Sentimentanalysis.R

#read data of 10000 tweets having “#blockchain”  
tweets<-read.csv("btweets.csv")

#load text mining library

# tm is text mining library  
library("tm")

## Warning: package 'tm' was built under R version 4.0.4

## Loading required package: NLP

#Clean up data:  
#Corpus is collection of texts  
twitterCorpus <-Corpus(VectorSource(tweets$text))  
#display 1 to 10 tweets  
inspect(twitterCorpus[1:10])

## <<SimpleCorpus>>  
## Metadata: corpus specific: 1, document level (indexed): 0  
## Content: documents: 10  
##   
## [1] RT @mastitsky: Querying rich transaction data on the #TronNetwork with my recently release R package tronr couldn't be simpler. \n#blockcha…   
## [2] #STORJ hit 5288 https://t.co/Bzd59lZFbA\n#crypto #bitcoin #blockchain #forex #cryptocurrency   
## [3] #STORJ (Binance)\nBuy zone 4677-4998\nSell zone 5248-5789-5986-6995-8975 https://t.co/XkjAwX37JK\n#crypto #bitcoin… https://t.co/Sz9qSyK4Br   
## [4] RT @cobadisec: Top 10 @owasp\n#webapplication #DevSecOps #devops \n#tech #BREAKING\n#BreakingNews #Kali #Linux #hacking #tools #Termux\n#cybers…   
## [5] RT @1centarmy: so now #blockstream will own the hashrate , the keys to the github , liquid and ln.\nhow is this not centralized.\nhow is this…   
## [6] RT @WaCo\_Token: WaCo App soon! Our team at work!<U+0001F331><U+267B><U+FE0F><U+0001F30F><U+0001F680><U+0001F680>\n#cryptocurrency #token #greeneconomy #recycling #crypto #cryptoworld #blockchain #bl…  
## [7] RT @Mazi55480281: Ready for Upcoming Events Today on @RageFanSocial ?\n Poolz IDO\nZeeDO IDO &amp; UniSwap Listing\nWatch this short video to know…   
## [8] RT @cryptomemes\_7: With #NFTs now in the spotlight, what’s next? Don’t sleep on #blockchain and #CloudComputing for it will embed #crypto w…   
## [9] You missed one more quick profit\n\nHere is one more profit for premium members in 2 hours https://t.co/1A6hPthB1C… https://t.co/Nzr6RpDe5t   
## [10] RT @RaverCrypto: $GO will 100X<U+0001F48E><U+0001F680>\n\n#GoChain is run by some of the most #reputable #organizations in the world.\n\nWorking together to make pos…

# convert the texts to lowetcase  
twitterCorpus<- tm\_map(twitterCorpus, content\_transformer(tolower))

## Warning in tm\_map.SimpleCorpus(twitterCorpus, content\_transformer(tolower)):  
## transformation drops documents

#remove stopwords such as and, the etc., language english  
twitterCorpus<- tm\_map(twitterCorpus,removeWords,stopwords("en"))

## Warning in tm\_map.SimpleCorpus(twitterCorpus, removeWords, stopwords("en")):  
## transformation drops documents

#remove numbers from tweets  
twitterCorpus<- tm\_map( twitterCorpus,removeNumbers)

## Warning in tm\_map.SimpleCorpus(twitterCorpus, removeNumbers): transformation  
## drops documents

#remove punctuations   
twitterCorpus<- tm\_map( twitterCorpus,removePunctuation)

## Warning in tm\_map.SimpleCorpus(twitterCorpus, removePunctuation): transformation  
## drops documents

#remove http urls   
removeURL<- function(x) gsub("http[[:alnum:]]\*", "", x)   
twitterCorpus<- tm\_map(twitterCorpus,content\_transformer(removeURL))

## Warning in tm\_map.SimpleCorpus(twitterCorpus, content\_transformer(removeURL)):  
## transformation drops documents

#remove remaining http functions  
removeURL<- function(x) gsub("edua[[:alnum:]]\*", "", x)   
twitterCorpus<- tm\_map(twitterCorpus,content\_transformer(removeURL))

## Warning in tm\_map.SimpleCorpus(twitterCorpus, content\_transformer(removeURL)):  
## transformation drops documents

# remove non "American standard code for information interchange (curly quotes and ellipsis)"  
# using function from package "textclean"   
removeNonAscii<-function(x) textclean::replace\_non\_ascii(x)   
twitterCorpus<-tm\_map(twitterCorpus,content\_transformer(removeNonAscii))

## Warning in tm\_map.SimpleCorpus(twitterCorpus,  
## content\_transformer(removeNonAscii)): transformation drops documents

#remove some non essential words  
twitterCorpus<- tm\_map(twitterCorpus,removeWords,c("amp","ufef",  
 "ufeft","uufefuufefuufef","uufef","s"))

## Warning in tm\_map.SimpleCorpus(twitterCorpus, removeWords, c("amp", "ufef", :  
## transformation drops documents

#remove white space  
twitterCorpus<- tm\_map(twitterCorpus,stripWhitespace)

## Warning in tm\_map.SimpleCorpus(twitterCorpus, stripWhitespace): transformation  
## drops documents

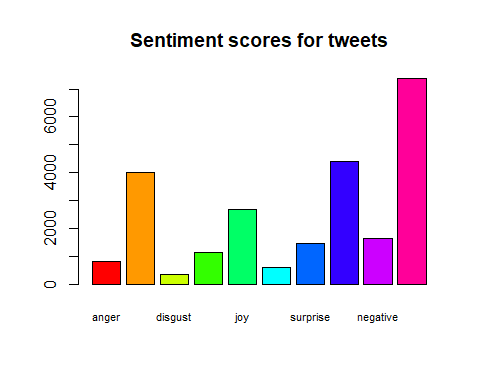
# Now inspect the cleaned tweets  
inspect(twitterCorpus[1:10])

## <<SimpleCorpus>>  
## Metadata: corpus specific: 1, document level (indexed): 0  
## Content: documents: 10  
##   
## [1] rt mastitsky querying rich transaction data tronnetwork recently release r package tronr simpler blockcha...   
## [2] storj hit crypto bitcoin blockchain forex cryptocurrency   
## [3] storj binance buy zone sell zone crypto bitcoin...   
## [4] rt cobadisec top owasp webapplication devsecops devops tech breaking breakingnews kali linux hacking tools termux cybers...   
## [5] rt centarmy now blockstream will hashrate keys github liquid ln centralized ...   
## [6] rt wacotoken waco app soon team workufubufefuffufuf cryptocurrency token greeneconomy recycling crypto cryptoworld blockchain bl...  
## [7] rt mazi ready upcoming events today ragefansocial poolz ido zeedo ido uniswap listing watch short video know...   
## [8] rt cryptomemes nfts now spotlight ' next don't sleep blockchain cloudcomputing will embed crypto w...   
## [9] missed one quick profit one profit premium members hours ...   
## [10] rt ravercrypto go will xufeuf gochain run reputable organizations world working together make pos...

#Sentiment analysis:  
#syuzhet is an sentiment analysis library  
library("syuzhet")

## Warning: package 'syuzhet' was built under R version 4.0.4

# find count of 8 emotions  
emotions<-get\_nrc\_sentiment(twitterCorpus$content)  
barplot(colSums(emotions),cex.names = .7,  
 col = rainbow(10),  
 main = "Sentiment scores for tweets"  
)



# sentiment positiviy rating  
get\_sentiment(twitterCorpus$content[1:10])

## [1] 0.90 -0.25 0.00 -0.60 0.00 0.00 1.20 0.00 -0.50 0.75

sent<-get\_sentiment(twitterCorpus$content)  
sentimentTweets<-dplyr::bind\_cols(tweets,data.frame(sent))  
  
# mean of sentiment positivity  
meanSent<-mean(sentimentTweets$sent)  
  
#display the mean of 10000 tweets   
meanSent

## [1] 0.526155