

Twitter zzh-gj-jz association

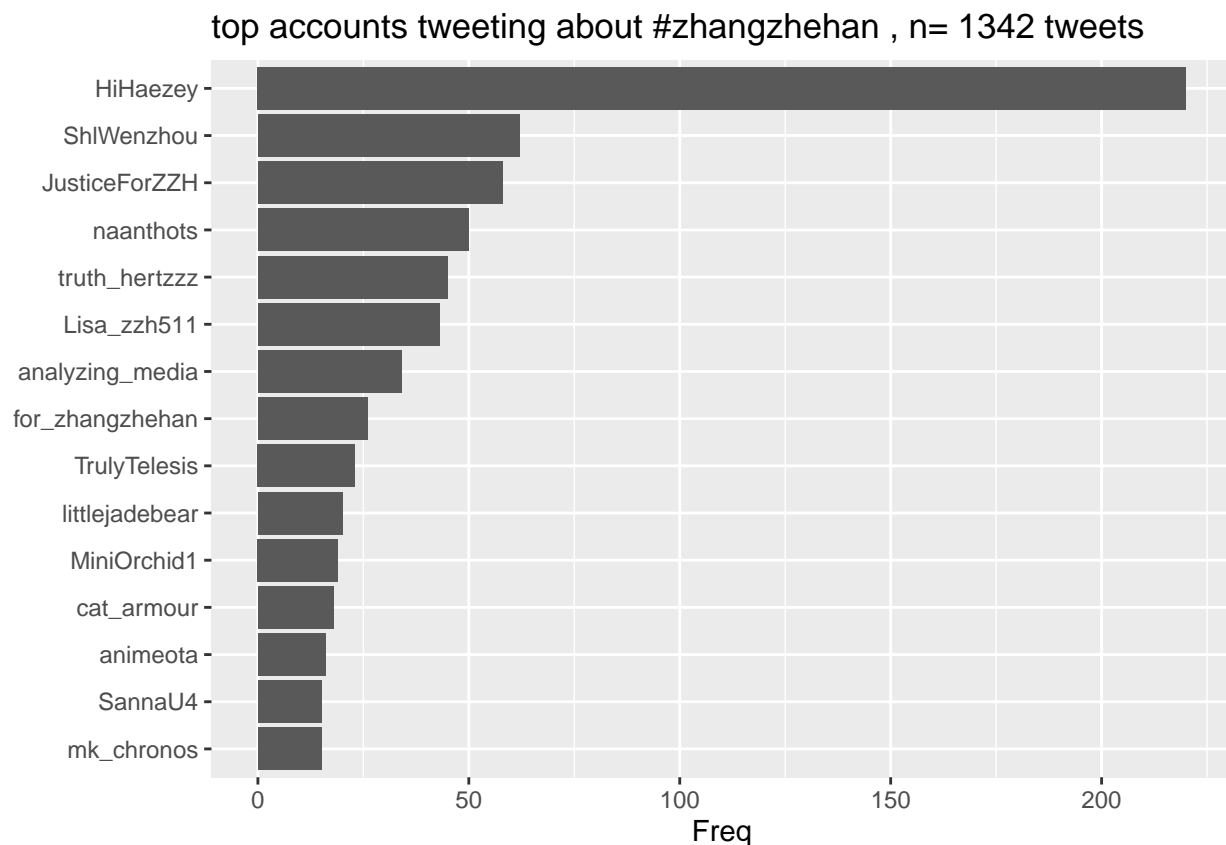
26 May, 2022

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

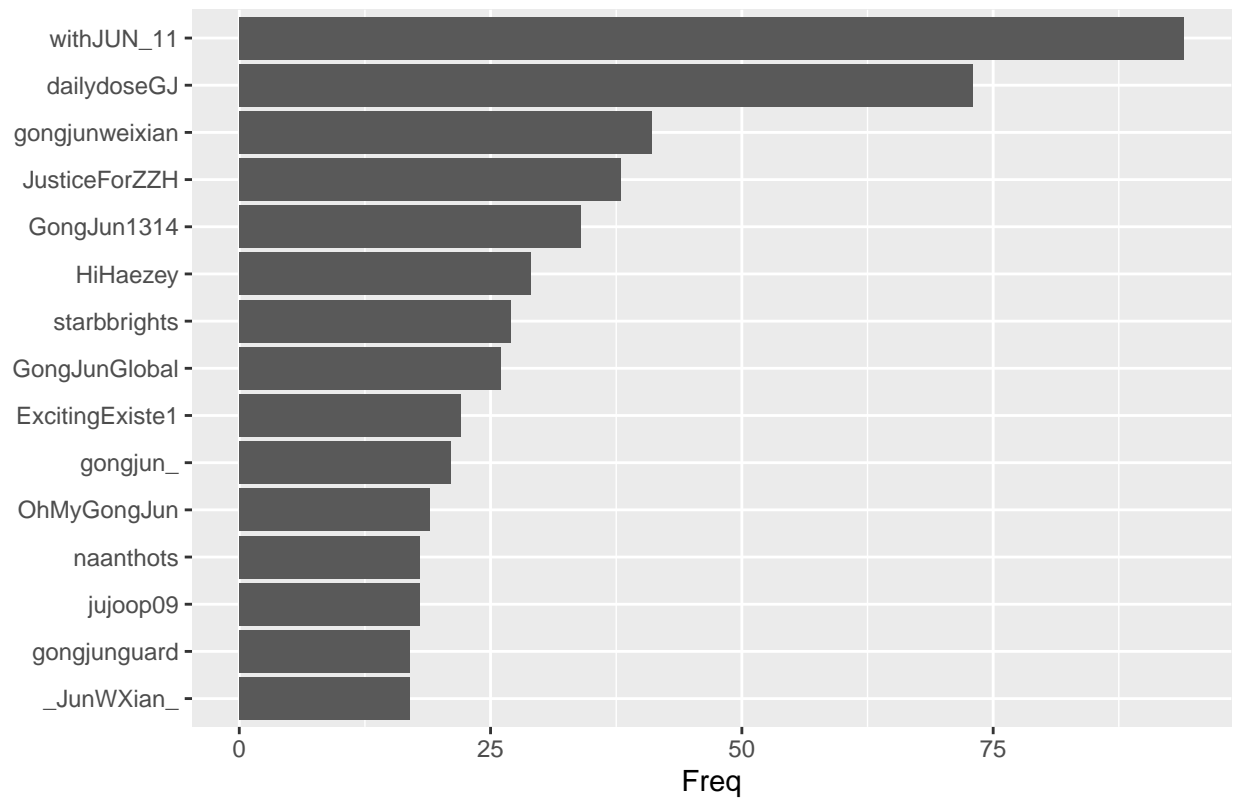
I will then plot the accounts with the most tweets about these topics.

Selecting by Freq



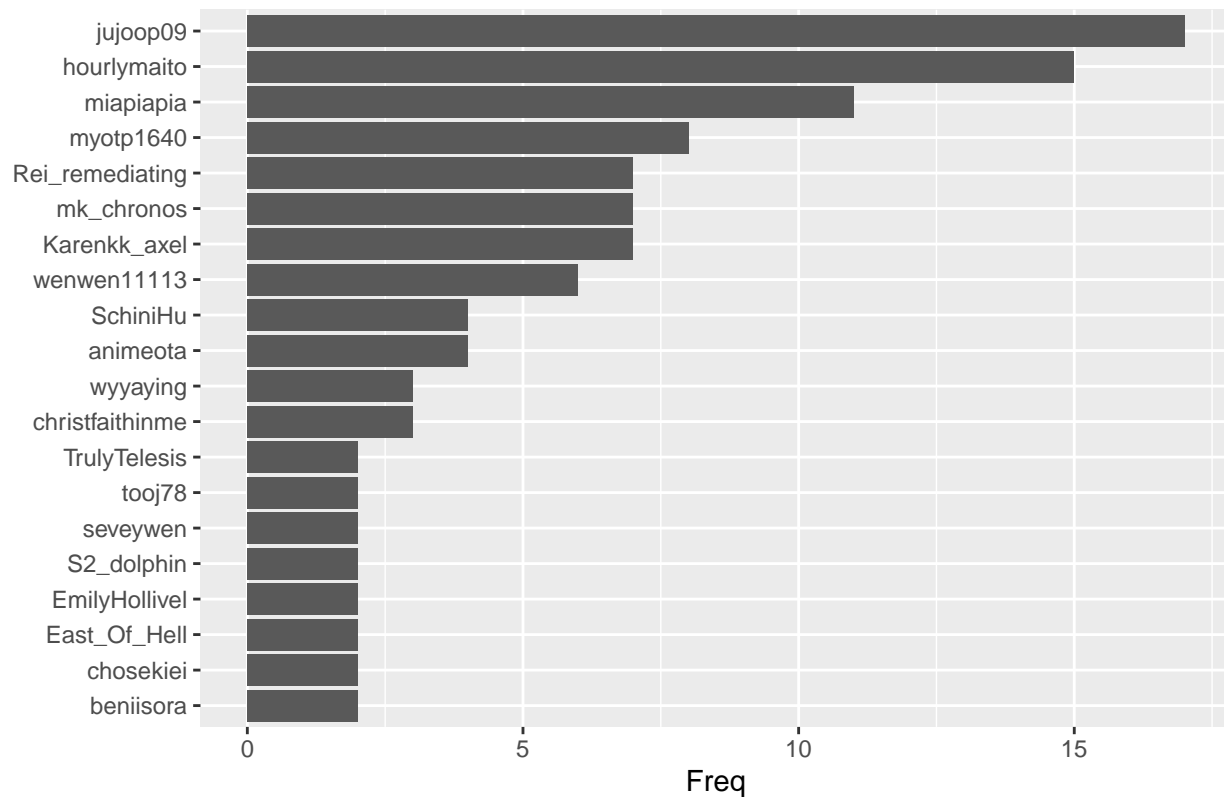
Selecting by Freq

top accounts tweeting about #gongjun , n= 1055 tweets



Selecting by Freq

top accounts tweeting about #junzhe , n= 153 tweets

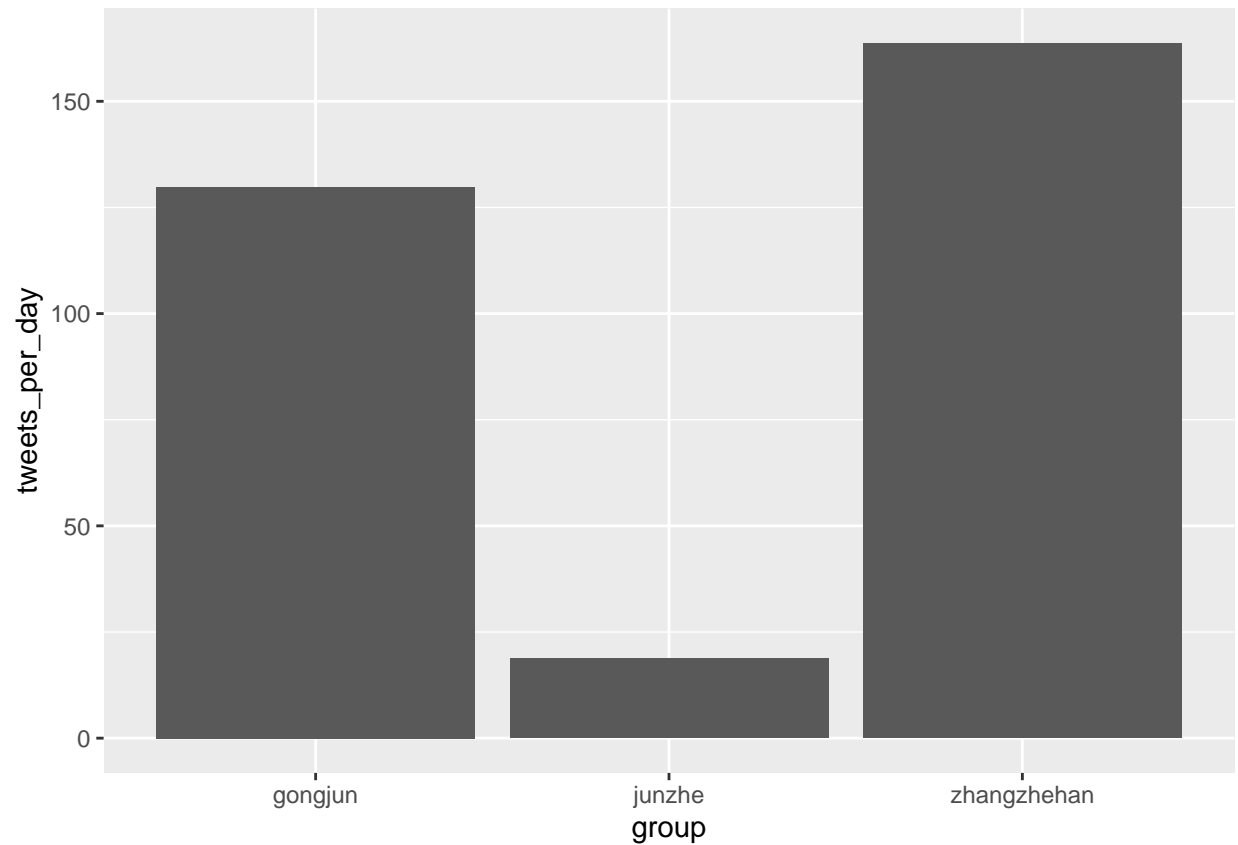


Quick info on which timespan the datasets cover and how many tweets per day

```
## [1] "dataset: #zhangzhehan"
## [1] "number of tweets: 1342"
## [1] "start: 2022-05-18 11:28:28"
## [1] "end: 2022-05-26 16:17:22"
## Time difference of 8.200625 days
## [1] "tweets per day: 163.65"
```

```
## [1] "dataset: #gongjun"
## [1] "number of tweets: 1055"
## [1] "start: 2022-05-18 12:21:31"
## [1] "end: 2022-05-26 15:26:31"
## Time difference of 8.128472 days
## [1] "tweets per day: 129.79"
```

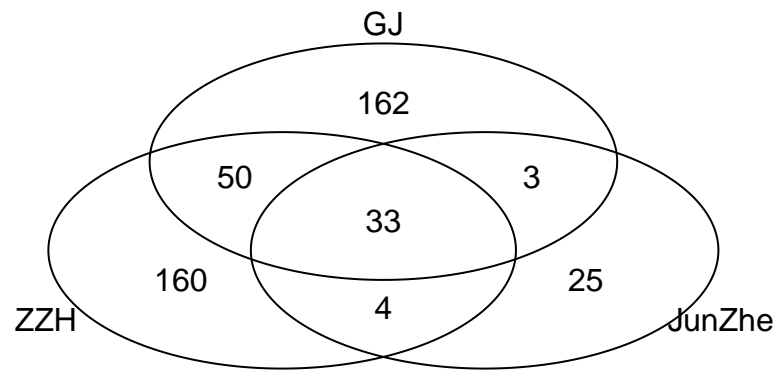
```
## [1] "dataset: #junzhe"
## [1] "number of tweets: 153"
## [1] "start: 2022-05-18 11:28:28"
## [1] "end: 2022-05-26 14:43:27"
## Time difference of 8.135405 days
## [1] "tweets per day: 18.81"
```



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
```

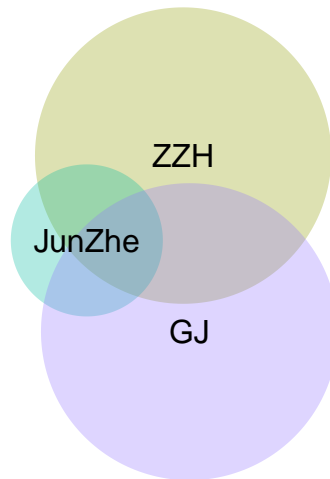


```
## [1] 247
```

```
## [1] 65
```

```
## [1] 248
```

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

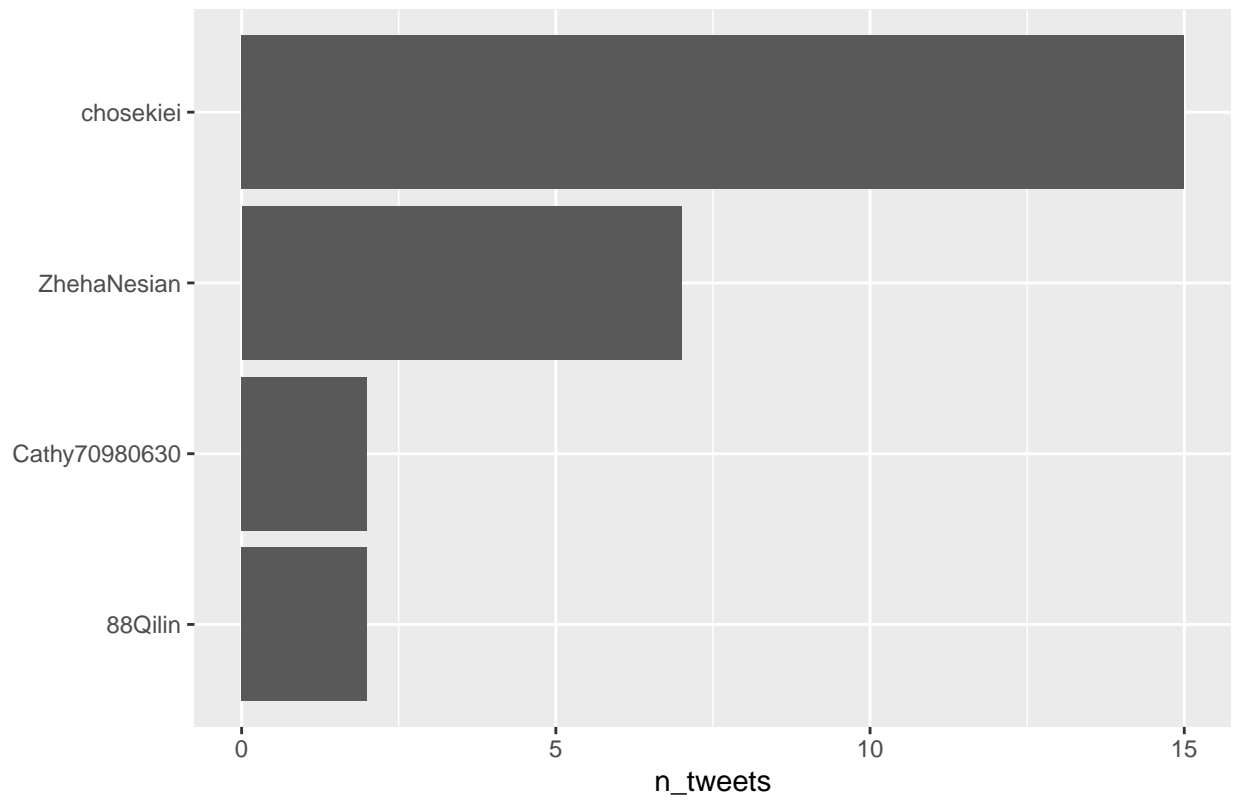


This Venn diagram only shows numbers of accounts tweeting about the three hashtags However I also want to see how many tweets are coming from each section of the Venn diagram.

```
## 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
```

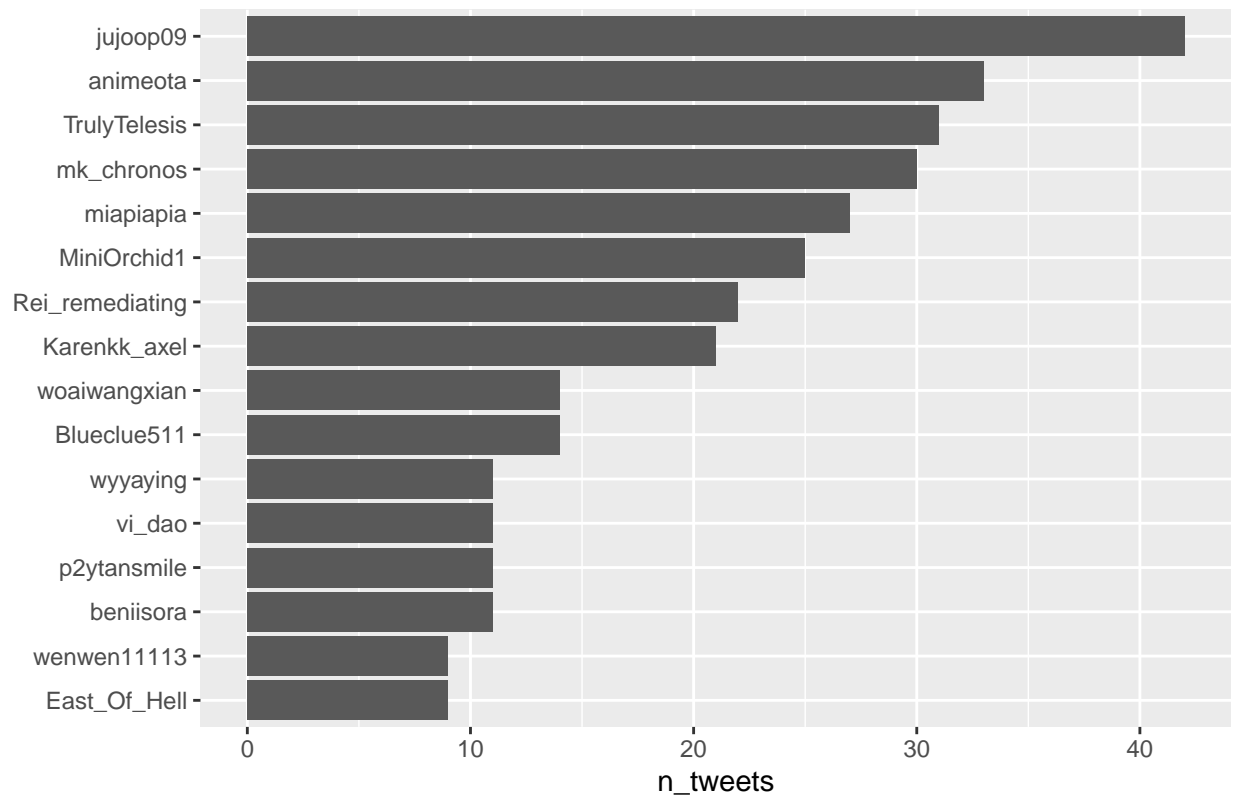
```
## [[1]]
```

top accounts tweeting about ZZH:JunZhe , n= 26 tweets



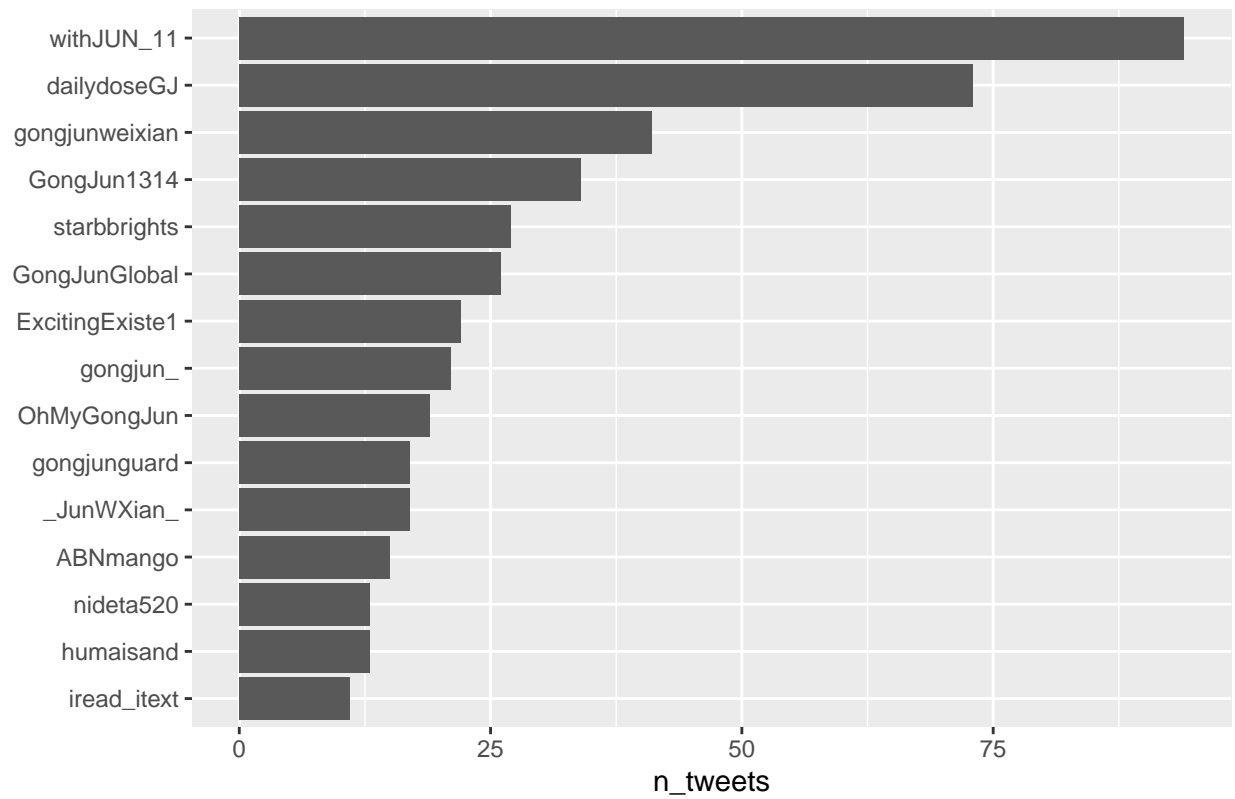
```
##  
## [[2]]
```

top accounts tweeting about ZZH:JunZhe:GJ , n= 396 tweets



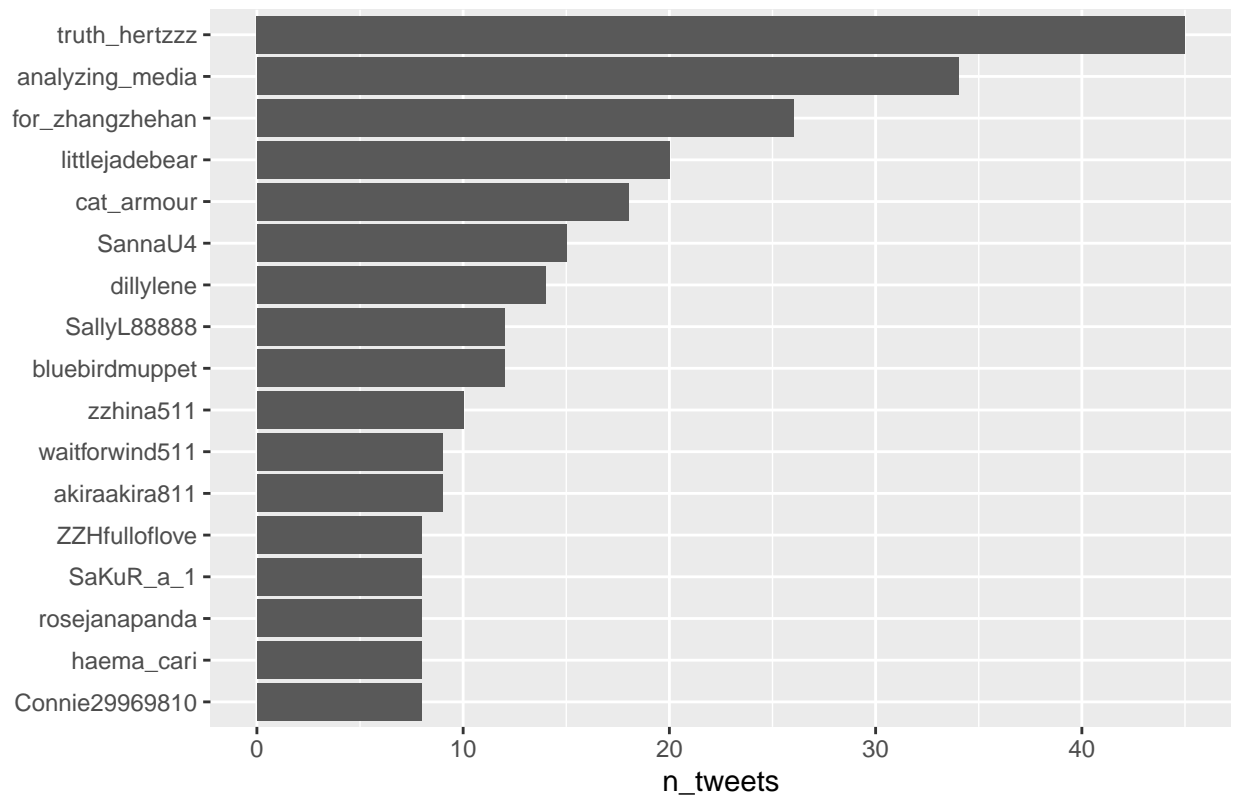
```
##  
## [[3]]
```


top accounts tweeting about GJ , n= 743 tweets



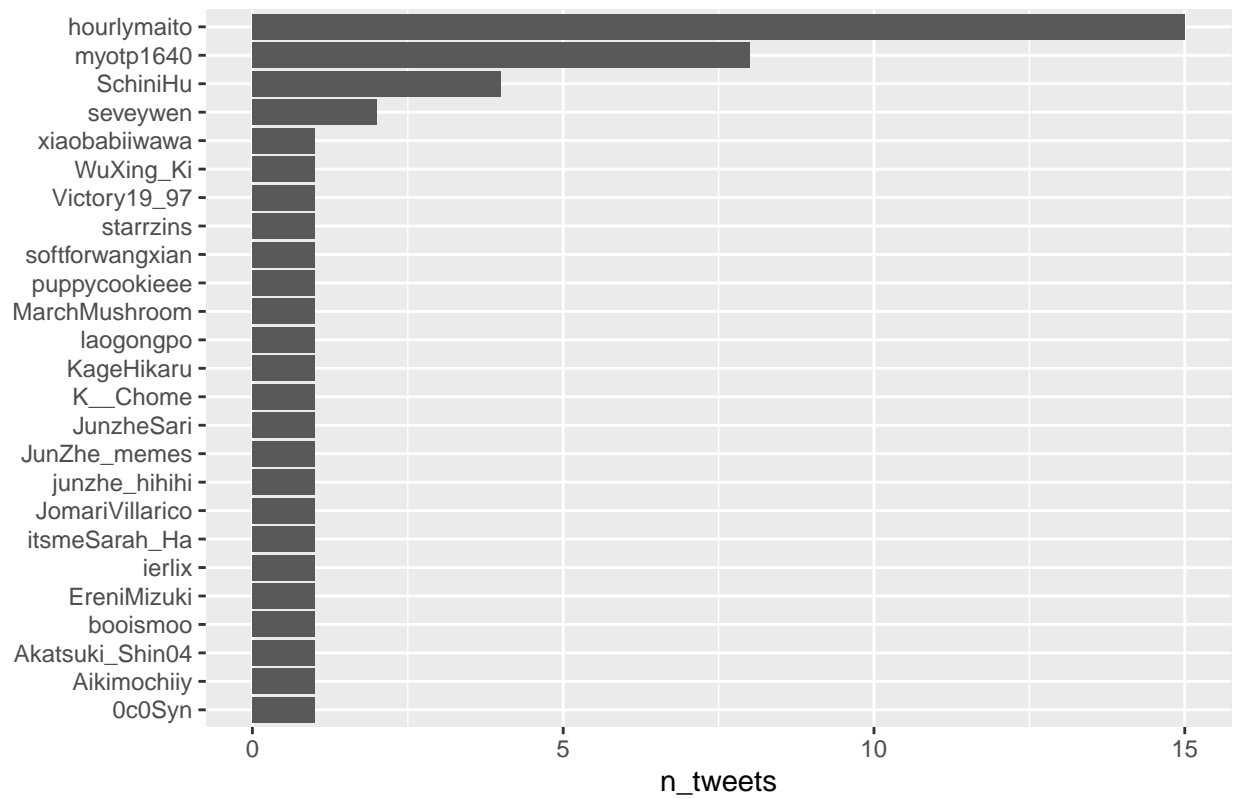
```
##  
## [[4]]
```

top accounts tweeting about ZZH , n= 527 tweets



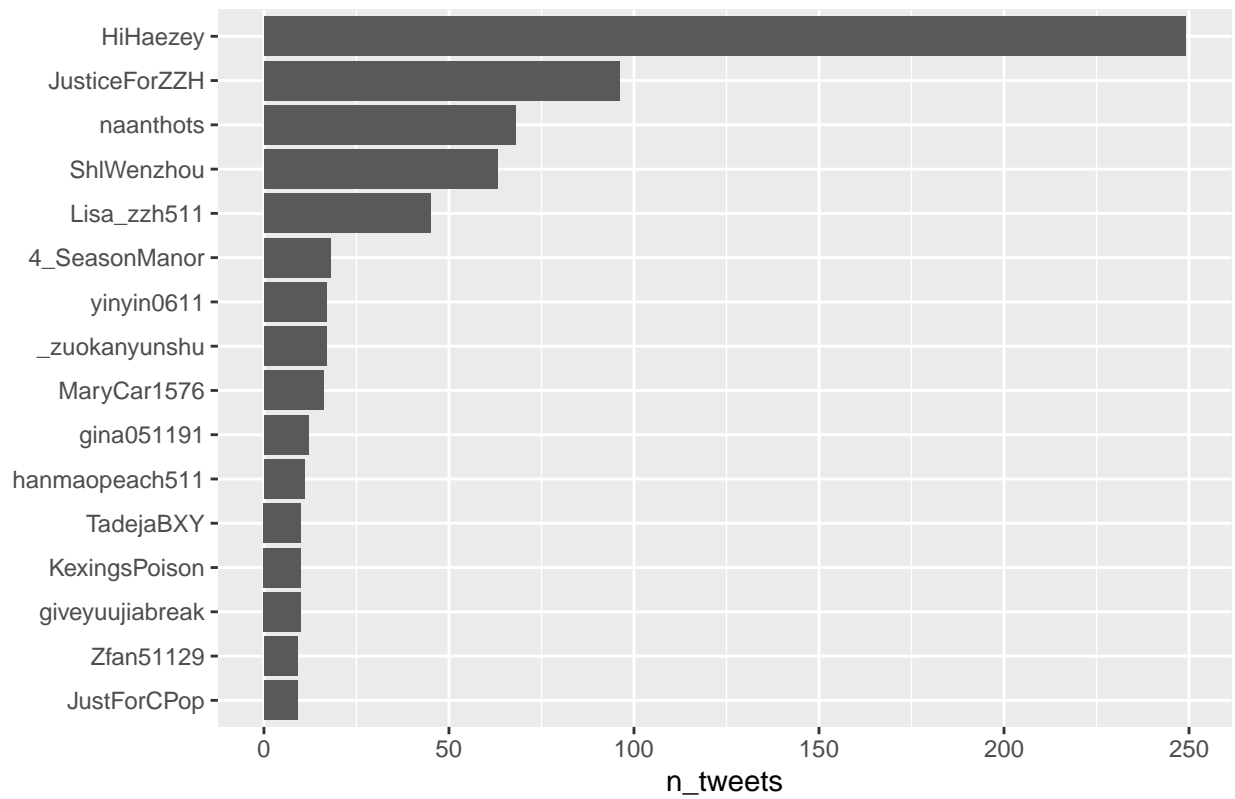
```
##  
## [[5]]
```

top accounts tweeting about JunZhe , n= 50 tweets

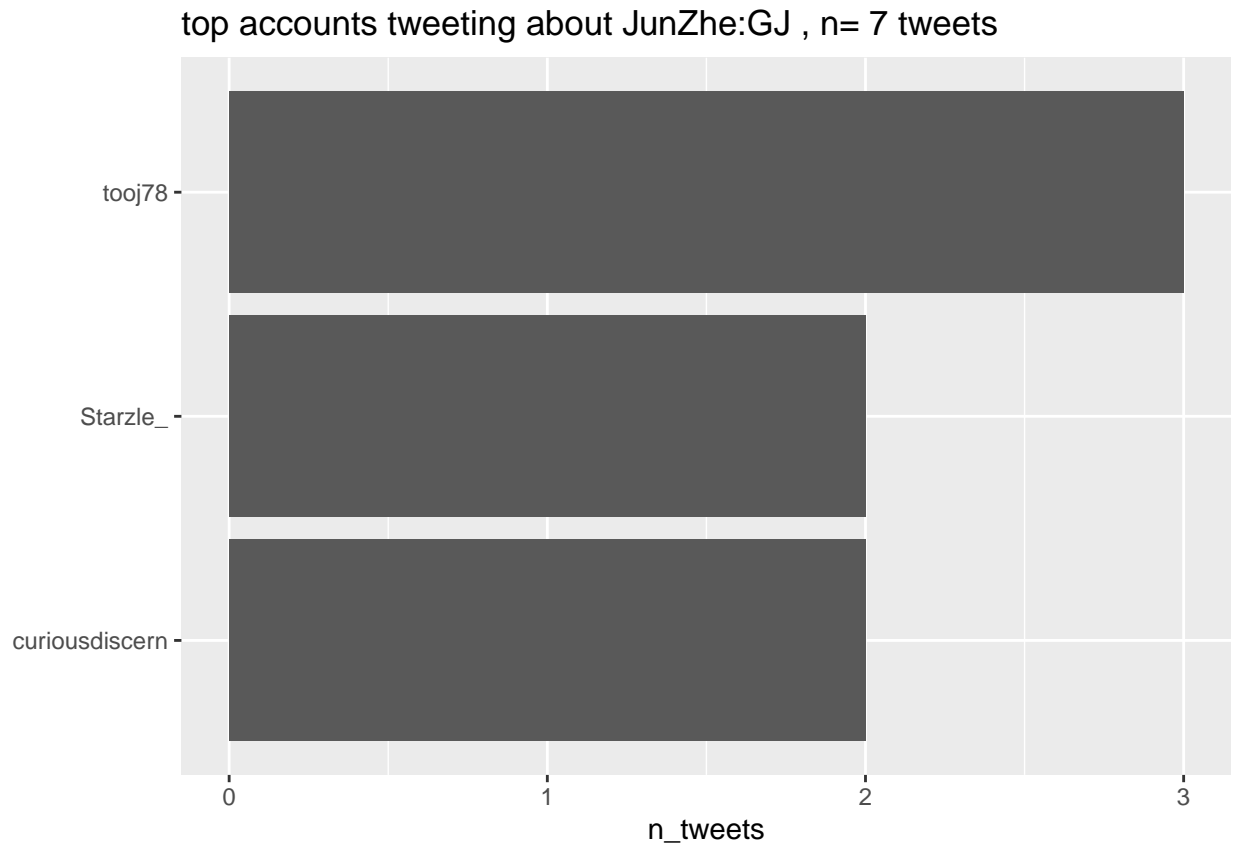


```
##
## [[6]]
```

top accounts tweeting about ZZH:GJ , n= 801 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

```
##
##           GJ JunZhe JunZhe:GJ ZZH ZZH:GJ ZZH:JunZhe ZZH:JunZhe:GJ
## #gongjun    743     0         3  0   199         0         110
## #junzhe      0    50         4  0     0         5         94
## #zhangzhehan 0     0         0 527  602        21        192
```

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

```
## [1] 0.3926975
```

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

```
## [1] 0.4485842
```

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

```
## [1] 0.1587183
```

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

```
## [1] 0.7042654
```

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

```
## [1] 0.1886256
```

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

```
## [1] 0.107109
```

