

A1

ayaz haider

4/25/2021

```
library(twitterR)
```

```
## Warning: package 'twitterR' was built under R version 4.0.5
```

```
library(ROAuth)
```

```
## Warning: package 'ROAuth' was built under R version 4.0.5
```

```
library(tm)
```

```
## Warning: package 'tm' was built under R version 4.0.5
```

```
## Loading required package: NLP
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.0.5
```

```
##
```

```
## Attaching package: 'ggplot2'
```

```
## The following object is masked from 'package:NLP':
```

```
##
```

```
##      annotate
```

```
library("devtools")
```

```
## Warning: package 'devtools' was built under R version 4.0.5
```

```
## Loading required package: usethis
```

```
## Warning: package 'usethis' was built under R version 4.0.5
```

```
library(xts)
```

```
## Warning: package 'xts' was built under R version 4.0.5
```

```

## Loading required package: zoo

## Warning: package 'zoo' was built under R version 4.0.5

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##      as.Date, as.Date.numeric

library(data.table)

## Warning: package 'data.table' was built under R version 4.0.5

##
## Attaching package: 'data.table'

## The following objects are masked from 'package:xts':
##
##      first, last

library(topicmodels)

## Warning: package 'topicmodels' was built under R version 4.0.5

library(sentiment)

## Loading required package: RCurl

## Loading required package: rjson

## Loading required package: plyr

## Warning: package 'plyr' was built under R version 4.0.5

##
## Attaching package: 'plyr'

## The following object is masked from 'package:twitterR':
##
##      id

library(wordcloud)

## Warning: package 'wordcloud' was built under R version 4.0.5

## Loading required package: RColorBrewer

```

```

library(RColorBrewer)

setup_twitter_oauth("jPRTv2TEp4DnffcAsHOY5YT2U", "Lmfb7x6sPXQ4rkCjQ1kzABr8KsRZdEJb0JDxFWUcNnhopGCsq0",

## [1] "Using direct authentication"

tweets <- userTimeline("@OfficialDGISPR", n = 300)

(n.tweet <- length(tweets))

## [1] 300

tweets.df <- twListToDF(tweets)

tweets.df[2, c("id", "created", "screenName", "replyToSN", "favoriteCount", "retweetCount", "longitude"

##           id          created    screenName    replyToSN
## 2 1380149106748837906 2021-04-08 13:22:43 OfficialDGISPR OfficialDGISPR
##   favoriteCount retweetCount longitude latitude
## 2           1727           397         NA      NA
##
## 2 ...Forum also reviewed internal security situation including ongoing third wave of COVID-19 &

writeLines(strwrap(tweets.df$text[1], 60))

## DG ISPR Press Conference - 26th Apr 2021
##
## https://t.co/AKJdMTkvwF

# Replace blank space ("rt")
tweets.df$text <- gsub("rt", "", tweets.df$text)
# Replace @UserName
tweets.df$text <- gsub("@\\w+", "", tweets.df$text)
# Remove punctuation
tweets.df$text <- gsub("[[:punct:]]", "", tweets.df$text)
# Remove links
tweets.df$text <- gsub("http\\w+", "", tweets.df$text)
# Remove tabs
tweets.df$text <- gsub("[ |\\t]{2,}", "", tweets.df$text)
# Remove blank spaces at the beginning
tweets.df$text <- gsub("^ ", "", tweets.df$text)
# Remove blank spaces at the end
tweets.df$text <- gsub(" $", "", tweets.df$text)
# Remove numbers
tweets.df$text <- gsub('[[:digit:]]+', '', tweets.df$text)
# Convert all text to lower case
tweets.df$text <- tolower(tweets.df$text)

```



```
mcorpus<- tm_map(mcorpus, stemDocument)
```

```
## Warning in tm_map.SimpleCorpus(mcorpus, stemDocument): transformation drops  
## documents
```

```
writeLines(strwrap(mcorpus[[1]]$content, 60))
```

```
## dg ispr press conferenceth apr
```

```
stemCompletion2 <- function(x, dictionary) {  
  x <- unlist(strsplit(as.character(x), ' '))  
  ## Unexpectedly, stemCompletion completes an empty string to  
  ## a word in dictionary. Remove empty string to avoid above issue.  
  x <- x[x != ""]  
  x <- stemCompletion(x, dictionary=dictionary)  
  x <- paste(x, sep="", collapse=" ")  
  PlainTextDocument(stripWhitespace(x))  
}
```

```
mcorpus <- lapply(mcorpus, stemCompletion2, dictionary=mcorpusCopy)
```

```
writeLines(strwrap(mcorpus[[1]]$content, 60))
```

```
## dg ispr press conferenceth apr
```

```
wordFreq <- function(corpus, word) {  
  results <- lapply(corpus,  
    function(x) { grep(as.character(x), pattern=paste0("nn<",word)) })  
  sum(unlist(results))  
}
```

```
mcorpus <- Corpus(VectorSource(tweets.df$text))  
replaceWord <- function(corpus, oldword, newword) {  
  tm_map(corpus, content_transformer(gsub),  
    pattern=oldword, replacement=newword)  
}
```

```
mcorpus <- replaceWord(mcorpus, "pak", "pakistan")
```

```
## Warning in tm_map.SimpleCorpus(corpus, content_transformer(gsub), pattern =  
## oldword, : transformation drops documents
```

```
mcorpus <- replaceWord(mcorpus, "ik", "imrankhan")
```

```
## Warning in tm_map.SimpleCorpus(corpus, content_transformer(gsub), pattern =  
## oldword, : transformation drops documents
```

```
mcorpus <- replaceWord(mcorpus, "ns", "nawazsharif")
```

```
## Warning in tm_map.SimpleCorpus(corpus, content_transformer(gsub), pattern =  
## oldword, : transformation drops documents
```

```
tdm <- TermDocumentMatrix(mcorpus) #, control = list(wordLengths = c(1, Inf)))  
tdm
```

```
## <<TermDocumentMatrix (terms: 1649, documents: 300)>>  
## Non-/sparse entries: 4133/490567  
## Sparsity : 99%  
## Maximal term length: 31  
## Weighting : term frequency (tf)
```

```
idx <- which(dimnames(tdm)$Terms %in% c("patwari", "lahore", "naya"))  
#as.matrix(tdm[idx, 21:30])  
(freq.terms <- findFreqTerms(tdm, lowfreq = 2))
```

```
## [1] "conferenceth"      "ispr"      "press"  
## [4] "also"              "amp"        "covid"  
## [7] "forum"             "including"  "internal"  
## [10] "of..."           "ongoing"    "reaffirmed"  
## [13] "resolve"           "reviewed"   "security"  
## [16] "situation"         "wave"       "administration"  
## [19] "civil"             "earlier"    "expressed"  
## [22] "for"               "satisfaction" "their"  
## [25] "ultimate"          "detailed"   "reiterated"  
## [28] "review"            "the"        "along"  
## [31] "eastern"           "environment" "especially"  
## [34] "fire"              "post"       "ccc"  
## [37] "chaired"           "coas"       "comprehenawazsharifive"  
## [40] "ghq"               "global"     "held"  
## [43] "paicipants"        "regional"   "today"  
## [46] "undeook"           "any"        "benefit"  
## [49] "country"           "designawazsharif" "from"  
## [52] "have"              "hostile"    "region"  
## [55] "towards"           "will"       "will..."  
## [58] "continue"          "pakistanistan" "relationawazsharif"  
## [61] "russia"            "said"       "that"  
## [64] "values"            "with"       "achievements"  
## [67] "acknowledged"      "againawazsharift" "contributionawazsharif"  
## [70] "dignitary"         "pakistanistan's" "terrorism"  
## [73] "visiting"          "war"        "bajwa"  
## [76] "called"            "during"     "foreign"  
## [79] "general"           "javed"      "minister"  
## [82] "qamar"             "th..."    "day"  
## [85] "one"               "<U+0627><U+06CC><U+06A9>" "are"  
## [88] "contributed"       "honour"     "nation"  
## [91] "our"               "they"       "women"  
## [94] "always"            "feb"        "motherland"  
## [97] "pakistan"          "suppo"      "combat"  
## [100] "formation"         "highest"    "preparedness"
```

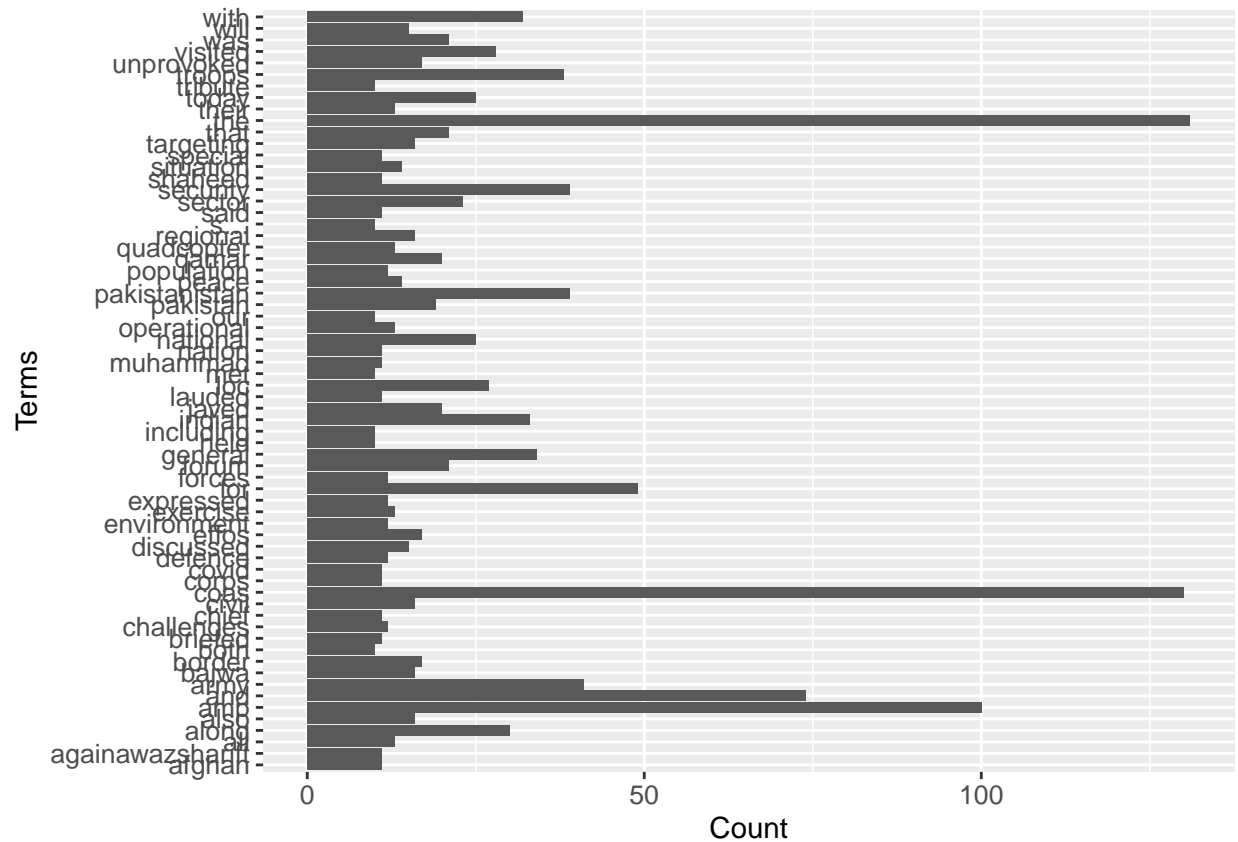
## [103]	"readiness"	"training"	"battlefield"
## [106]	"been"	"conditionawazsharif"	"exercise"
## [109]	"field"	"over"	"this"
## [112]	"troops"	"under"	"forces"
## [115]	"given"	"infantry"	"integrated"
## [118]	"was"	"area"	"near"
## [121]	"visited"	"wherein"	"witnessed"
## [124]	"admiral"	"chairman"	"cnawazsharif"
## [127]	"countries"	"international"	"jcsc"
## [130]	"cjcscc"	"commitment"	"emphasised"
## [133]	"its"	"peaceful"	"shared"
## [136]	"appreciated"	"effos"	"multinational"
## [139]	"navy"	"sai..."	"and"
## [142]	"improving"	"joint"	"maritime"
## [145]	"operationawazsharif"	"sea"	"threats"
## [148]	"through"	"upon"	"..."
## [151]	"a..."	"being"	"noh"
## [154]	"chiefs"	"committee"	"conduct"
## [157]	"nadeem"	"raza"	"staff"
## [160]	"dividends"	"improved"	"incl"
## [163]	"numerous"	"complete"	"design"
## [166]	"peace"	"concluded"	"challenges"
## [169]	"interest"	"keeping"	"national"
## [172]	"responawazsharife"	"afghan"	"dilating"
## [175]	"right"	"self"	"kashmir"
## [178]	"noted"	"solidarity"	"heldghq"
## [181]	"i..."	"kashmiris"	"struggle"
## [184]	"valiant"	"conducted"	"flight"
## [187]	"successful"	"surface"	"test"
## [190]	"condolences"	"expresses"	"heafelt"
## [193]	"may"	"mother"	"passing"
## [196]	"t..."	"earned"	"enduring"
## [199]	"hard"	"highlighting"	"shall"
## [202]	"stability"	"while"	"corps"
## [205]	"lauded"	"leas"	"peshawar"
## [208]	"police"	"border"	"management"
## [211]	"update"	"balochistan"	"prosperity"
## [214]	"due"	"focus"	"future"
## [217]	"pakistanis..."	"strategic"	"blood"
## [220]	"garrison"	"incident"	"justice"
## [223]	"mayrs"	"measures"	"taken"
## [226]	"command"	"long"	"met"
## [229]	"quetta"	"visit"	"and..."
## [232]	"enemy"	"president"	"prime"
## [235]	"indigenously"	"s..."	"all"
## [238]	"families"	"secure"	"supreme"
## [241]	"tribute"	"edge"	"elements"
## [244]	"levels"	"meet"	"other"
## [247]	"power"	"vital"	"loc"
## [250]	"must"	"professional"	"working"
## [253]	"current"	"spectrum"	"threat"
## [256]	"determined"	"ever"	"freedom"
## [259]	"indian"	"suppress"	"brave"
## [262]	"chakothi"	"down"	"intr..."

## [265] "pakistanistanarmy"	"quadcopter"	"sector"
## [268] "shot"	"spying"	"nefarious"
## [271] "operation"	"progress"	"search"
## [274] "such"	"last"	"night"
## [277] "terrorists"	"anniversary"	"can"
## [280] "eah"	"high"	"morale"
## [283] "operational"	"success"	"both"
## [286] "emerging"	"face"	"geostrategic"
## [289] "capacity"	"air"	"china"
## [292] "force"	"witness"	"army"
## [295] "new"	"established"	"clearly"
## [298] "make"	"deliberately"	"resoed"
## [301] "unprovoked"	"muhammad"	"naimrankhan"
## [304] "nishane"	"sacrifice"	"shaheed"
## [307] "venerates"	"armed"	"helicopters"
## [310] "defence"	"the..."	"weaponawazsharif"
## [313] "were"	"which"	"cas"
## [316] "gujranwala"	"haider"	"his"
## [319] "hussain"	"cfvs"	"initiated"
## [322] "khuiratta"	"responded"	"amp..."
## [325] "maj"	"valour"	"odds"
## [328] "best"	"military"	"out"
## [331] "demise"	"former"	"pakistan..."
## [334] "sad"	"between"	"cooperation"
## [337] "enhancement"	"brotherly"	"greatly"
## [340] "china's"	"projects"	"p..."
## [343] "fenghe"	"matters"	"meeting"
## [346] "mutual"	"wei"	"arrival"
## [349] "contingent"	"bilateral"	"paicularly"
## [352] "prevailing"	"tajimrankhanistan"	"bagsar"
## [355] "cfv"	"civilian"	"indianarmy"
## [358] "innocent"	"population"	"targeting"
## [361] "gen"	"syed"	"abdul"
## [364] "appointments"	"aziz"	"comd"
## [367] "in..."	"lieutenant"	"major"
## [370] "promoted"	"chief"	"oppounities"
## [373] "people"	"these"	"commanders"
## [376] "directed"	"effor..."	"enawazsharifure"
## [379] "wake"	"firing"	"strong"
## [382] "ajk"	"state"	"geo"
## [385] "presided"	"grief"	"amb"
## [388] "discussed"	"bajwacoas"	"condolence"
## [391] "afg"	"process"	"sit"
## [394] "abreast"	"developments"	"faculty"
## [397] "interacting"	"stay"	"about"
## [400] "briefed"	"e..."	"create"
## [403] "fencing"	"highlighted"	"pakistanafghan"
## [406] "speaking"	"interacted"	"federation"
## [409] "op..."	"pakistanistanrussian"	"procedures"
## [412] "rescue"	"special"	"college"
## [415] "nasir"	"khanjar"	"rakhchimrankhanri"
## [418] "sectors"	"fuher"	"pledged"
## [421] "polio"	"despite"	"gates"
## [424] "pandemic"	"affected"	"child"

## [427]	"workers"	"coverage"	"drive"
## [430]	"enawazsharifuring"	"proper"	"reach"
## [433]	"suppoing"	"bill"	"away"
## [436]	"afghanistan"	"ambassador"	"help"
## [439]	"services"	"khan"	"druzhbav"
## [442]	"paicipating"	"tarbela"	"chinese"
## [445]	"ceremony"	"russian"	"opening"
## [448]	"only"	"she"	"woman"
## [451]	"rashid"	"enhance"	"c..."
## [454]	"proud"	"two"	"conference"
## [457]	"ali"	"extend"	"has"
## [460]	"medical"	"battle"	"local"
## [463]	"areas"	"latest"	"strat"
## [466]	"bahawalpur"	"welfare"	"quality"
## [469]	"who"	"men"	"officers"
## [472]	"administrative"	"competition"	"across"
## [475]	"karachi"	"khalilzad"	"reconciliation"
## [478]	"representative"	"zalmay"	"sco"
## [481]	"serve"	"facility"	"skardu"
## [484]	"amir"	"professionalism"	"alive"
## [487]	"focused"	"induction"	"recent"
## [490]	"tank"	"duty"	"laying"
## [493]	"life"	"line"	"appr..."
## [496]	"more"	"terrorist"	"commander"
## [499]	"ihshan"	"automatics"	"moars"
## [502]	"nishanehaider"	"bajw..."	"leadership"
## [505]	"sides"	"had"	"flood"
## [508]	"relief"	"as..."	"campaign"
## [511]	"control"	"locust"	"evolving"
## [514]	"to..."	"than"	"officer"
## [517]	"resources"	"effectively"	"urban"
## [520]	"issues"	"level"	"where"
## [523]	"better"	"spirit"	"eid"
## [526]	"spent"	"time"	"ceasefire"
## [529]	"violation"	"capabilities"	"hit"
## [532]	"outstanding"	"speed"	"heavy"
## [535]	"coas..."	"fight"	"violationcfv"
## [538]	"old"	"civilianawazsharif"	"got"
## [541]	"injured"	"every"	"paid"
## [544]	"glowing"	"additional"	"movement"
## [547]	"social"	"possible"	"rangers"
## [550]	"sindh"	"lives"	"addressed"
## [553]	"led"	"team"	"laid"
## [556]	"civ"	"nimrankhanial"	"torkham"
## [559]	"steps"	"evacuated"	"peacekeepers"
## [562]	"pia"		

```
term.freq <- rowSums(as.matrix(tdm))
term.freq <- subset(term.freq, term.freq >= 10)
df <- data.frame(term = names(term.freq), freq = term.freq)
```

```
ggplot(df, aes(x=term, y=freq)) + geom_bar(stat="identity") +
  xlab("Terms") + ylab("Count") + coord_flip() +
  theme(axis.text=element_text(size=10))
```



```
m <- as.matrix(tdm)
word.freq <- sort(rowSums(m), decreasing = T)

pal <- brewer.pal(9, "BuGn")[-(1:4)]

wordcloud(words = names(word.freq), freq = word.freq, min.freq = 1, scale=c(3,0.2), colors=brewer.pal(8,
```



```
findAssocs(tdm, "nishanehaider", 0.1)
```

##	\$nishanehaider						
##	bhatti	honours	l...	raja	stellar	bravery	moally
##	0.71	0.71	0.71	0.71	0.71	0.71	0.71
##	mohammad	ofmajor	resilience	tufail	wou...	being	despite
##	0.71	0.71	0.71	0.71	0.71	0.50	0.50
##	shaheed	aziz	war	supreme	major	his	tribute
##	0.42	0.40	0.35	0.26	0.26	0.24	0.21
##	sacrifice	nation					
##	0.20	0.19					

```
findAssocs(tdm, "imran", 0.1)
```

##	\$imran					
##	clear	connectivity	regarding	vision	khan	has
##	1.00	1.00	1.00	1.00	0.58	0.58
##	prime	region	given	a...	minister	said
##	0.50	0.44	0.44	0.40	0.35	0.30
##	peace	that				
##	0.24	0.21				

```
dtm <- as.DocumentTermMatrix(tdm)
```

```
rowTotals <- apply(dtm , 1, sum)
dtm.new <- dtm[rowTotals> 0, ]
lda <- LDA(dtm.new, k = 7)
```

```
term <- terms(lda, 3) # first 7 terms of every topic
(term <- apply(term, MARGIN = 2, paste, collapse = ", "))
```

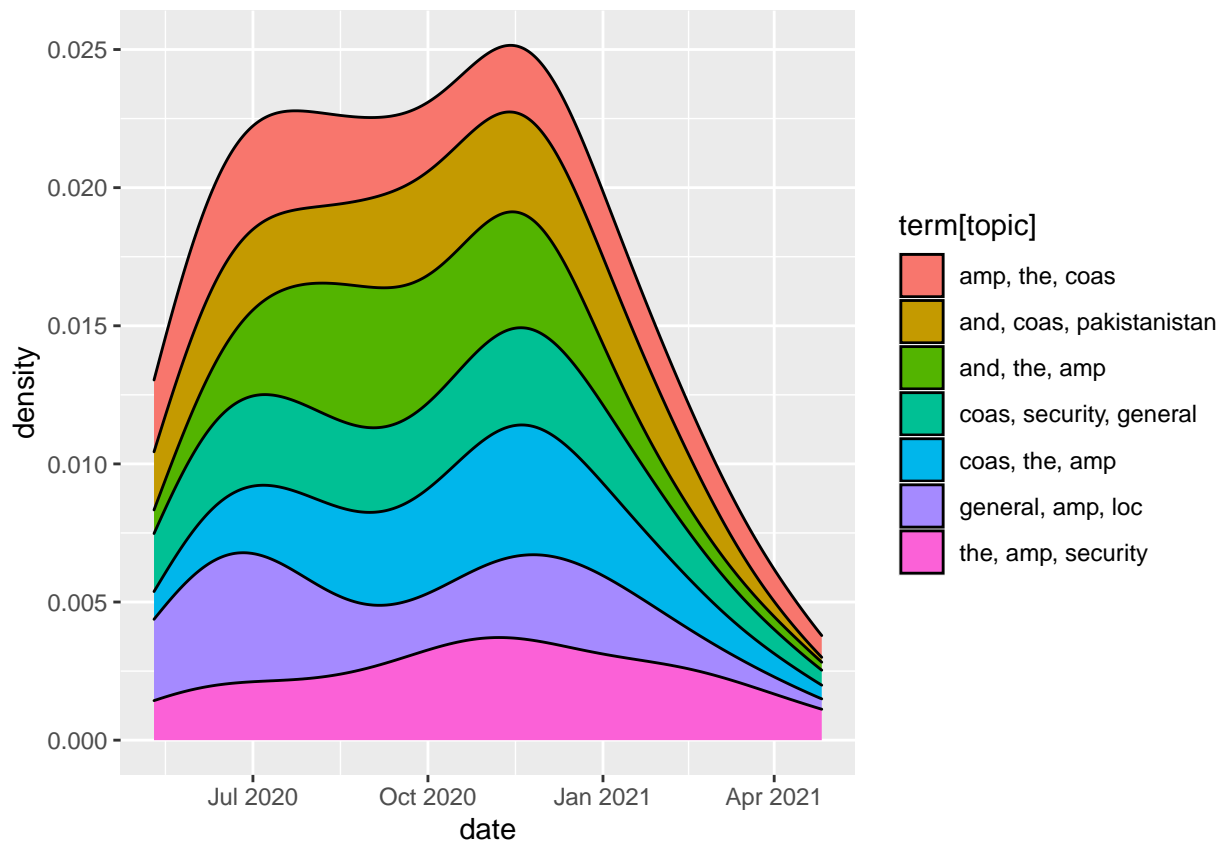
```
##                Topic 1                Topic 2
##      "general, amp, loc"      "amp, the, coas"
##                Topic 3                Topic 4
## "and, coas, pakistanistan" "coas, security, general"
##                Topic 5                Topic 6
##      "and, the, amp"      "coas, the, amp"
##                Topic 7
##      "the, amp, security"
```

```
topics <- topics(lda)
```

```
topics <- data.frame(date=as.IDate(tweets.df$created[c(1:400,1)]), topic=topics[c(1:400,1)])
```

```
ggplot(topics, aes(date, fill = term[topic])) +
  geom_density(position = "stack")
```

```
## Warning: Removed 100 rows containing non-finite values (stat_density).
```



```
sentiments <- sentiment(tweets.df$text)
table(sentiments$polarity)
```

```
##
## negative  neutral  positive
##        11        286         3
```

```
sentiments$score <- 1
sentiments$score[sentiments$polarity == "positive"] <- 4
sentiments$score[sentiments$polarity == "negative"] <- -4
sentiments$date <- as.IDate(tweets.df$created) # as.ITime(tweets.df$created)
result <- aggregate(score ~ date, data = sentiments, sum)

plot(result, type = "l")
```

