

# SYRIA - Sentiment Analysis

Aravind

March 2, 2018

## Api Info

```
api_key <- '###nHS#####YMy####o#J#k##'  
api_secret <- '#####8ghL#####51Bv7UY###7DW8RE####dQDi##h#'  
access_token <- "1#####-lpiqi0irm#####jKL5CgyY31Slvmk#####"  
access_token_secret <- "#####4KMv4a###g8tr#####Mvpa###"
```

Use your own API keys and tokens.

### Setup Twitter API

```
library(twitterR)  
setup_twitter_oauth(api_key, api_secret, access_token, access_token_secret)
```

```
## [1] "Using direct authentication"
```

## Getting tweets on Syria War

```
tweets <- searchTwitter('#SyriaWar', n=2000, lang = 'en')  
head(tweets)
```

```
## [[1]]  
## [1] "Hari54016313: @Suriya_offl @VigneshShivN @anirudhofficial @kegvraja@Keerth  
yOfficial @meramyakrishnan @menongautham... https://t.co/zVb9y1NEdr"  
##  
## [[2]]  
## [1] "singh300_singh: RT @cityheadlines07: Sikh charity group #KhalsaAid extends  
help to Syrian refugees @Khalsa_Aid \n\n#SyriaWar #Syrianchildren #Syria #Syrian..."  
##  
## [[3]]  
## [1] "RJ_Prateeksha: https://t.co/hvKYJIot7S \nHights of #cruelty #HumansWantHum  
anity #HumanRightsDiedInSyria #SaveSyrianChildren... https://t.co/eIN2Go8EoS"  
##  
## [[4]]  
## [1] "shaikhmazhar21: RT @cityheadlines07: Sikh charity group #KhalsaAid extends  
help to Syrian refugees @Khalsa_Aid \n\n#SyriaWar #Syrianchildren #Syria #Syrian..."  
##  
## [[5]]  
## [1] "srk141119: RT @Najjo786: This picture tell us the whole story of #SyriaWar.  
https://t.co/JC0wD1AN6D"  
##  
## [[6]]  
## [1] "PersianGulf4eva: @ExtinctMedia @potkazar #USA signed the fake fomented\ "Is  
lamic revolution\" when they flew BBC promoted Indian-briti...https://t.co/VomnsLz4j  
U"
```

```
tweetsdf <- twListToDF(tweets)
write.csv(tweetsdf, file="syria.csv", row.names=F)
```

## Read file

```
syriawar <- read.csv("syria.csv", header = T)
str(syriawar)
```

```
## 'data.frame':    2000 obs. of  16 variables:
##  $ text          : Factor w/ 1018 levels "'In Russia, the law prohibits mercenar
y armies. They have been fighting for years in the east of Ukraine, in S"|_trunca
ted_,...: 215 617 398 617 682 165 729 535 613 613 ...
##  $ favorited     : logi FALSE FALSE FALSE FALSE FALSE FALSE ...
##  $ favoriteCount: int  0  0  1  0  0  1  0  0  0  0 ...
##  $ replyToSN     : Factor w/  79 levels "21stCenWar","aajtak",...: 68 NA NA NA NA 1
9 NA NA NA NA ...
##  $ created       : Factor w/ 1988 levels "2018-02-27 02:44:45",...: 1988 1987 1986
1985 1984 1983 1982 1981 1980 1979 ...
##  $ truncated     : logi TRUE FALSE TRUE FALSE FALSE TRUE ...
##  $ replyToSID    : num 9.7e+17 NA NA NA NA ...
##  $ id            : num 9.7e+17 9.7e+17 9.7e+17 9.7e+17 9.7e+17 ...
##  $ replyToUID    : num 3.07e+09 NA NA NA NA ...
##  $ statusSource  : Factor w/  25 levels "<ahref=\"http://blackberry.com/twitter\"
rel=\"nofollow\">Twitter for BlackBerry®</a>",...: 9 9 9 9 9 7 9 10 9 9 ...
##  $ screenName    : Factor w/ 1725 levels "_Bharath_Cherry",...: 531 1414 1225 1362
1450 1086 361 495 817 152 ...
##  $ retweetCount  : int  0 43  0 43 85  0  5  0  6  6 ...
##  $ isRetweet     : logi FALSE TRUE FALSE TRUE TRUE FALSE ...
##  $ retweeted     : logi FALSE FALSE FALSE FALSE FALSE FALSE ...
##  $ longitude     : logi  NA NA NA NA NA NA ...
##  $ latitude      : logi  NA NA NA NA NA NA ...
```

## Build Corpus

```
require(tm)
```

```
## Loading required package: tm
```

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

```
corpus <- iconv(syriawar$text, to = "utf-8")
corpus <- Corpus(VectorSource(corpus))
inspect(corpus[1:5])
```

```
## <<SimpleCorpus>>
## Metadata:  corpus specific: 1, document level (indexed): 0
## Content:  documents: 5
##
## [1] @Suriya_offl @VigneshShivN @anirudhofficial @kegvraja @KeerthyOfficial@mera
myakrishnan @menongauthamâ€¦ https://t.co/zVb9y1NEdr
## [2] RT @cityheadlines07: Sikh charity group #KhalsaAid extends help to Syrianre
fugees @Khalsa_Aid \n\n#SyriaWar  #Syrianchildren #Syria #Syrianâ€¦
## [3] https://t.co/hvKYJIot7S \nHights of #cruelty  #HumansWantHumanity #HumanRigh
tsDiedInSyria #SaveSyrianChildrenâ€¦ https://t.co/eIN2Go8EoS
## [4] RT @cityheadlines07: Sikh charity group #KhalsaAid extends help to Syrianre
fugees @Khalsa_Aid \n\n#SyriaWar  #Syrianchildren #Syria #Syrianâ€¦
## [5] RT @Najjo786: This picture tell us the whole story of #SyriaWar.https://t.c
o/JC0wD1AN6D
```

## Clean text

```
corpus <- tm_map(corpus,tolower)
inspect(corpus[1:5])
```

```
## <<SimpleCorpus>>
## Metadata:  corpus specific: 1, document level (indexed): 0
## Content:  documents: 5
##
## [1] @suriya_offl @vigneshshivn @anirudhofficial @kegvraja @keerthyofficial@mera
myakrishnan @menongauthamâ€¦ https://t.co/zvb9y1nedr
## [2] rt @cityheadlines07: sikh charity group #khalsaaaid extends help to syrianre
fugees @khalsa_aid \n\n#syriawar  #syrianchildren #syria #syrianâ€¦
## [3] https://t.co/hvkyjiot7s \nhights of #cruelty  #humanswanthumanity #humanrigh
tsdiedinsyria #savesyrianchildrenâ€¦ https://t.co/ein2go8eos
## [4] rt @cityheadlines07: sikh charity group #khalsaaaid extends help to syrianre
fugees @khalsa_aid \n\n#syriawar  #syrianchildren #syria #syrianâ€¦
## [5] rt @najjo786: this picture tell us the whole story of #syriawar.https://t.c
o/jc0wd1an6d
```

```
corpus <- tm_map(corpus,removePunctuation)
inspect(corpus[1:5])
```

```
## <<SimpleCorpus>>
## Metadata: corpus specific: 1, document level (indexed): 0
## Content: documents: 5
##
## [1] suriyaoffl vigneshshivn anirudhofficial kegvraja keerthyofficialmeramyakris
hnan menongauthamâ€¦ httpstcozvb9ylnedr
## [2] rt cityheadlines07 sikh charity group khalsaaaid extends help to syrianrefug
ees khalsaaaid \n\nsyriawar syrianchildren syria syrianâ€¦
## [3] httpstcohvkyjiot7s \nhights of cruelty humanswanthumanity humanrightsdiedin
syria savesyrianchildrenâ€¦ httpstcoein2go8eos
## [4] rt cityheadlines07 sikh charity group khalsaaaid extends help to syrianrefug
ees khalsaaaid \n\nsyriawar syrianchildren syria syrianâ€¦
## [5] rt najjo786 this picture tell us the whole story of syriawarhttpstcojco0wdla
n6d
```

```
corpus <- tm_map(corpus,removeNumbers)
inspect(corpus[1:5])
```

```
## <<SimpleCorpus>>
## Metadata: corpus specific: 1, document level (indexed): 0
## Content: documents: 5
##
## [1] suriyaoffl vigneshshivn anirudhofficial kegvraja keerthyofficialmeramyakris
hnan menongauthamâ€¦ httpstcozvbbynedr
## [2] rt cityheadlines sikh charity group khalsaaaid extends help to syrianrefugee
s khalsaaaid \n\nsyriawar syrianchildren syria syrianâ€¦
## [3] httpstcohvkyjiots \nhights of cruelty humanswanthumanity humanrightsdiedins
yria savesyrianchildrenâ€¦ httpstcoeingoeos
## [4] rt cityheadlines sikh charity group khalsaaaid extends help to syrianrefugee
s khalsaaaid \n\nsyriawar syrianchildren syria syrianâ€¦
## [5] rt najjo this picture tell us the whole story of syriawarhttpstcojcwddand
```

```
cleanset <- tm_map(corpus,removeWords, stopwords('english'))
inspect(cleanset[1:5])
```

```
## <<SimpleCorpus>>
## Metadata: corpus specific: 1, document level (indexed): 0
## Content: documents: 5
##
## [1] suriyaoffl vigneshshivn anirudhofficial kegvraja keerthyofficialmeramyakris
hnan menongauthamâ€¦ httpstcozvbbynedr
## [2] rt cityheadlines sikh charity group khalsaaaid extends help syrian refugees
khalsaaaid \n\nsyriawar syrianchildren syria syrianâ€¦
## [3] httpstcohvkyjiots \nhights cruelty humanswanthumanity humanrightsdiedinsyr
ia savesyrianchildrenâ€¦ httpstcoeingoeos
## [4] rt cityheadlines sikh charity group khalsaaaid extends help syrian refugees
khalsaaaid \n\nsyriawar syrianchildren syria syrianâ€¦
## [5] rt najjo picture tell us whole story syriawar httpstcojcwddand
```

```
removeURL <- function(x) gsub('http[[:alnum:]]*', '', x)
cleanset <-tm_map(cleanset, content_transformer(removeURL))
inspect(cleanset[1:5])
```

```
## <<SimpleCorpus>>
## Metadata: corpus specific: 1, document level (indexed): 0
## Content: documents: 5
##
## [1] suriyaoffl vigneshshivn anirudhofficial kegvraja keerthyofficialmeramyakris
hnan menongauthamâ€¦
## [2] rt cityheadlines sikh charity group khalsaaaid extends help syrian refugees
khalsaaaid \n\nsyriawar syrianchildren syria syrianâ€¦
## [3] \nhights cruelty humanswanthumanity humanrightsdiedinsyria savesyrianchild
renâ€¦
## [4] rt cityheadlines sikh charity group khalsaaaid extends help syrian refugees
khalsaaaid \n\nsyriawar syrianchildren syria syrianâ€¦
## [5] rt najjo picture tell us whole story syriawar
```

## Remove common words and useless words

```
cleanset <- tm_map(cleanset, removeWords, c('syriawar', 'syria', 'syrian', 'ubaubfubau
bcdubaubaubc',
                                           'ubeubaubcduba', 'ubbubbfubbubcd',
                                           'eduaubdedubufeduaubdedubufeduaubdedub
uf',
                                           'eduaubdedubua', 'uauuuuuuuucc', 'uauuu',
                                           'iamka',
                                           'behindwoods'))
```

## Remove whitespace

```
cleanset <- tm_map(cleanset, stripWhitespace)
inspect(cleanset[1:5])
```

```
## <<SimpleCorpus>>
## Metadata: corpus specific: 1, document level (indexed): 0
## Content: documents: 5
##
## [1] suriyaoffl vigneshshivn anirudhofficial kegvraja keerthyofficialmeramyakris
hnan menongauthamâ€¦
## [2] rt cityheadlines sikh charity group khalsaaaid extends help refugees khalsaai
d syrianchildren syrianâ€¦
## [3] hights cruelty humanswanthumanity humanrightsdiedinsyria savesyrianchildren
â€¦
## [4] rt cityheadlines sikh charity group khalsaaaid extends help refugees khalsaai
d syrianchildren syrianâ€¦
## [5] rt najjo picture tell us whole story
```

# Converting into structured data

## Term Document Matrix

```
tdm <- TermDocumentMatrix(cleanset)
tdm
```

```
## <<TermDocumentMatrix (terms: 3096, documents: 2000)>>
## Non-/sparse entries: 16820/6175180
## Sparsity           : 100%
## Maximal term length: 143
## Weighting           : term frequency (tf)
```

```
tdm <- as.matrix(tdm)
tdm[1:10,1:20]
```

```
##              Docs
## Terms          1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
## anirudhofficial 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## keerthyofficial 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## kegvraja        1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## menongauthamâ€ 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## meramyakrishnan 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## suriyaooffl     1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## vigneshshivn    1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## charity         0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## cityheadlines   0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## extends         0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

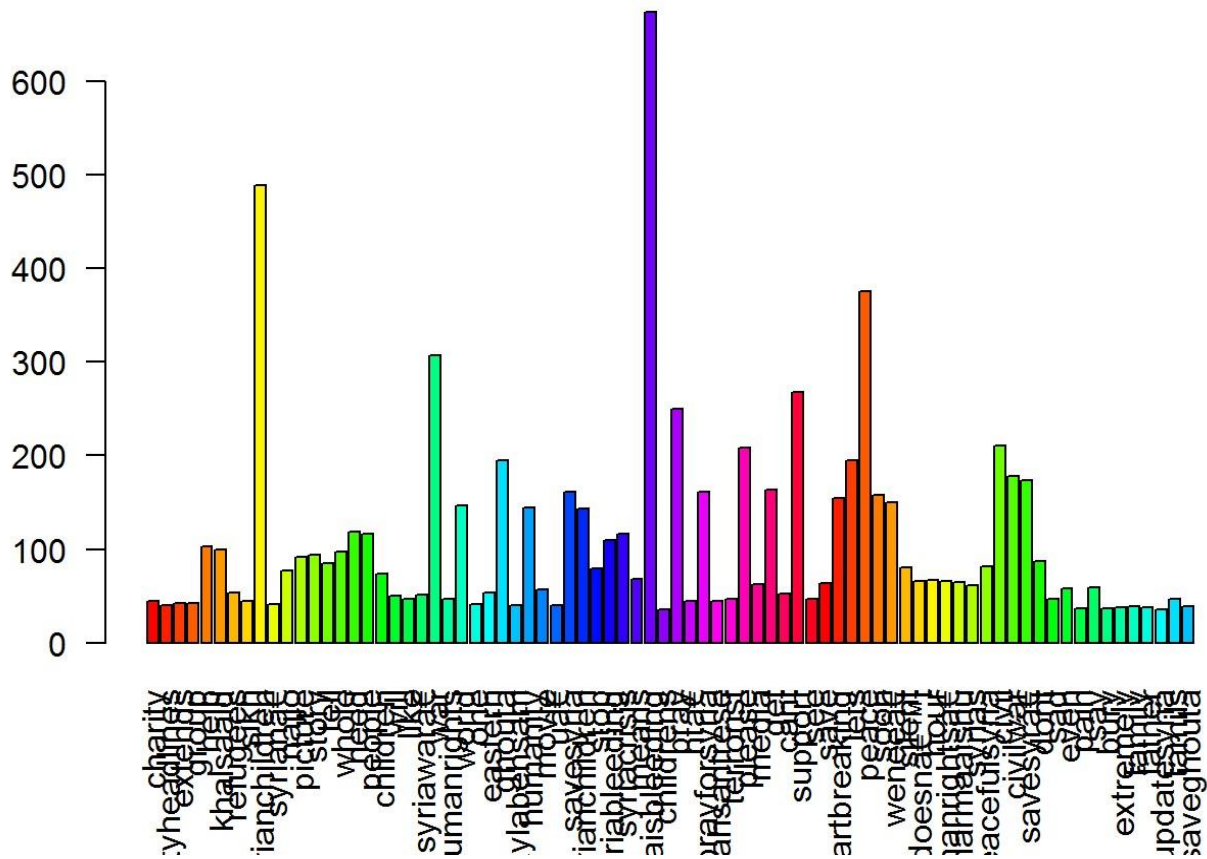
## Visualization

### Bar plots

```
w <- rowSums(tdm)
w <- subset(w, w>=35)
w
```

##	charity	cityheadlines	extends
##	44	40	42
##	group	help	khalsaaaid
##	42	103	100
##	refugees	sikh	syrianchildren
##	53	44	489
##	syrianâ€	najjo	picture
##	41	77	92
##	story	tell	whole
##	94	85	97
##	need	people	children
##	119	116	74
##	will	like	syriawarâ€
##	50	47	51
##	war	humanrights	world
##	307	47	147
##	one	eastern	ghouta
##	41	54	195
##	aylabehsam	humanity	movie
##	40	144	57
##	uâ€	savesyria	savesyrianchildren
##	40	161	143
##	stop	syriableeding	syriacrisis
##	79	110	116
##	means	syriaisbleeding	childrens
##	68	674	36
##	pray	htâ€	prayforsyria
##	250	44	161
##	sarahsantrese	terrorist	please
##	44	47	208
##	media	get	cant
##	62	163	52
##	support	see	save
##	268	47	64
##	heartbreaking	lets	peace
##	155	195	375
##	soon	weneeâ€	silent
##	158	150	80
##	doesnâ€™t	hour	humanrightsâ€
##	66	67	66
##	sharmaaishu	syrias	weneedpeacefulsyria
##	65	61	81
##	civil	civilwar	savesyrâ€
##	210	178	173
##	dont	sad	even
##	87	47	58
##	pain	say	bury
##	37	59	37
##	extremely	family	father
##	38	39	38
##	liveupdatesyria	tamils	saveghouta
##	36	47	39

```
barplot(w,
      las = 2,
      col = rainbow(50))
```



## Word cloud

```
library(wordcloud)
```

```
## Loading required package: RColorBrewer
```

```
w <- sort(rowSums(tdm), decreasing = TRUE)
set.seed(222)
wordcloud(words = names(w),
  freq=w,
  max.words = 300,
  random.order = F,
  min.freq = 20,
  colors=brewer.pal(8, 'Dark2'),
  rot.per=0.25)
```





## Wordcloud 2

```
library(wordcloud2)
w<- data.frame(names(w), w)
colnames(w) <- c('Word', 'freq')
head(w)
```

##		Word	freq
##	syriaaisbleeding	syriaaisbleeding	674
##	syrianchildren	syrianchildren	489
##	peace	peace	375
##	war	war	307
##	support	support	268
##	pray	pray	250

```
wordcloud2(w,  
            size = 0.7,  
            shape = 'W',  
            rotateRatio = 0.5,  
            minSize = 1)
```



## Sentiment Analysis

```
library(syuzhet)
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
```

```
## The following object is masked from 'package:base':
##
##     date
```

```
library(ggplot2)
```

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

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

```
library(scales)
```

```
##
## Attaching package: 'scales'
```

```
## The following object is masked from 'package:syuzhet':
##
##     rescale
```

```
library(reshape2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:lubridate':
##
##     intersect, setdiff, union
```

```
## The following objects are masked from 'package:twitterR':
##
##     id, location
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
#obtain sentiment scores
tweets <- iconv(syriawar$text, to='utf-8')
s <- get_nrc_sentiment(tweets)
head(s)
```

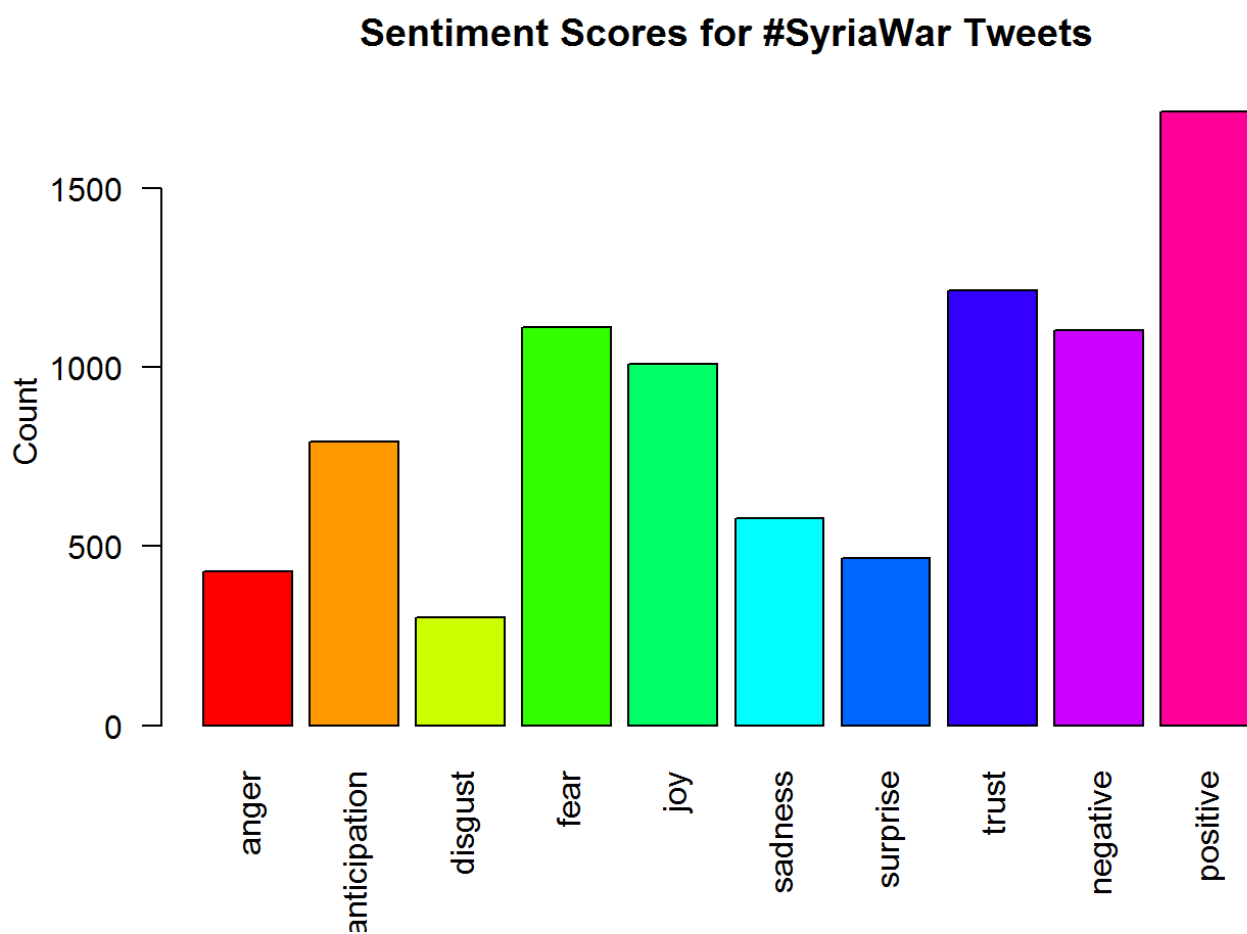
```
##      anger anticipation  disgust  fear joy  sadness  surprise  trust  negative
## 1      0             0          0   0   0          0          0      0          0
## 2      0             0          0   0   1          0          0      0          0
## 3      1             0          1   1   0          1          0      0          1
## 4      0             0          0   0   1          0          0      0          0
## 5      0             0          0   0   0          0          0      0          0
## 6      1             1          0   1   0          1          1      0          2
##      positive
## 1           0
## 2           2
## 3           0
## 4           2
## 5           0
## 6           1
```

```
tweets[2]
```

```
## [1] "RT @cityheadlines07: Sikh charity group #KhalsaAid extends help to Syrian r  
efugees @Khalsa_Aid \n\n#SyriaWar #Syrianchildren #Syria #Syrianâ€¦"
```

## bar plot

```
barplot(colSums(s),  
        las=2,  
        col = rainbow(10),  
        ylab='Count',  
        main = 'Sentiment Scores for #SyriaWar Tweets')
```



The Positive count here shows that numerous people have showed their support for Syria during this chaos. Tweets like save Syria, join hands, support Syria, etc.