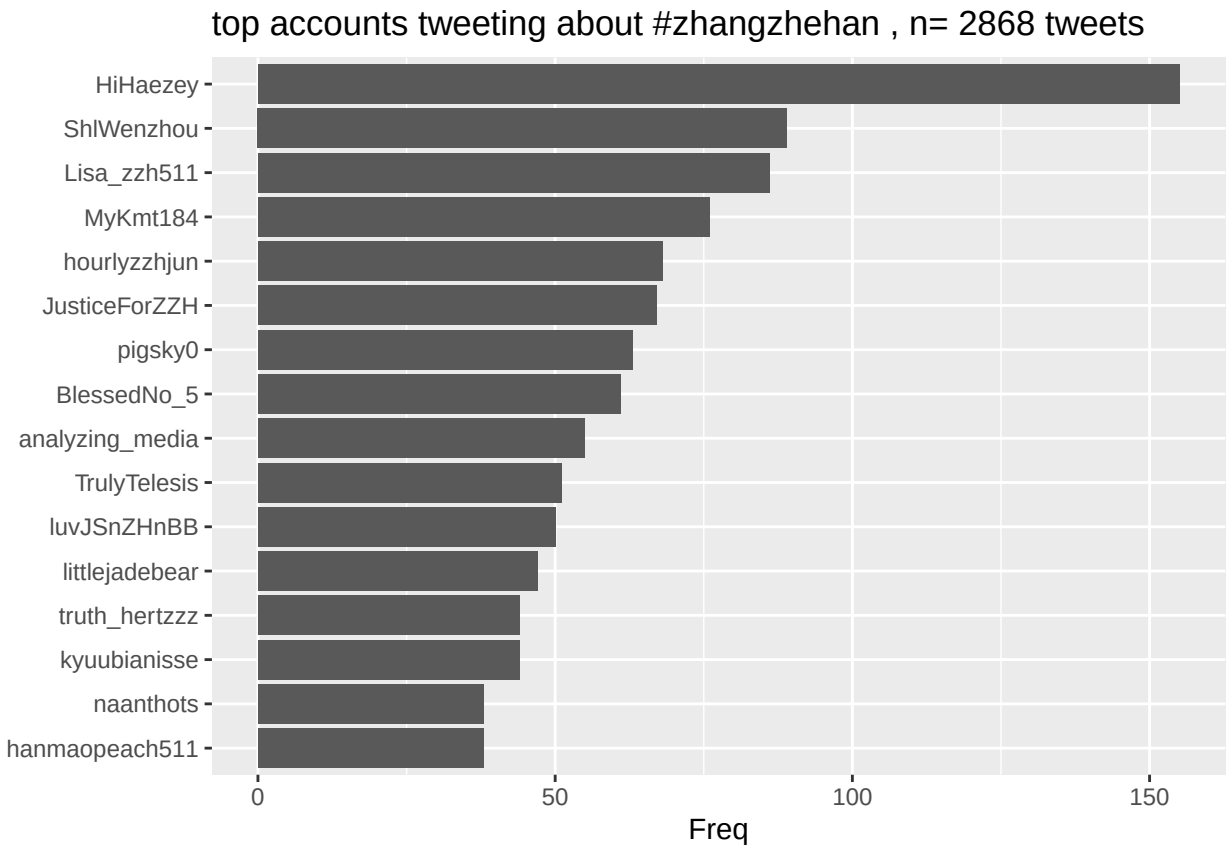
After playing around a bit I decided to use only hashtags instead of full names to focus on dedicated tweets instead of casual mentions.

First let's get the tweets via Twitter API. Please note that there is a rate limit of up to 18000 tweets per day. Nonetheless, I never manage to get the maximum of 6000 tweets per query for some reason. Therefore this search is limited to the last couple of days.

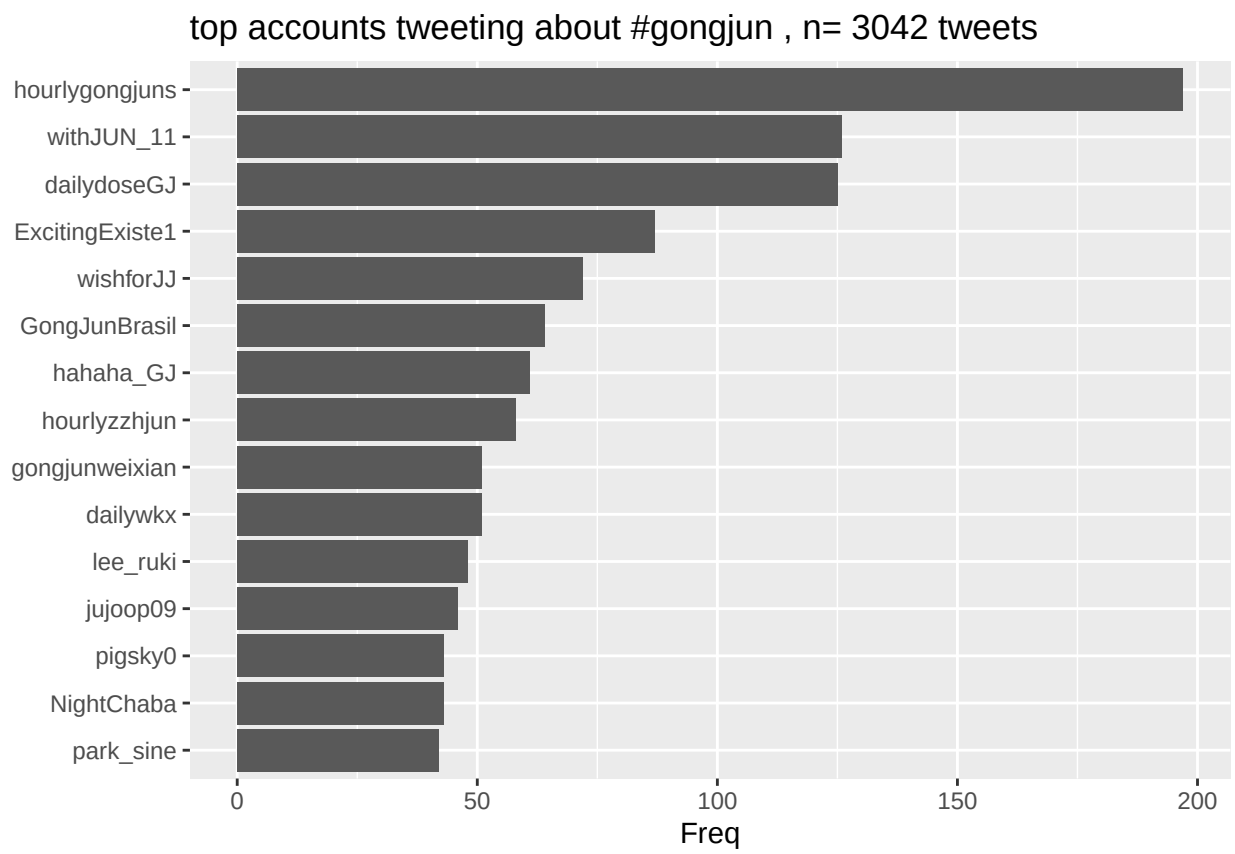
I will then plot the accounts with the most tweets about these topics.

Note: I previously fltered for tweets in English language, now no more language filter applied.

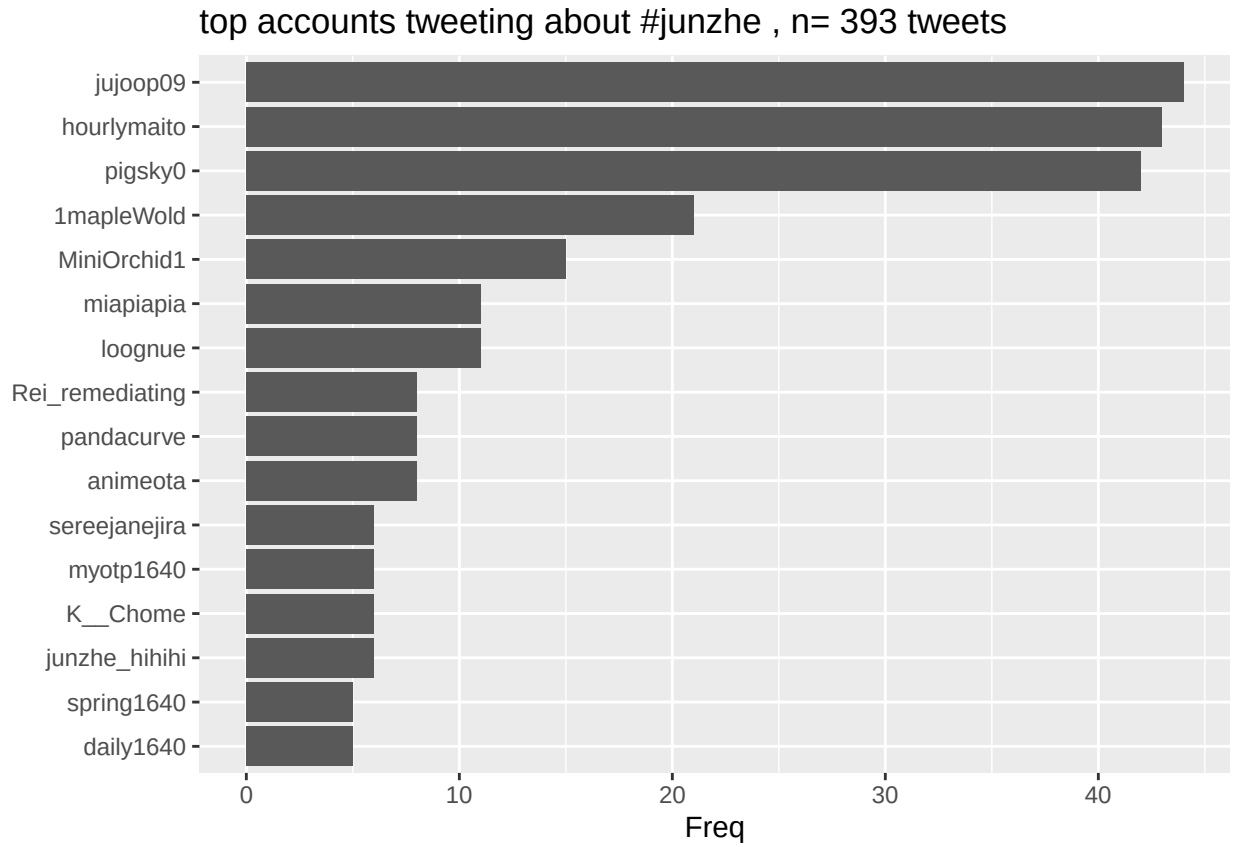
```
## Selecting by Freq
```



```
## Selecting by Freq
```

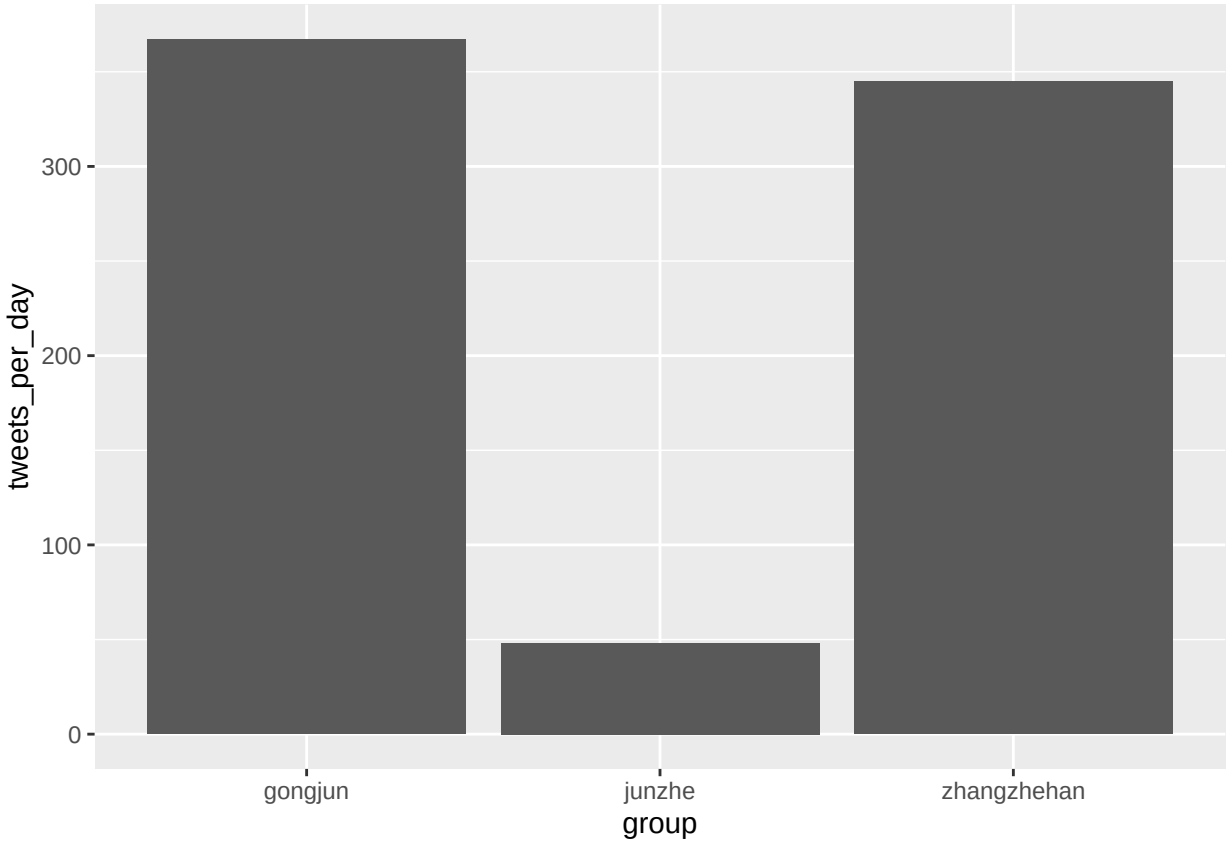



```
## Selecting by Freq
```



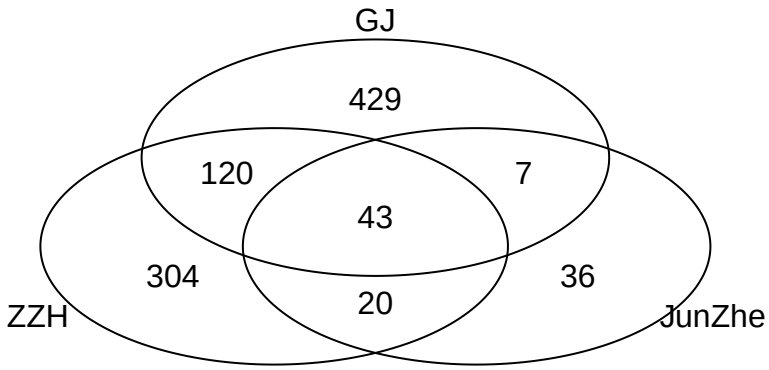
Quick info on which timespan the datasets cover and how many tweets per day

```
## [1] "dataset: #zhangzhehan"
## [1] "number of tweets: 2868"
## [1] "start: 2022-05-22 12:33:56"
## [1] "end: 2022-05-30 20:03:20"
## Time difference of 8.312083 days
## [1] "tweets per day: 345.04"

## [1] "dataset: #gongjun"
## [1] "number of tweets: 3042"
## [1] "start: 2022-05-22 13:11:58"
## [1] "end: 2022-05-30 20:03:14"
## Time difference of 8.285602 days
## [1] "tweets per day: 367.14"

## [1] "dataset: #junzhe"
## [1] "number of tweets: 393"
## [1] "start: 2022-05-22 15:26:26"
## [1] "end: 2022-05-30 19:31:54"
## Time difference of 8.170463 days
## [1] "tweets per day: 48.1"
```



Next let's make a Venn diagram to see the overlap between accounts tweeting about #junzhe, #zhangzhehan or #gongjun. I have to use 2 Venn diagram packages because gplots gives me the numbers+intersections but only venneuler lets me make a beautiful diagram with proportional circle areas.

```
##
## Attaching package: 'gplots'

## The following object is masked from 'package:stats':
##
##     lowess
```

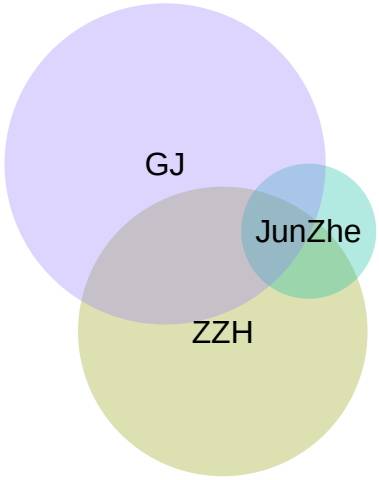



```
## [1] 487
```

```
## [1] 106
```

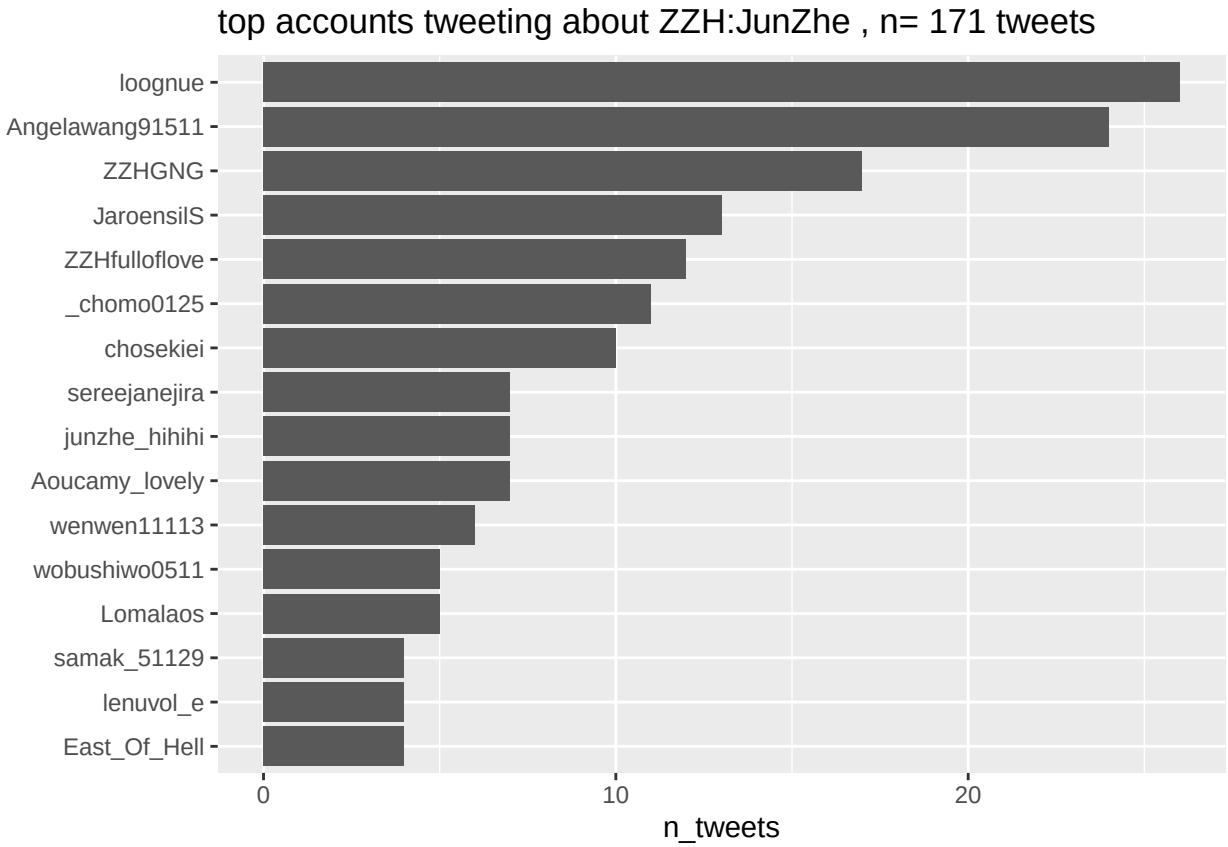
```
## [1] 599
```

```
## Loading required package: rJava
```

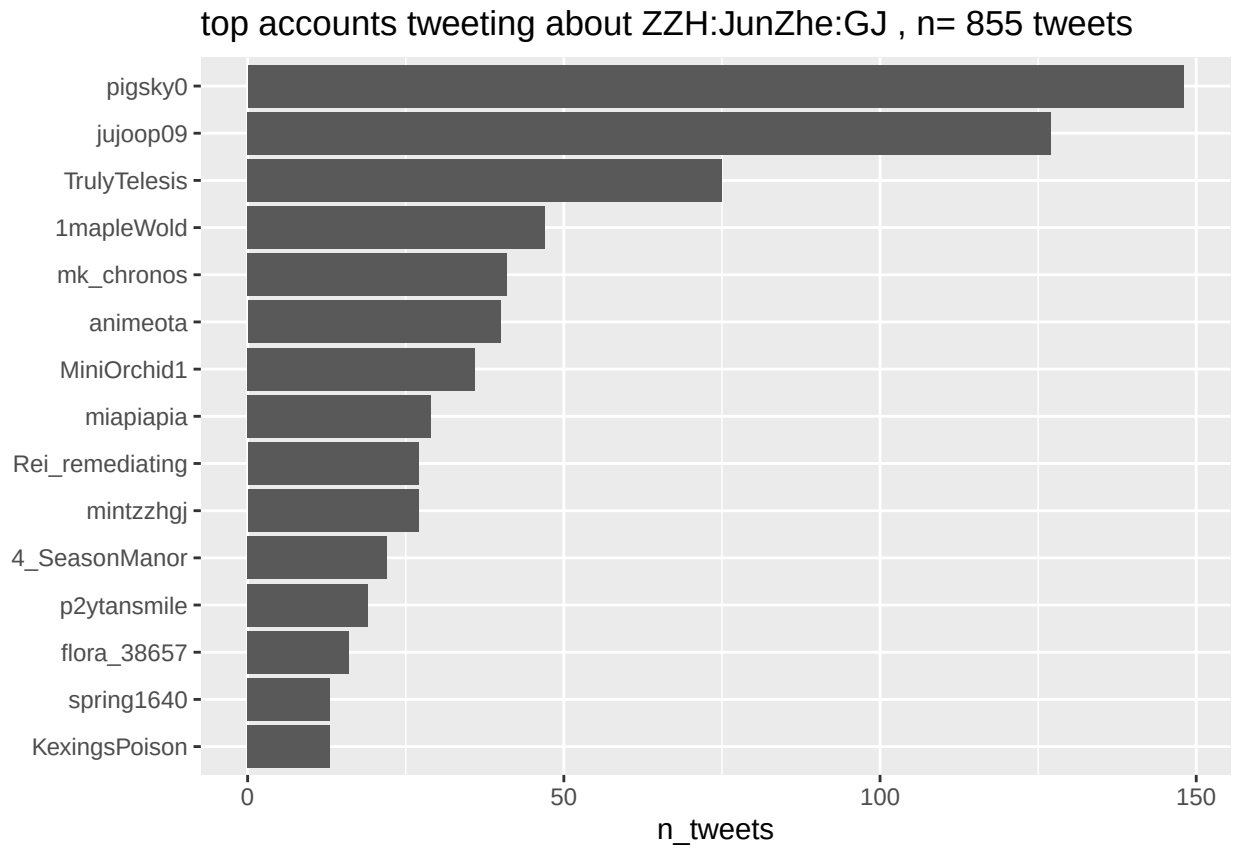


This Venn diagram only shows numbers of accounts tweeting about the three hashtags However I also want to see how many tweets are coming from each section of the Venn diagram.

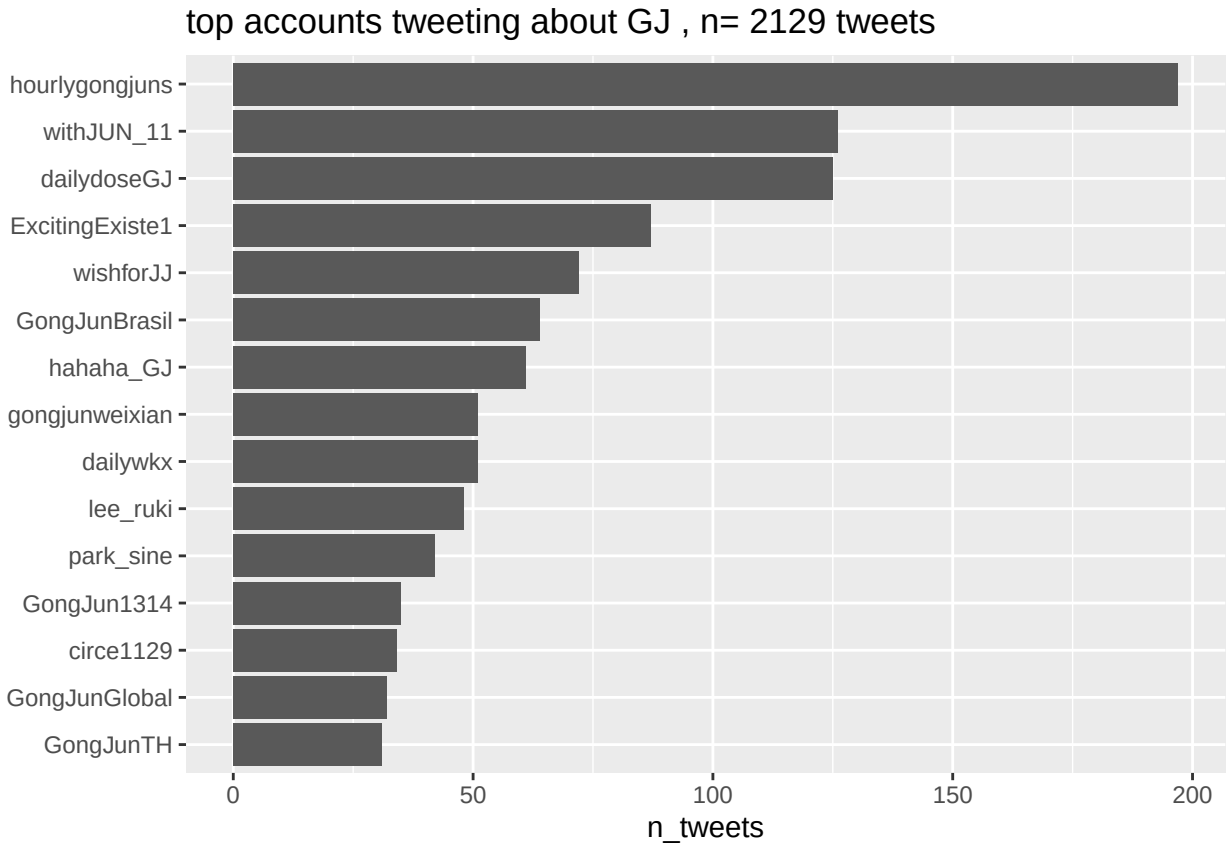
```
## Selecting by n_tweets
## Selecting by n_tweets
## Selecting by n_tweets
## Selecting by n_tweets
## Selecting by n_tweets
## Selecting by n_tweets
## Selecting by n_tweets

## [[1]]
```


top accounts tweeting about ZZH:JunZhe , n= 171 tweets
loognue
Angelawang91511
ZZHGNG
JaroensilS
ZZHfulloflove
_chomo0125
chosekiei
sereejanejira
junzhe_hihihi
Aoucamy_lovely
wenwen11113
wobushiwo0511
Lomalaos
samak_51129
lenuvol_e
East_Of_Hell
0
10
20
n_tweets

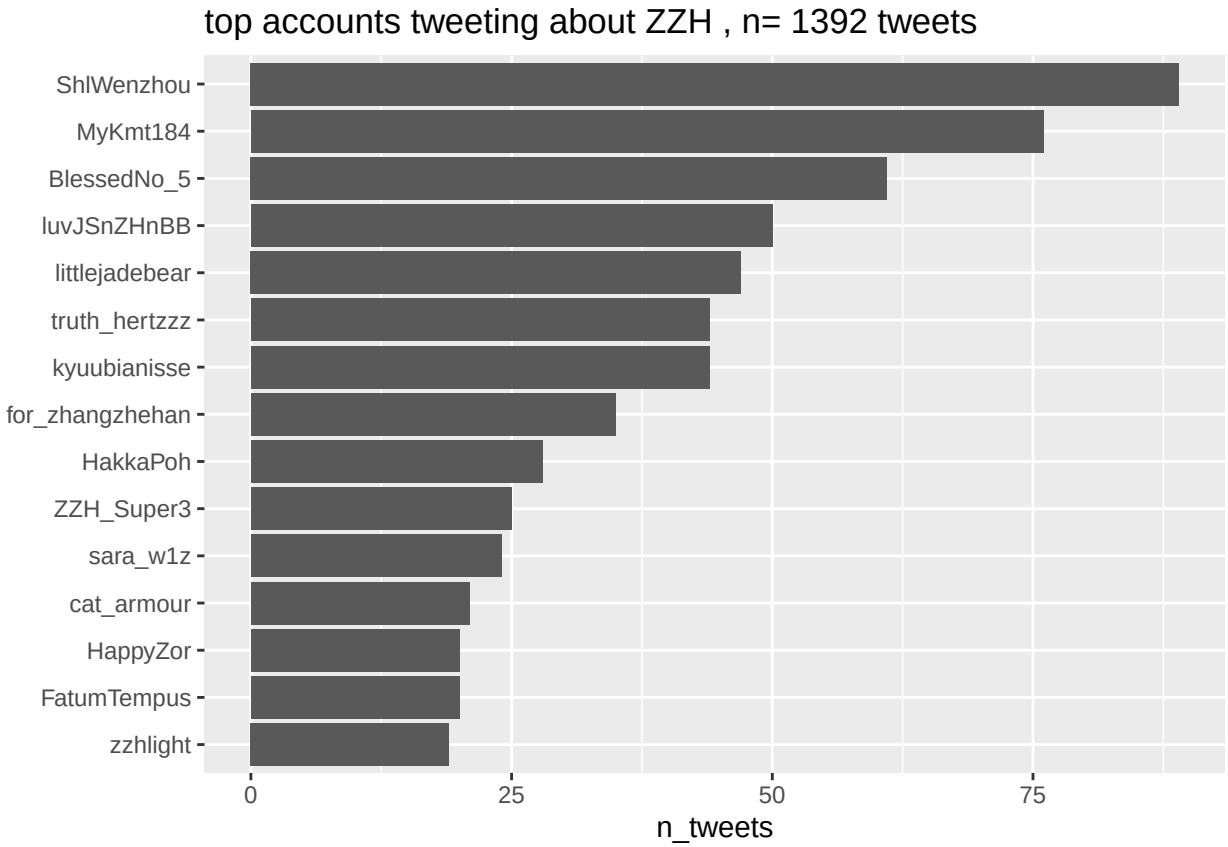
```
##
## [[2]]
```


top accounts tweeting about ZZH:JunZhe:GJ , n= 855 tweets
pigsky0
jujoop09
TrulyTelesis
1mapleWold
mk_chronos
animeota
MiniOrchid1
miapiapia
Rei_remediating
mintzzhgj
4_SeasonManor
p2ytansmile
flora_38657
spring1640
KexingsPoison
0
50
100
150
n_tweets

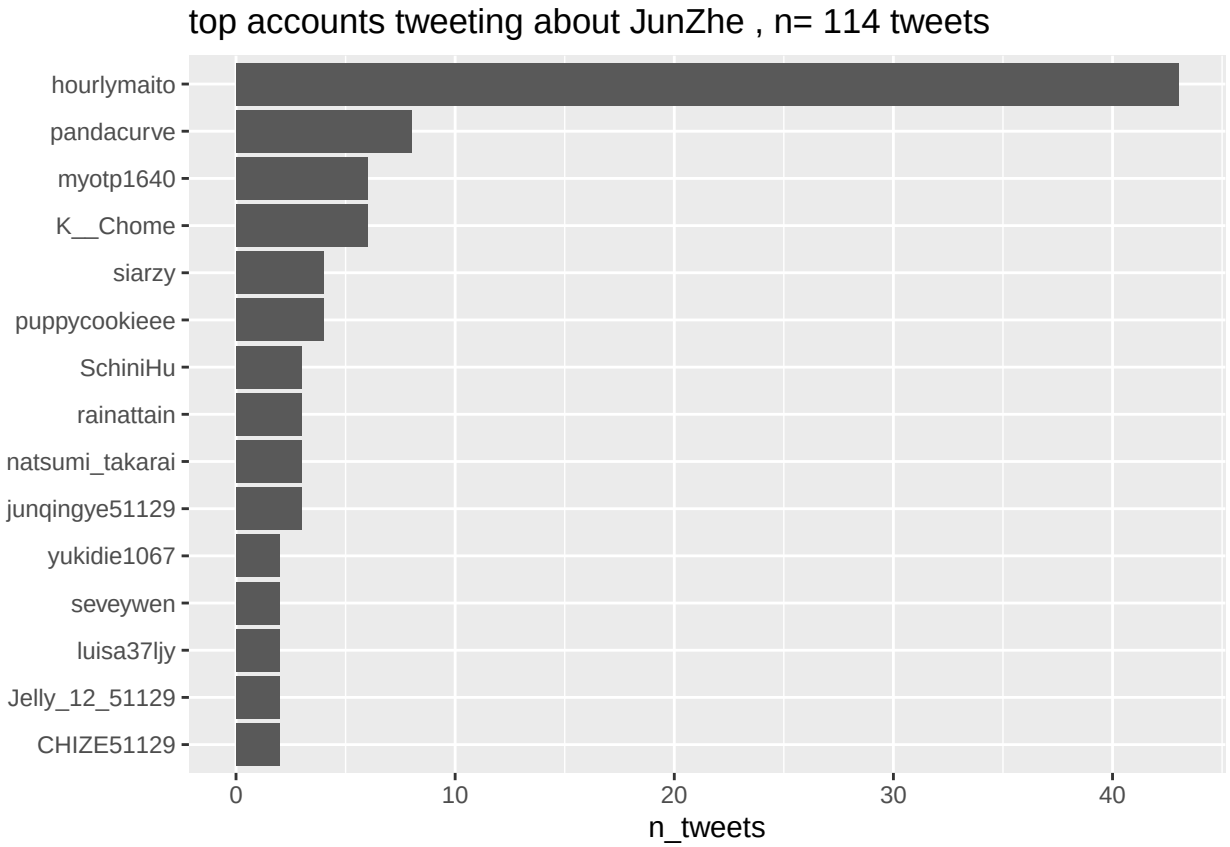
```
##
## [[3]]
```


top accounts tweeting about GJ , n= 2129 tweets
hourlygongjuns
withJUN_11
dailydoseGJ
ExcitingExiste1
wishforJJ
GongJunBrasil
hahaha_GJ
gongjunweixian
dailywkx
lee_ruki
park_sine
GongJun1314
circe1129
GongJunGlobal
GongJunTH
0
50
100
150
200
n_tweets

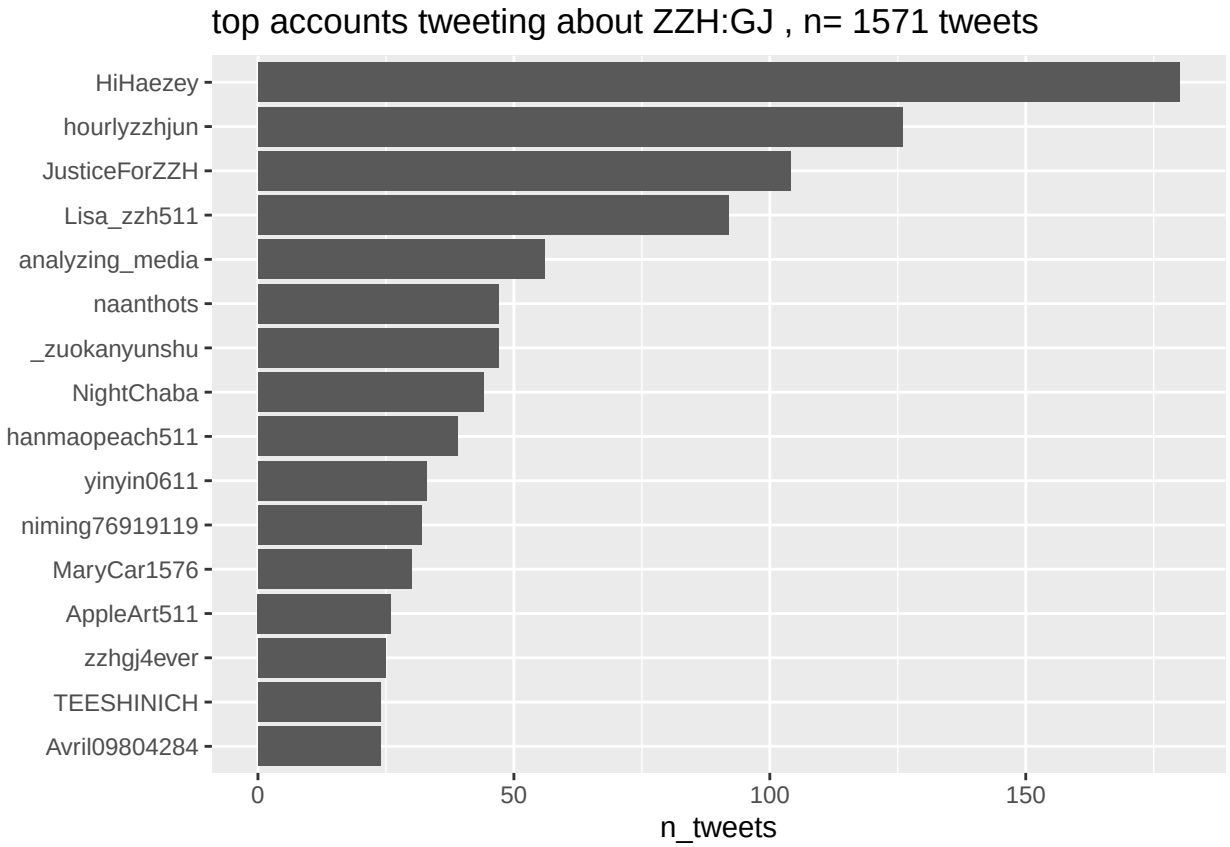
```
##
## [[4]]
```


top accounts tweeting about ZZH , n= 1392 tweets
ShlWenzhou
MyKmt184
BlessedNo_5
luvJSnZHnBB
littlejadebear
truth_hertzzz
kyuubianisse
for_zhangzhehan
HakkaPoh
ZZH_Super3
sara_w1z
cat_armour
HappyZor
FatumTempus
zzhlight
0
25
50
75
n_tweets

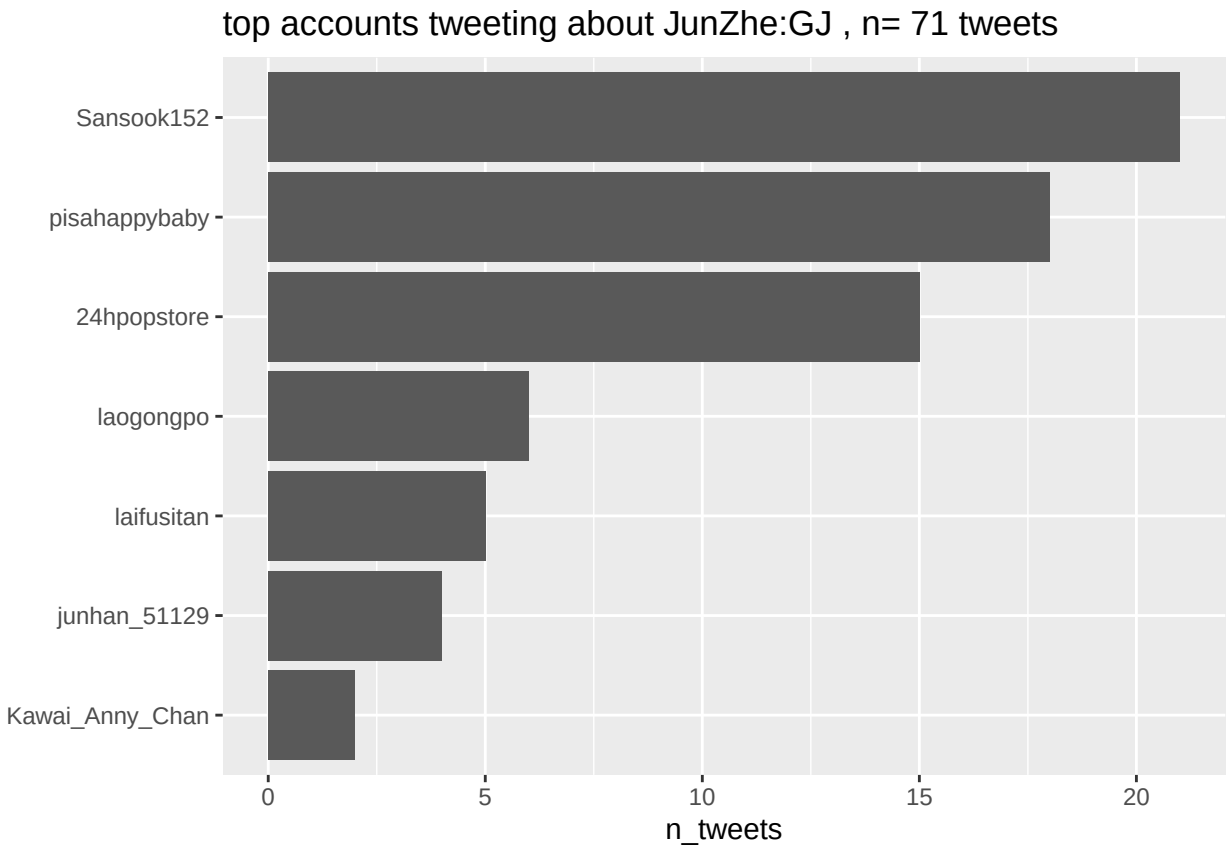
```
##
## [[5]]
```


top accounts tweeting about JunZhe , n= 114 tweets
hourlymaito
pandacurve
myotp1640
K__Chome
siarzy
puppycookieee
SchiniHu
rainattain
natsumi_takarai
junqingye51129
yukidie1067
seveywen
luisa37ljy
Jelly_12_51129
CHIZE51129
0
10
20
30
40
n_tweets

```
##
## [[6]]
```


top accounts tweeting about ZZH:GJ , n= 1571 tweets
HiHaezey
hourlyzzhjun
JusticeForZZH
Lisa_zzh511
analyzing_media
naanthots
_zuokanyunshu
NightChaba
hanmaopeach511
yinyin0611
niming76919119
MaryCar1576
AppleArt511
zzhgj4ever
TEESHINICH
Avril09804284
0
50
100
150
n_tweets

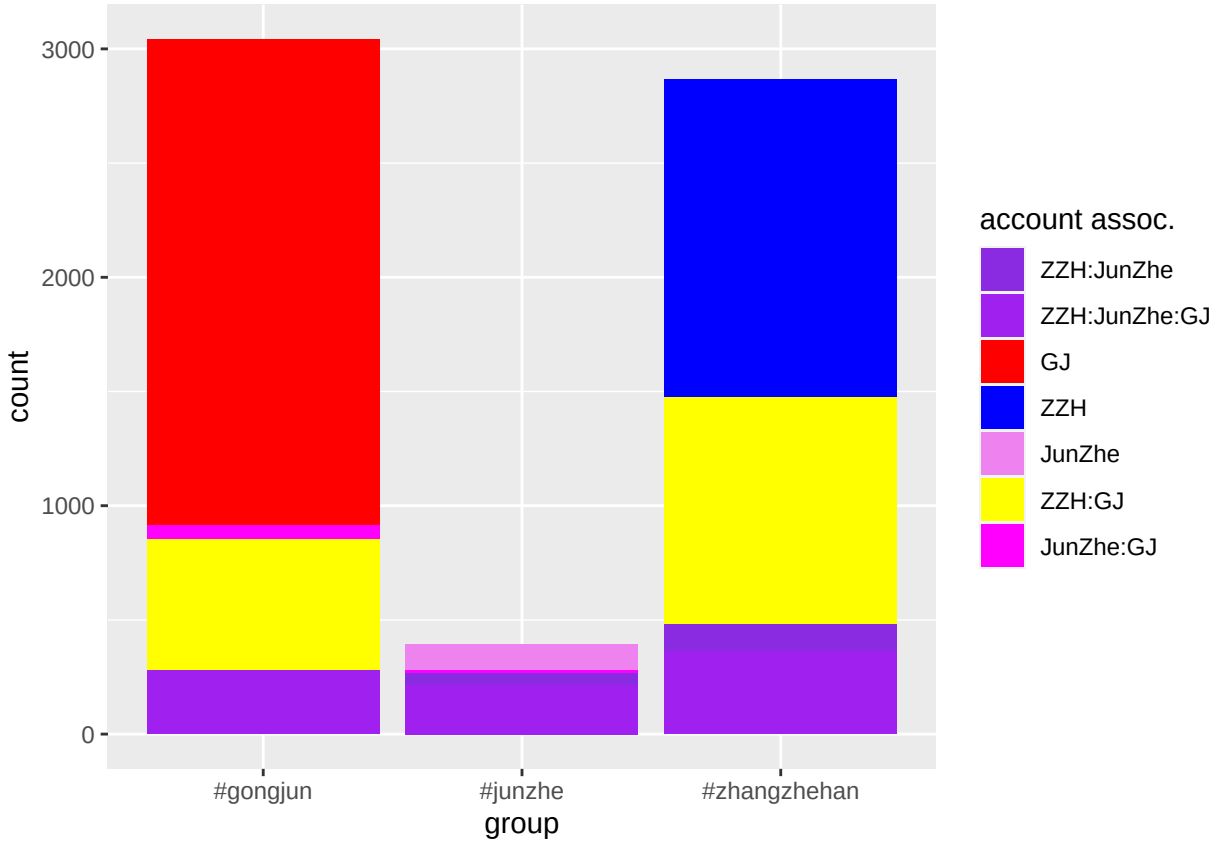
```
##
## [[7]]
```



Now I have information on which accounts are also tweeting about the other two hashtags I looked at. Will now add this information to bar plot showing tweet counts per hashtag

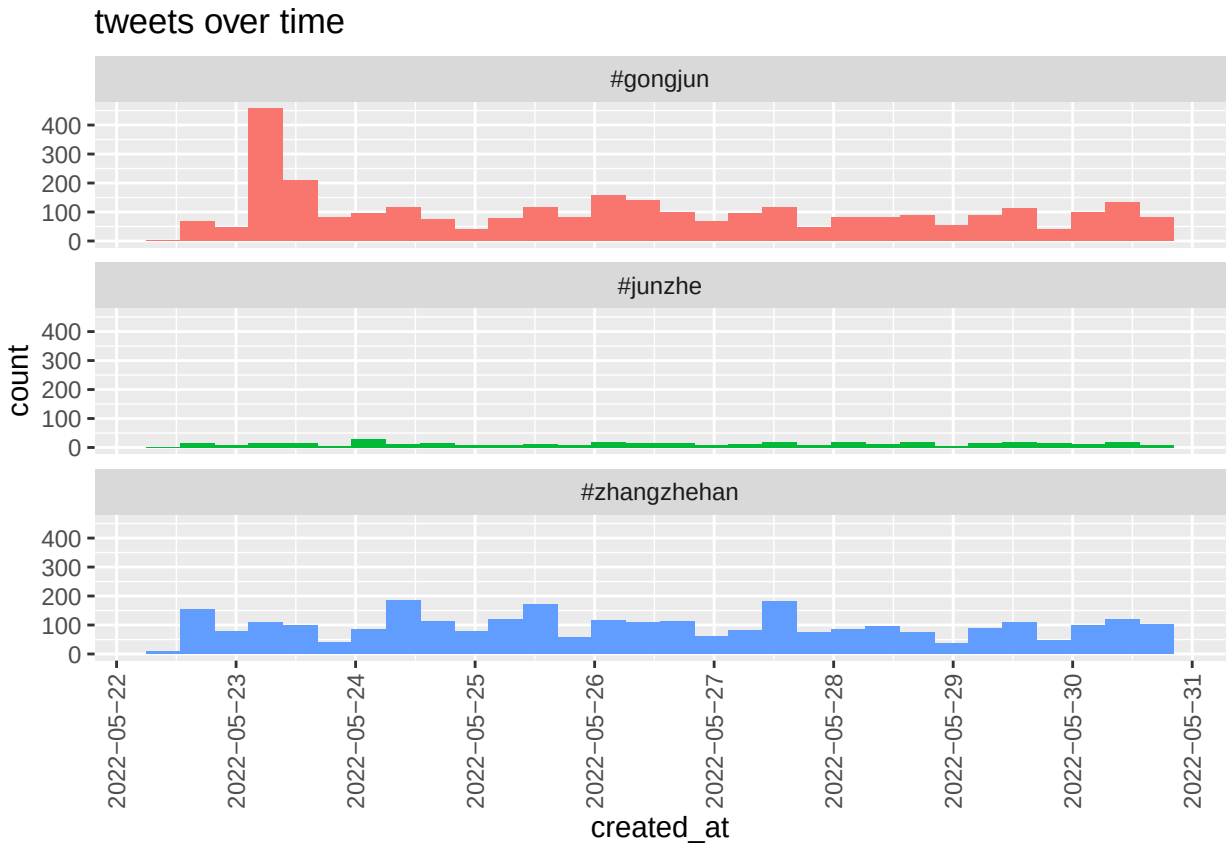
```
## [1] "percent of ZZH tweets from 'pure' ZZH accounts"

## [1] 0.4853556

## [1] "percent of ZZH tweets from GJ-mentioning accounts"

## [1] 0.3469317

## [1] "percent of ZZH tweets from CP-related accounts"

## [1] 0.1677127

## [1] "percent of GJ tweets from 'pure' GJ accounts"

## [1] 0.6998685

## [1] "percent of GJ tweets from ZZH-mentioning accounts"

## [1] 0.1893491

## [1] "percent of GJ tweets from CP-related accounts"
```

```
## [1] 0.1107824
```



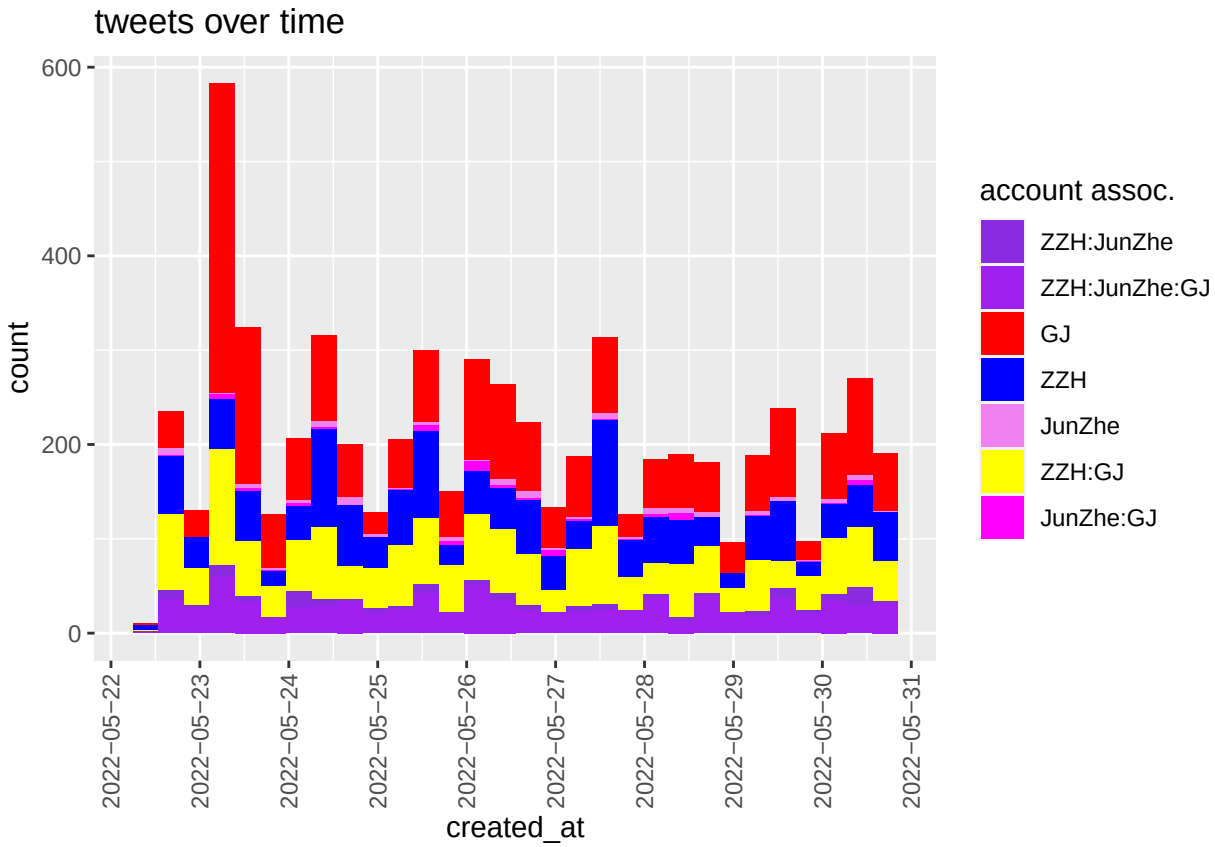
Let's also look at tweets over time for all three hashtags

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Same but this time fill color by venn slice

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```


tweets over time
600
400
200
0
count
2022−05−22
2022−05−23
2022−05−24
2022−05−25
2022−05−26
2022−05−27
2022−05−28
2022−05−29
2022−05−30
2022−05−31
created_at
account assoc.
ZZH:JunZhe
ZZH:JunZhe:GJ
GJ
ZZH
JunZhe
ZZH:GJ
JunZhe:GJ