Name-Jayanta Ghosh

Roll no-190112

Write a program in R for mining Twitter to identify tweets for a specific period and identify trends and named entities.

Practical Overview

Text and Tweets Processing using R

With the increasing importance of computational text analysis in research, many researchers face the challenge of learning how to use advanced software that enables this text analysis. Currently, one of the most popular environments for computational methods and the emerging field of "data science" is the R statistical software. However, for researchers that are not well-versed in programming, learning how to use R can be a challenge, and performing text analysis in particular can seem daunting. In this step by step guide one can learn that performing text analysis in R is not as hard as some might fear.

R usage for extracting Tweets and Packages

R was specifically designed for statistical analysis, which makes it highly suitable for data science applications. Although the learning curve for programming with R can be steep, especially for people without prior programming experience, the tools now available for carrying out text analysis in R make it easy to perform powerful, cutting-edge text analytics using only a few simple commands. One of the keys to R's explosive growth has been its densely populated collection of extension software libraries, known in R terminology as packages, supplied and maintained by R's extensive user community.

Each package extends the functionality of the base R language and core packages, and in addition to functions and data must include documentation and examples, often in the form of vignettes demonstrating the use of the package. The best-known package repository, the Comprehensive R Archive Network (CRAN), currently has over 10,000 packages that are published.

Twitter Data

Twitter is an online microblogging tool that disseminates more than 400 million messages per day, including vast amounts of information about almost all industries from entertainment to sports, health to business etc. One of the best things about Twitter — indeed, perhaps its greatest appeal — is in its accessibility. It's easy to use both for sharing information and for collecting it. Twitter provides unprecedented access to our lawmakers and to our celebrities, as well as to news as it's happening. Twitter represents an important data source for the business models of huge companies as well.

All the above characteristics make twitter a best place to collect real time and latest data to analyse and do any sought of research for real life situations.

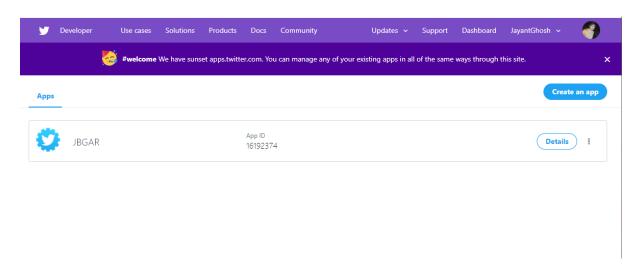
Practical Implementation

First, I will steer through my Applications of Twitter

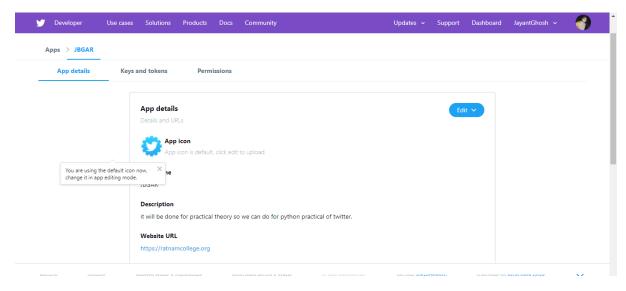
Link:- https://apps.twitter.com/

As I have performed in TYCS Practical that's why my App was already there so it will automatically open the developer

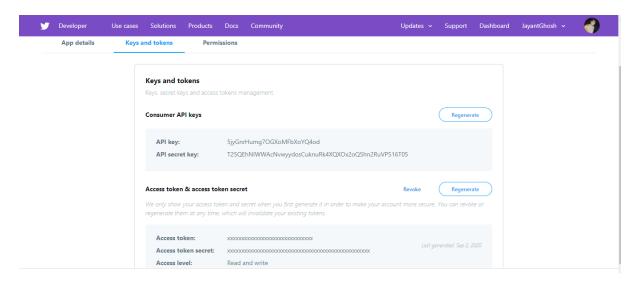
link:- https://developer.twitter.com/en/apps



As we can see my App name was which I am created was JBGAR and ID 16192374 and you can create an App it's your choice as I have already an app I don't go for it.



My App Details



My keys and tokens of Twitter Apps i can revoke your access tokens and secret keys.

Doing this steps you can access your tokens and keys after doing this all we will switch to RStudio or RGUI. I have done in RGUI.

INSTALL AND LOAD R PACKAGES:

R comes with a standard set of packages. A number of other packages are available for download and installation. For the purpose of this post, we will need the following packages:

ROAuth: Provides an interface to the OAuth 1.0 specification, allowing users to authenticate via OAuth to the server of their choice.

TwitteR: Provides an interface to the Twitter web API.

ggplot: It can be used to declare the input data frame for a graphic and to specify the set of plot aesthetics intended to be common throughout all subsequent layers unless specifically overridden.

TopicModels: package takes a Document-Term Matrix as input and produces a model that can be tided by tidytext, such that it can be manipulated and visualized with dplyr and ggplot2.

Syuzhet: A R package for the extraction of sentiment and sentiment-based plot arcs from text.

Stringi: The R package for fast, correct, consistent, portable, as well as convenient string/text processing in every locale and any native character encoding.

SnowballC: An R interface to the C 'libstemmer' library that implements Porter's word stemming algorithm for collapsing words to a common root to aid comparison of vocabulary.

Tidyr: Tidy data is data that's easy to work with: it's easy to munge (with dplyr), visualise (with ggplot2 or ggvis) and model (with R's hundreds of modelling packages).

Wordcloud: Text mining methods allow us to highlight the most frequently used keywords in a paragraph of texts. One can create a word cloud, also referred as text cloud or tag cloud, which is a visual representation of text data.

The procedure of creating word clouds is very simple in R if you know the different steps to execute. The text mining package (tm) and the word cloud generator package (wordcloud) are available in R for helping us to analyse texts and to quickly visualize the keywords as a word cloud.

tm: The tm package utilizes the Corpus as its main structure. A corpus is simply a collection of documents, but like most things in R, the corpus has specific attributes that enable certain types of analysis.

Let's start by installing and loading all the required packages.

```
install.packages("twitteR")
install.packages("ROAuth")
install.packages("tm")
install.packages("SnowballC")
install.packages("stringi")
install.packages("topicmodels")
install.packages("syuzhet")
```

```
R Console
                                                                      - - X
> install.packages("twitteR")
--- Please select a CRAN mirror for use in this session ---
also installing the dependencies 'sys', 'askpass', 'bit', 'curl', 'jsonlite', '$
 There is a binary version available but the source version is later:
     binary source needs compilation
bit64 4.0.4 4.0.5
 Binaries will be installed
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/sys 3.4.zip'
Content type 'application/zip' length 59844 bytes (58 KB)
downloaded 58 KB
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/askpass 1.1$
Content type 'application/zip' length 243567 bytes (237 KB)
downloaded 237 KB
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/bit 4.0.4.z$
Content type 'application/zip' length 634352 bytes (619 KB)
downloaded 619 KB
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/curl 4.3.zi$
Content type 'application/zip' length 4126188 bytes (3.9 MB)
```

```
package 'curl' successfully unpacked and MD5 sums checked
 package 'jsonlite' successfully unpacked and MD5 sums checked
 package 'mime' successfully unpacked and MD5 sums checked
package 'openssl' successfully unpacked and MD5 sums checked
package 'R6' successfully unpacked and MD5 sums checked
package 'bit64' successfully unpacked and MD5 sums checked
package 'rjson' successfully unpacked and MD5 sums checked
package 'DBI' successfully unpacked and MD5 sums checked
package 'httr' successfully unpacked and MD5 sums checked
package 'twitteR' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
         C:\Users\hp\AppData\Local\Temp\RtmpoNY5Cm\downloaded packages
>
> install.packages("ROAuth")
also installing the dependencies 'bitops', 'RCurl', 'digest'
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/bitops 1.0-$
Content type 'application/zip' length 38414 bytes (37 KB)
downloaded 37 KB
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/RCurl 1.98-$
Content type 'application/zip' length 3044987 bytes (2.9 MB)
downloaded 2.9 MB
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/digest 0.6.$
Content type 'application/zip' length 245013 bytes (239 KB)
downloaded 239 KB
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/ROAuth 0.9.$
Content type 'application/zip' length 133500 bytes (130 KB)
downloaded 130 KB
package 'bitops' successfully unpacked and MD5 sums checked
package 'RCurl' successfully unpacked and MD5 sums checked
package 'digest' successfully unpacked and MD5 sums checked
package 'ROAuth' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
```

package 'sys' successfully unpacked and MD5 sums checked package 'askpass' successfully unpacked and MD5 sums checked package 'bit' successfully unpacked and MD5 sums checked

```
> install.packages("tm")
 also installing the dependencies 'Rcpp', 'xml2', 'BH'
 trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/Rcpp 1.0.$
 Content type 'application/zip' length 3265833 bytes (3.1 MB)
 downloaded 3.1 MB
 trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/xml2 1.3.$
 Content type 'application/zip' length 3005723 bytes (2.9 MB)
 downloaded 2.9 MB
 trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/BH 1.72.0$
 Content type 'application/zip' length 18270708 bytes (17.4 MB)
 downloaded 17.4 MB
 trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/tm 0.7-7.$
 Content type 'application/zip' length 1553910 bytes (1.5 MB)
 downloaded 1.5 MB
 package 'Rcpp' successfully unpacked and MD5 sums checked
 package 'xml2' successfully unpacked and MD5 sums checked
 package 'BH' successfully unpacked and MD5 sums checked
 package 'tm' successfully unpacked and MD5 sums checked
 The downloaded binary packages are in
> install.packages("syuzhet")
also installing the dependencies 'assertthat', 'utf8', 'cli', 'crayon', 'fans$
  There is a binary version available but the source version is later:
      binary source needs compilation
vctrs 0.3.3 0.3.4
  Binaries will be installed
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/asserttha$
Content type 'application/zip' length 54929 bytes (53 KB)
downloaded 53 KB
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/utf8 1.1.$
Content type 'application/zip' length 208779 bytes (203 KB)
downloaded 203 KB
> install.packages("SnowballC")
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/SnowballC$
Content type 'application/zip' length 450058 bytes (439 KB)
downloaded 439 KB
package 'SnowballC' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
        C:\Users\hp\AppData\Local\Temp\RtmpoNY5Cm\downloaded packages
```

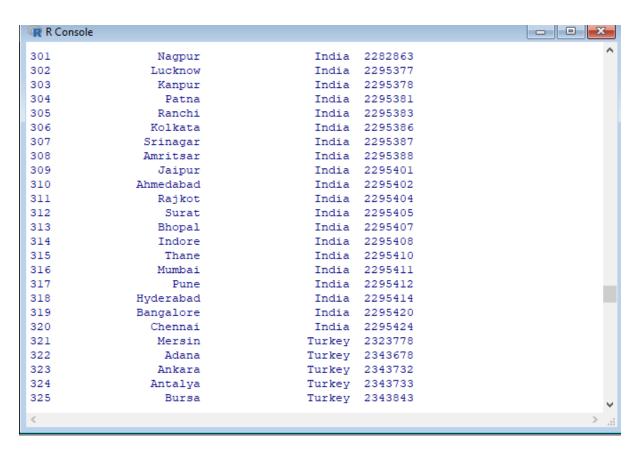
```
o:/oseis/ub/apppaca/bocai/iemp/kcmponiscm/downioaded/packages
 > install.packages("stringi")
 trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/stri
 Content type 'application/zip' length 15217051 bytes (14.5 MB)
 downloaded 14.5 MB
 package 'stringi' successfully unpacked and MD5 sums checked
 The downloaded binary packages are in
         C:\Users\hp\AppData\Local\Temp\RtmpoNY5Cm\downloaded packages
>
> install.packages("topicmodels")
also installing the dependency 'modeltools'
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/modeltool$
Content type 'application/zip' length 207116 bytes (202 KB)
downloaded 202 KB
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/topicmode$
Content type 'application/zip' length 1372952 bytes (1.3 MB)
downloaded 1.3 MB
package 'modeltools' successfully unpacked and MD5 sums checked
package 'topicmodels' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
        C:\Users\hp\AppData\Local\Temp\RtmpoNY5Cm\downloaded_packages
>
Then execute the libraries
library("twitteR")
library("tm")
library("SnowballC")
library("stringi")
library("topicmodels")
library("syuzhet")
library("ROAuth")
Connect your twitter account to R, in order to extract the required tweets.
consumer_key <- '5jyGnrHumg7OGXoMFbXoYQ4od'
consumer_secret <- 'T25QEhNIWWAcNvwyydosCuknuRk4XQXOx2oQShn2RuVPS16T05'
access_token <- '1112948760575963137-CqKgU1gxoe9QSBR6CDtLwHaMB4Krxs'
access_secret <- 'FgQB7ECr1ll7x4PmaBVy29oZUiwnNmSHo5zp7rLCL0agj'
setup_twitter_oauth(consumer_key, consumer_secret, access_token, access_secret)
```

```
> library("topicmodels")
>
> library("tm")
Loading required package: NLP
> library("NLP")
> library("syuzhet")
> library("SnowballC")
> library("stringi")
> consumer_key <- '5jyGnrHumg7OGXoMFbXoYQ4od'
> consumer_secret <- 'T25QEhNIWWAcNvwyydosCuknuRk4XQXOx2oQShn2RuVPS16T05'
> access token <- '1112948760575963137-CqKgUlgxoe9QSBR6CDtLwHaMB4Krxs'
> access secret <- 'FgQB7ECr1117x4PmaBVy29oZUiwnNmSHo5zp7rLCL0agj'
> setup twitter oauth(consumer key, consumer secret, access token, access secre$
> setup_twitter_oauth(consumer_key, consumer_secret, access_token, access_secre$
[1] "Using direct authentication"
Use a local file ('.httr-oauth'), to cache OAuth access credentials between R s$
1: Yes
2: No
Selection: Yes
Adding .httr-oauth to .gitignore
```

Trend Locations

availableTrendLocations()

```
M V COURDIE
> consumer_key <- '5jyGnrHumg70GXoMFbXoYQ4od'
> consumer_secret <- 'T25QEhNIWWAcNvwyydosCuknuRk4XQXOx2oQShn2RuVPS16T05'
> access token <- '1112948760575963137-CqKgUlgxoe9QSBR6CDtLwHaMB4Krxs'
> access_secret <- 'FgQB7ECrlll7x4PmaBVy29oZUiwnNmSHo5zp7rLCL0agj'
> setup_twitter_oauth(consumer_key, consumer_secret, access_token, access_secre$
[1] "Using direct authentication"
> availableTrendLocations()
                                     country
                                              woeid
               Worldwide
1
                Winnipeg
                                      Canada
                                                2972
3
                                                 3369
                 Ottawa
                                      Canada
4
                  Quebec
                                       Canada
                                                 3444
5
               Montreal
                                      Canada
                                                 3534
                                                 4118
6
                Toronto
                                      Canada
                Edmonton
                                       Canada
                                                 8676
                                                 8775
8
                Calgary
                                      Canada
              Vancouver
9
                                      Canada
                                                 9807
                             United Kingdom
United Kingdom
10
              Birmingham
                                                 12723
11
              Blackpool
                                                 12903
                             United Kingdom 13383
             Bournemouth
                              United Kingdom 13911
United Kingdom 13963
13
               Brighton
14
                Bristol
                             United Kingdom
                Cardiff
                                               15127
15
16
               Coventry
                              United Kingdom 17044
17
                  Derby
                               United Kingdom
                                                 18114
```



India WOEID:- 23424848

Now we will see today's trends of Twitter

ny <- getTrends(23424848)

head(ny,n = 10)

```
name
1
         #ThursdayMotivation
2
        #CBIUnfoldSSRMystery
3
4
           #NewIndiaFitIndia
5
           #BollywoodCleanup
6
               #HansrajMeena
7
             Farooq Abdullah
8 KIM TAEHYUNG WE LOVE YOU
9
                       Rafale
10
                   pillowtalk
                                                               ur1
             http://twitter.com/search?q=%23ThursdayMotivation
2
            http://twitter.com/search?q=%23CBIUnfoldSSRMystery
3
                                http://twitter.com/search?q=Gigi
4
               http://twitter.com/search?q=%23NewIndiaFitIndia
5
               http://twitter.com/search?q=%23BollywoodCleanup
6
                    http://twitter.com/search?q=%23HansrajMeena
7
             http://twitter.com/search?q=%22Farooq+Abdullah%22
8
  http://twitter.com/search?q=%22KIM+TAEHYUNG+WE+LOVE+YOU%22
9
                             http://twitter.com/search?q=Rafale
10
                         http://twitter.com/search?g=pillowtalk
Extract Tweets of a specific user
test user <- getUser('@Pritighosh27')
test user$id
test user$getDescription()
test user$getFollowersCount()
userTimeline(user = " Pritighosh27", n = 5)
milor, object test_user not round
> test user <- getUser('@PritiGhosh27')
> test user$id
[1] "1110949225381199872"
> test user$getDescription()
[1] "own rules #foodie #selfielover #army\U0001f49c #btslover\U0001f49c"
> test user$getFollowersCount()
[1] 5
> userTimeline(user = "PritiGhosh27", n = 5)
[1] "PritiGhosh27: I purple\U0001f49c you namjoona \U0001f60a\U0001f60a cutie b$
[1] "PritiGhosh27: Wish you a very very happy birthday hope your all wishes com$
[[3]]
[1] "PritiGhosh27: #NewProfilePic"
```

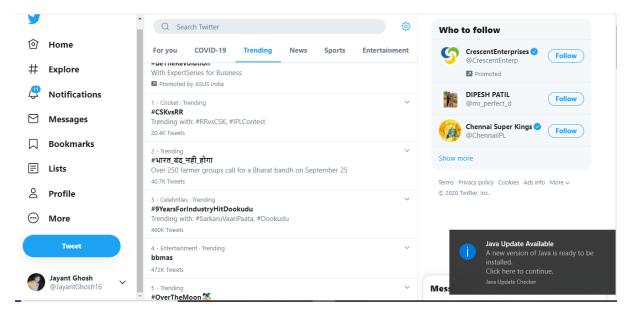
oupun Jooseele

> ny <- getTrends(23424848)</pre>

> head(ny,n = 10)

>

Now we will go to find latest twitter trends as I have performed in 22 September 2020 at 8PM.



So I have performed first four latest trends to find Data of their Tweets, Retweets and Everything.

Extracting tweets using a particular hashtag:

```
tweets m <- searchTwitter("#CSKVSRR", n=1000,lang = "en")
```

tweets_ba <- searchTwitter("#भारत_बंद_नही_होगा", n=1000,lang = "en")

tweets_d <- searchTwitter("#9YearsForIndustryHitDookudu", n=1000,lang = "en")

tweets_bt <- searchTwitter("#bbmas", n=1000,lang = "en")</pre>

Then we will Convert this extracted data to a dataframe which makes it more readable and easier to work with.

CSKRR_tweets <- twListToDF(tweets_m)

Bharat_tweets <- twListToDF(tweets_ba)

Dook_tweets <- twListToDF(tweets_d)

bbmas_tweets <- twListToDF(tweets_bt)

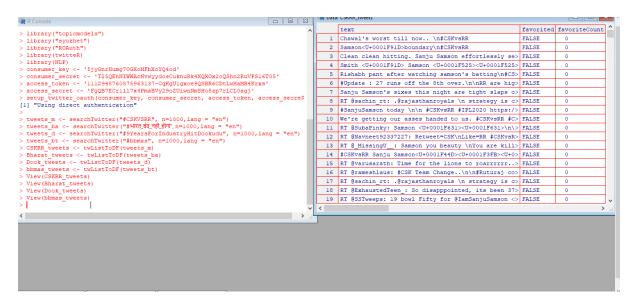
And now we will view the data of each Hashtags trends

View(CSKRR_tweets)

View(Bharat_tweets)

View(Dook_tweets)

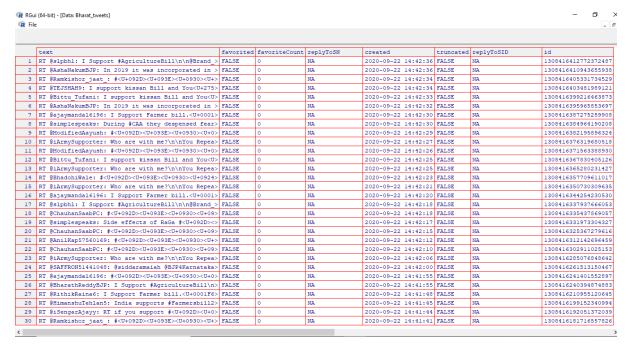
View(bbmas_tweets)



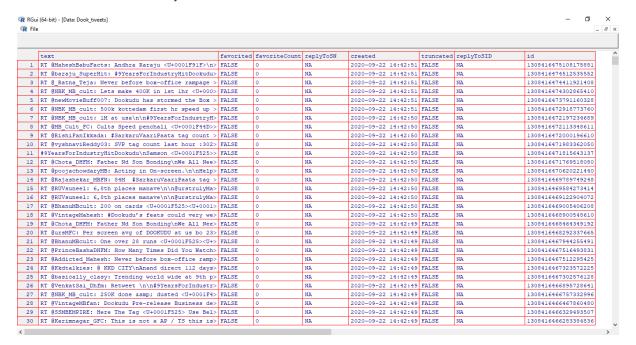
Data of CSKVSRR Trends

	text	favorited	favoriteCount	replyToSN	created	truncated	replyToSID	id
1	Chawal's worst till now \n#CSKvsRR	FALSE	0	NA	2020-09-22 14:42:26	FALSE	NA	13084163699317555
2	Samson <u+0001f91d>boundary\n#CSKvsRR</u+0001f91d>	FALSE	0	NA	2020-09-22 14:42:26	FALSE	NA	13084163697221099
3	Clean clean hitting. Sanju Samson effortlessly se>	FALSE	0	NA	2020-09-22 14:42:25	TRUE	NA	13084163675242782
4	Smith <u+0001f91d> Samson <u+0001f525><u+0001f525></u+0001f525></u+0001f525></u+0001f91d>	FALSE	0	NA	2020-09-22 14:42:25	FALSE	NA	13084163669706301
5	Rishabh pant after watching samson's batting\n#CS>	FALSE	0	NA	2020-09-22 14:42:25	FALSE	NA	13084163668867399
6	#Update : 27 runs off the 8th over.\n\nRR are hig>	FALSE	0	NA	2020-09-22 14:42:25	FALSE	NA	13084163660353167
7	Sanju Samson's sixes this night are tight slaps o>	FALSE	0	NA	2020-09-22 14:42:25	FALSE	NA	13084163658339983
8	RT @sachin_rt: .@rajasthanroyals \n strategy is c>	FALSE	0	NA	2020-09-22 14:42:25	FALSE	NA	13084163654061015
9	#SanjuSamson today \n\n #CSKvsRR #IPL2020 https:/>	FALSE	0	NA	2020-09-22 14:42:25	FALSE	NA	13084163649154375
10	We're getting our asses handed to us. #CSKvsRR #C>	FALSE	0	NA	2020-09-22 14:42:25	FALSE	NA	13084163639255695
11	RT @SubaPinky: Samson <u+0001f631><u+0001f631>\n\></u+0001f631></u+0001f631>	FALSE	0	NA	2020-09-22 14:42:24	FALSE	NA	13084163633761075
12	RT @Navneet92337227: Retweet=CSK\nLike=RR #CSKvsR>	FALSE	0	NA	2020-09-22 14:42:24	FALSE	NA	13084163633341849
13	RT @_MissingU: Samson you beauty \nYou are kill>	FALSE	0	NA	2020-09-22 14:42:24	FALSE	NA	13084163618534850
14	#CSKvsRR Sanju Samson <u+0001f44d><u+0001f3fb><u+0></u+0></u+0001f3fb></u+0001f44d>	FALSE	0	NA	2020-09-22 14:42:24	FALSE	NA	13084163606245990
15	RT @varusarath: Time for the lions to roarrrrrr>	FALSE	0	NA	2020-09-22 14:42:23	FALSE	NA	13084163593411420
16	RT @rameshlaus: #CSK Team Change\n\n#Ruturaj co>	FALSE	0	NA	2020-09-22 14:42:23	FALSE	NA	13084163574998876
17	RT @sachin_rt: .@rajasthanroyals \n strategy is c>	FALSE	0	NA	2020-09-22 14:42:21	FALSE	NA	1308416350533095
18	RT @ExhaustedTeen_: So disapppointed, its been 37>	FALSE	0	NA	2020-09-22 14:42:21	FALSE	NA	13084163489601945
19	RT @SSTweeps: 19 bowl Fifty for @IamSanjuSamson <>	FALSE	0	NA	2020-09-22 14:42:21	FALSE	NA	13084163487547678
20	RT @sachin_rt: .@rajasthanroyals \n strategy is c>	FALSE	0	NA	2020-09-22 14:42:21	FALSE	NA	13084163478781665
21	RT @Anu_rag_Singh_: #CSKvsRR\nWhen You Hear Fake >	FALSE	0	NA	2020-09-22 14:42:20	FALSE	NA	13084163452902768
22	Sanju Samson is not hitting sixes against CSK, ra>	FALSE	0	NA	2020-09-22 14:42:20	TRUE	NA	13084163450050682
23	RT @sachin_rt: .@rajasthanroyals \n strategy is c>	FALSE	0	NA	2020-09-22 14:42:19	FALSE	NA	13084163427821035
24	IPL Match 4\n\nRR 96/1 in 8 Overs\n\nSamson &>	FALSE	0	NA	2020-09-22 14:42:19	TRUE	NA	13084163412930437
25	#RR Dealing with sixes <u+0001f525><u+0001f525>></u+0001f525></u+0001f525>	FALSE	0	NA	2020-09-22 14:42:19	FALSE	NA	13084163391371796
26	@rajasthanroyals Sanju Anna Class with Pure Mass<>	FALSE	0	rajasthanroyals	2020-09-22 14:42:18	FALSE	NA	13084163360166133
27	RT @irshad5005: #SanjuSamson is becoming a #IPL 1>	FALSE	0	NA	2020-09-22 14:42:18	FALSE	NA	13084163356433735
28	Free Hitttt\n\nSixxxxxx <u+0001f525><u+0001f525><u></u></u+0001f525></u+0001f525>	FALSE	0	NA	2020-09-22 14:42:18	FALSE	NA	13084163354504519
29	Unbelievable Wow!\n#CSKvsRR #sanju	FALSE	0	NA	2020-09-22 14:42:17	FALSE	NA	13084163337265725
30	Smith <u+0001f633><u+0001f494><u+0001f494>\n#CSKv></u+0001f494></u+0001f494></u+0001f633>	FALSE	0	NA	2020-09-22 14:42:17	FALSE	NA	13084163336552734

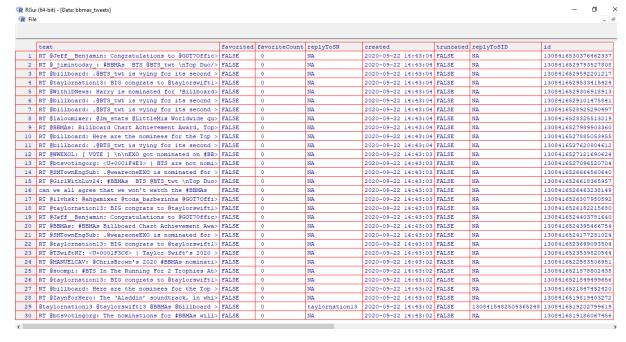
Data of भारत_बंद_नही_होगा trends



Data on 9YearsForIndustryHitDookudu trends



Data on bbmas trends



After extracting and carried out data of each trends we will go for Wordcloud to analyse the words used in trends of how much amount.

WordCloud

Word clouds (also known as text clouds or tag clouds) work in a simple way: the more a specific word appears in a source of textual data (such as a speech, blog post, or database), the bigger and bolder it appears in the word cloud.

So let's generate some word clouds and find out some of the frequent and important terms being used in the tweets we have extracted.

For that we have to first install wordcloud package and import he library

Install.packages("wordcloud")

Library(wordcloud)

```
> install.packages("Wordcloud")
--- Please select a CRAN mirror for use in this session ---
Warning messages:
1: package 'Wordcloud' is not available (for R version 4.0.2)
2: Perhaps you meant 'wordcloud' ?
> install.packages("wordcloud")
also installing the dependency 'RColorBrewer'
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/RColorBrewe$
Content type 'application/zip' length 55583 bytes (54 KB)
downloaded 54 KB
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/wordcloud 2$
Content type 'application/zip' length 785050 bytes (766 KB)
downloaded 766 KB
package 'RColorBrewer' successfully unpacked and MD5 sums checked
package 'wordcloud' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
        C:\Users\hp\AppData\Local\Temp\RtmpgrBGJa\downloaded packages
> install.packages("wordcloud"
+ )
--- Please select a CRAN mirror for use in this session ---
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/wordcloud 2$
Content type 'application/zip' length 785045 bytes (766 KB)
downloaded 766 KB
package 'wordcloud' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
       C:\Users\hp\AppData\Local\Temp\RtmpeyQN8S\downloaded packages
> library(wordcloud)
Loading required package: RColorBrewer
> library(RColorBrewer
>
Then we will go for code
CSKRR_text<- CSKRR_tweets$text
Bharat text<- Bharat tweets$text
Dook text<-Dook tweets$text
bbmas text<-bbmas tweets$text
CSKRR tweets.text.corpus<-CSKRR text
Bharat tweets.text.corpus<-Bharat text
Dook tweets.text.corpus<-Dook text
bbmas_tweets.text.corpus<-bbmas_text
```

```
> CSKRR_text<- CSKRR_tweets$text
> Bharat_text<- Bharat_tweets$text
> Dook_text<-Dook_tweets$text
> bbmas_text<-bbmas_tweets$text
>
> CSKRR_tweets.text.corpus<-CSKRR_text
> Bharat_tweets.text.corpus<-Bharat_text
> Dook_tweets.text.corpus<-Dook_text
> bbmas_tweets.text.corpus<-bbmas_text</pre>
```

wordcloud(CSKRR_tweets.text.corpus,min.freq = 10,colors=brewer.pal(8, "Dark2"),random.color = TRUE,max.words = 500)

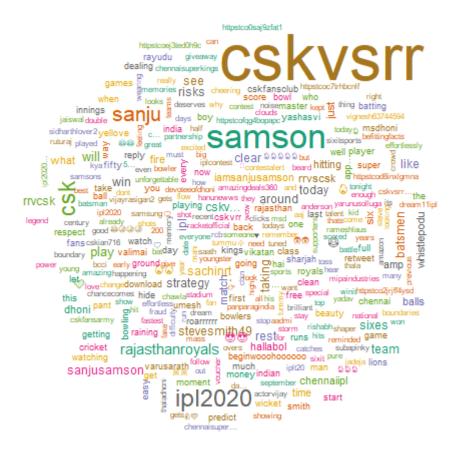
wordcloud(Bharat_tweets.text.corpus,min.freq = 10,colors=brewer.pal(8, "Dark2"),random.color = TRUE,max.words = 500)

wordcloud(Dook_tweets.text.corpus,min.freq = 10,colors=brewer.pal(8, "Dark2"),random.color = TRUE,max.words = 500)

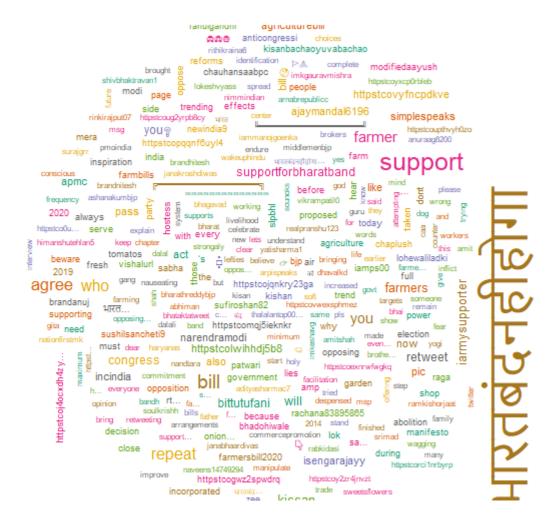
wordcloud(bbmas_tweets.text.corpus,min.freq = 10,colors=brewer.pal(8, "Dark2"),random.color = TRUE,max.words = 500)

```
> wordcloud(CSKRR_tweets.text.corpus,min.freq = 5,colors=brewer.pal(8, "Dark2")$
Warning messages:
1: In tm_map.SimpleCorpus(corpus, tm::removePunctuation) :
    transformation drops documents
2: In tm_map.SimpleCorpus(corpus, function(x) tm::removeWords(x, tm::stopwords($
    transformation drops documents
> wordcloud(Bharat_tweets.text.corpus,min.freq = 5,colors=brewer.pal(8, "Dark2"$
Warning messages:
1: In tm_map.SimpleCorpus(corpus, tm::removePunctuation) :
    transformation drops documents
2: In tm_map.SimpleCorpus(corpus, function(x) tm::removeWords(x, tm::stopwords($
    transformation drops documents
```

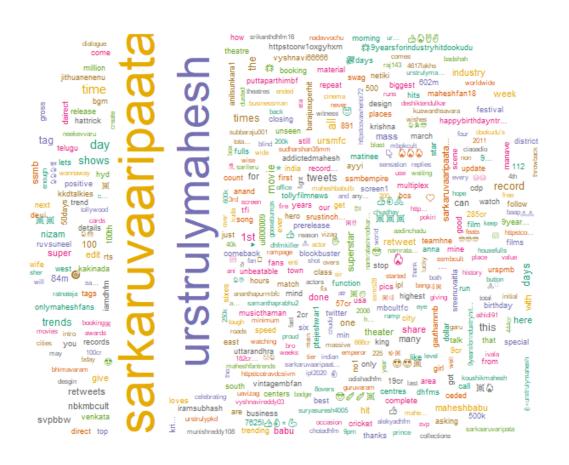
```
> wordcloud(Dook_tweets.text.corpus,min.freq = 5,colors=brewer.pal(8, "Dark2"),$
Warning messages:
1: In tm_map.SimpleCorpus(corpus, tm::removePunctuation) :
    transformation drops documents
2: In tm_map.SimpleCorpus(corpus, function(x) tm::removeWords(x, tm::stopwords($
    transformation drops documents
3: In wordcloud(Dook_tweets.text.corpus, min.freq = 5, colors = brewer.pal(8, :
        9yearsforindustryhitdookudu could not be fit on page. It will not be plotted.
> wordcloud(bbmas_tweets.text.corpus,min.freq = 5,colors=brewer.pal(8, "Dark2")$
Warning messages:
1: In tm_map.SimpleCorpus(corpus, tm::removePunctuation) :
    transformation drops documents
2: In tm_map.SimpleCorpus(corpus, function(x) tm::removeWords(x, tm::stopwords($
    transformation drops documents
```



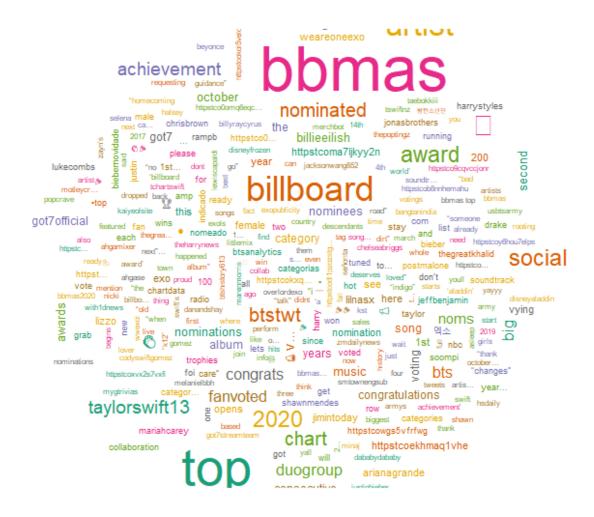
Worcloud of भारत_बंद_नही_होगा trend



Worcloud of 9YearsForIndustryHitDookudu trend



Wordcloud of bbmas trend



Sentiment Analysis

Now we will go for Sentiment Analysis for finding emotions psychology of human being through twitter data how people react to trends of their nature.

In recent years there has been a steady increase in interest from brands, companies and researchers in Sentiment Analysis and its application to business analytics. The business world today, as is the case in many data analytics streams, are looking for "business insight."

In relation to sentiment analysis, I am talking about insights into consumer behavior, what customers want, what are customers like and dislike about the products, what their buying signals are, what their decision process looks like etc because in the end of the its the customers for whose satisfaction these businesses work for I have used the inbuilt sentiment analyser in R, which uses the NRC sentiment dictionary to calculate the presence of eight different emotions and their corresponding valence in a text.

For to analyse sentiments we have to install tidyr package and import the libraries

Getting emotions using in-built function

```
mysentiment_CSKRR<-get_nrc_sentiment((CSKRR_text))
mysentiment_Bharat<-get_nrc_sentiment((Bharat_text))
mysentiment_Dook<-get_nrc_sentiment((Dook_text))
mysentiment_bbmas<-get_nrc_sentiment((bbmas_text))</pre>
```

Calculating total score for each sentiment

```
Sentimentscores_CSKRR<-data.frame(colSums(mysentiment_CSKRR[,]))
Sentimentscores_Bharat<-data.frame(colSums(mysentiment_Bharat[,]))
Sentimentscores_Dook<-data.frame(colSums(mysentiment_Dook[,]))
Sentimentscores_bbmas<-data.frame(colSums(mysentiment_bbmas[,]))

names(Sentimentscores_CSKRR)<-"Score"
Sentimentscores_CSKRR<-
cbind("sentiment"=rownames(Sentimentscores_CSKRR),Sentimentscores_CSKRR)
rownames(Sentimentscores_CSKRR)<-NULL

names(Sentimentscores_Bharat)<-"Score"
Sentimentscores_Bharat<-
cbind("sentiment"=rownames(Sentimentscores_Bharat),Sentimentscores_Bharat)
rownames(Sentimentscores_Bharat)<-NULL
```

```
Sentimentscores_Dook<-
cbind("sentiment"=rownames(Sentimentscores_Dook),Sentimentscores_Dook)
rownames(Sentimentscores_Dook)<-NULL

names(Sentimentscores_bbmas)<-"Score"

Sentimentscores_bbmas<-
cbind("sentiment"=rownames(Sentimentscores_bbmas),Sentimentscores_bbmas)
rownames(Sentimentscores_bbmas)<-NULL
```

```
> library(tidyr)
> mysentiment CSKRR<-get nrc sentiment((CSKRR text))
> mysentiment_Bharat<-get_nrc_sentiment((Bharat_text))
> mysentiment Dook<-get nrc sentiment((Dook text))
> mysentiment_bbmas<-get_nrc_sentiment((bbmas_text))
> Sentimentscores_CSKRR<-data.frame(colSums(mysentiment CSKRR[,]))
> Sentimentscores_Bharat<-data.frame(colSums(mysentiment Bharat[,]))
> Sentimentscores Dook<-data.frame(colSums(mysentiment Dook[,]))
> Sentimentscores bbmas<-data.frame(colSums(mysentiment bbmas[,]))
> names(Sentimentscores CSKRR)<-"Score"
> Sentimentscores CSKRR<-cbind("sentiment"=rownames(Sentimentscores CSKRR),Sent$
> rownames(Sentimentscores_CSKRR)<-NULL
> names(Sentimentscores_Bharat)<-"Score"
> Sentimentscores Bharat<-cbind("sentiment"=rownames(Sentimentscores Bharat),Se$
> rownames(Sentimentscores Bharat)<-NULL
> names(Sentimentscores Dook)<-"Score"
> Sentimentscores Dook<-cbind("sentiment"=rownames(Sentimentscores Dook),Sentim$
> rownames(Sentimentscores_Dook)<-NULL
> names(Sentimentscores bbmas)<-"Score"
> Sentimentscores bbmas<-cbind("sentiment"=rownames(Sentimentscores bbmas),Sent$
> rownames(Sentimentscores bbmas)<-NULL
```

Then we will install the package ggplot2 as newer version has ggplot2 we will download that and import the library

```
R Console
                                                                       ---
> install.packages("ggplot2")
also installing the dependencies 'ps', 'processx', 'callr', 'prettyunits', 'bac$
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/ps 1.3.4.zi$
Content type 'application/zip' length 702002 bytes (685 KB)
downloaded 685 KB
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/processx 3.$
Content type 'application/zip' length 1378554 bytes (1.3 MB)
downloaded 1.3 MB
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/callr 3.4.4$
Content type 'application/zip' length 377308 bytes (368 KB)
downloaded 368 KB
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/prettyunits$
Content type 'application/zip' length 37670 bytes (36 KB)
downloaded 36 KB
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/backports 1$
Content type 'application/zip' length 90803 bytes (88 KB)
downloaded 88 KB
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/desc 1.2.0.$
Content type 'application/zip' length 289205 bytes (282 KB)
```

Plotting the sentiments with scores

```
ggplot(data=Sentimentscores_CSKRR,aes(x=sentiment,y=Score))+geom_bar(aes(fill=sentiment),stat = "identity")+
theme(legend.position="none")+
xlab("Sentiments")+ylab("scores")+ggtitle("Sentiments of people behind the tweets on CSKVSRR Match")
ggplot(data=Sentimentscores_Bharat,aes(x=sentiment,y=Score))+geom_bar(aes(fill=sentiment),stat = "identity")+
theme(legend.position="none")+
xlab("Sentiments")+ylab("scores")+ggtitle("Sentiments of people behind the tweets on भारत_बंद_नही_होगा")
ggplot(data=Sentimentscores_Dook,aes(x=sentiment,y=Score))+geom_bar(aes(fill=sentiment),stat = "identity")+
theme(legend.position="none")+
xlab("Sentiments")+ylab("scores")+ggtitle("Sentiments of people behind the tweets on 9YearsForIndustryHitDookudu")
```

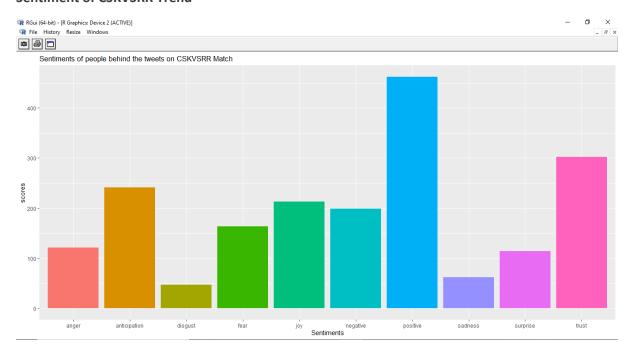
ggplot(data=Sentimentscores_bbmas,aes(x=sentiment,y=Score))+geom_bar(aes(fill=sentiment),stat
= "identity")+

theme(legend.position="none")+

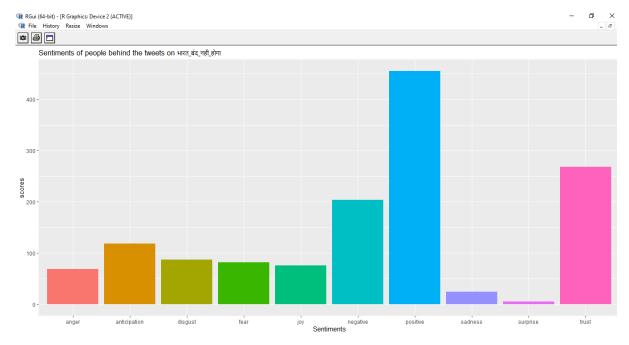
xlab("Sentiments")+ylab("scores")+ggtitle("Sentiments of people behind the tweets on BillBoard Music Awards 2020")

```
> library(ggplot2)
Attaching package: 'ggplot2'
The following object is masked from 'package:NLP':
    annotate
> ggplot(data=Sentimentscores CSKRR,aes(x=sentiment,y=Score))+geom bar(aes(fill$
    theme(legend.position="none")+
    xlab("Sentiments")+ylab("scores")+ggtitle("Sentiments of people behind the $
> ggplot(data=Sentimentscores Bharat,aes(x=sentiment,y=Score))+geom bar(aes(fil$
    theme(legend.position="none")+
    xlab("Sentiments")+ylab("scores")+ggtitle("Sentiments of people behind the $
> ggplot(data=Sentimentscores Bharat, aes(x=sentiment, y=Score))+geom bar(aes(fil$
    theme(legend.position="none")+
    xlab("Sentiments")+ylab("scores")+ggtitle("Sentiments of people behind the $
> ggplot(data=Sentimentscores_Dook,aes(x=sentiment,y=Score))+geom_bar(aes(fill=$
    theme(legend.position="none")+
    xlab("Sentiments")+ylab("scores")+ggtitle("Sentiments of people behind the $
> ggplot(data=Sentimentscores bbmas,aes(x=sentiment,y=Score))+geom bar(aes(fill$
    theme(legend.position="none")+
    xlab("Sentiments")+ylab("scores")+ggtitle("Sentiments of people behind the $
>
```

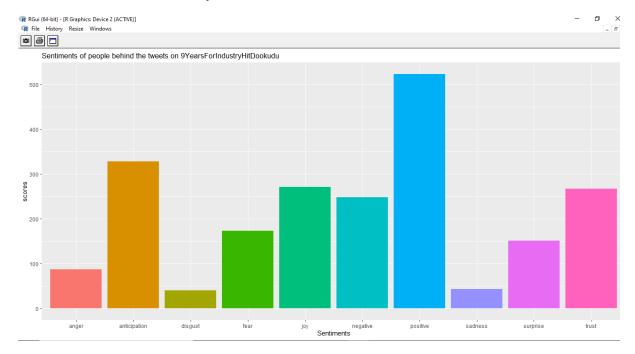
Sentiment of CSKVSRR Trend



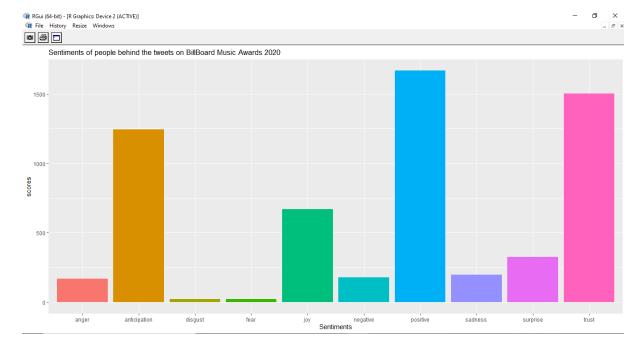
Sentiment of भारत_बंद_नही_होगा Trend



Sentiment of 9YearsForIndustryHitDookudu Trend



Sentiment of bbmas Trend



As i have completed with latest trends now we will execute specific trends which of my interest I have done with four dimensions means first trend I have done its related to Sport i.e IPL2020 Biggest trending topic, second I have done with Covid-19 Pandemic trend in which is we are facing for 6-7 months and its always buzzing topic, third I have gone with K-Pop Music which is now trending very drastically i.e BTS Dynamite Song and last have done trend of Biggest Superstar in world Shahrukh Khan.

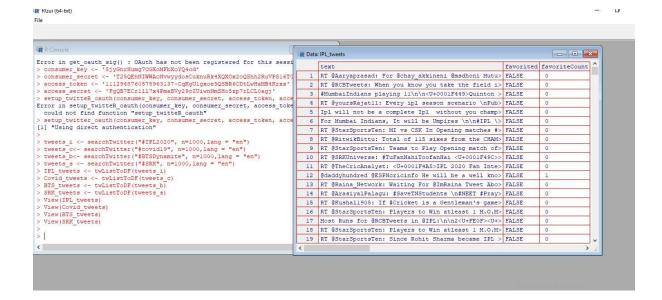
So for Specific trends:-

View(SRK tweets)

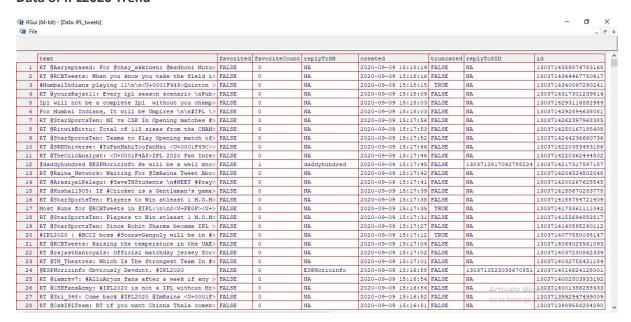
```
tweets_i <- searchTwitter("#IPL2020", n=1000,lang = "en")
tweets_c <- searchTwitter("#Covid19", n=1000,lang = "en")
tweets_b <- searchTwitter("#BTSDynamite", n=1000,lang = "en")
tweets_s <- searchTwitter("#SRK", n=1000,lang = "en")

IPL_tweets <- twListToDF(tweets_i)
Covid_tweets <- twListToDF(tweets_c)
BTS_tweets <- twListToDF(tweets_b)
SRK_tweets <- twListToDF(tweets_s)

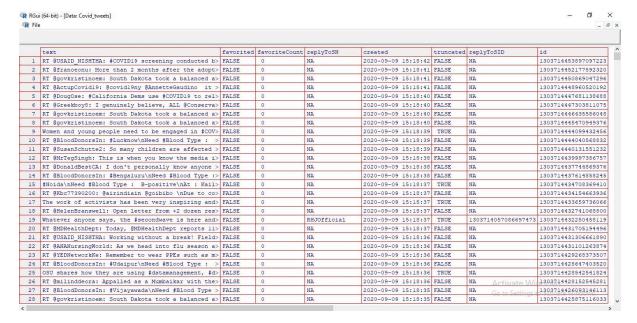
View(IPL_tweets)
View(Covid_tweets)
View(BTS_tweets)</pre>
```



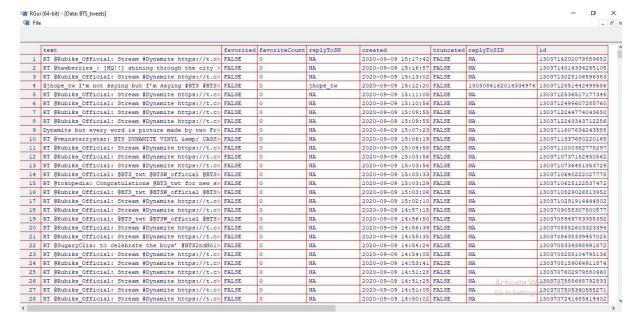
Data of IPL2020 Trend



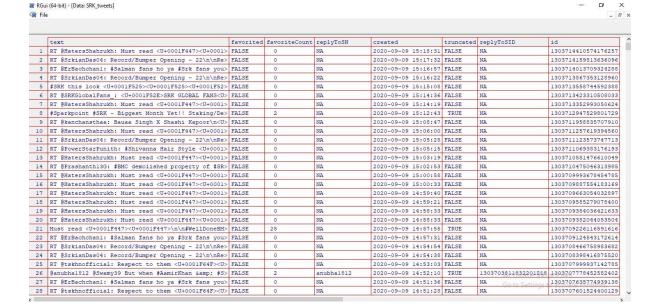
Data of Covid-19 Trend



Data of BTS Dynamite Trend



Data of SRK Trend



Wordcloud

IPL_text<- IPL_tweets\$text</pre>

Covid_text<- Covid_tweets\$text

```
BTS_text<-BTS_tweets$text
SRK_text<- SRK_tweets$text
IPL_text<-IPL_tweets.text.corpus
IPL tweets.text.corpus<-IPL text
Covid_tweets.text.corpus<-Covid_text
BTS_tweets.text.corpus<-BTS_text
SRK tweets.text.corpus<-SRK text
> CSKRR text<- CSKRR tweets$text
> Bharat text<- Bharat tweets$text
> Dook text<-Dook tweets$text
> bbmas_text<-bbmas_tweets$text
>
> CSKRR tweets.text.corpus<-CSKRR text
> Bharat_tweets.text.corpus<-Bharat_text
> Dook_tweets.text.corpus<-Dook_text
> bbmas_tweets.text.corpus<-bbmas_text
>
```

wordcloud(IPL_tweets.text.corpus,min.freq = 10,colors=brewer.pal(8, "Dark2"),random.color = TRUE,max.words = 500)

wordcloud(Covid_tweets.text.corpus,min.freq = 10,colors=brewer.pal(8, "Dark2"),random.color = TRUE,max.words = 500)

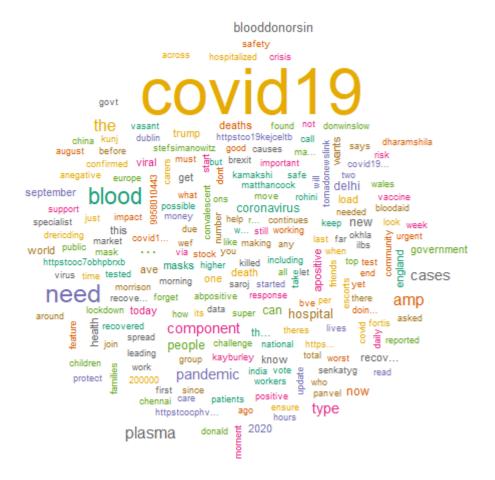
wordcloud(BTS_tweets.text.corpus,min.freq = 10,colors=brewer.pal(8, "Dark2"),random.color = TRUE,max.words = 500)

wordcloud(SRK_tweets.text.corpus,min.freq = 10,colors=brewer.pal(8, "Dark2"),random.color = TRUE,max.words = 500)

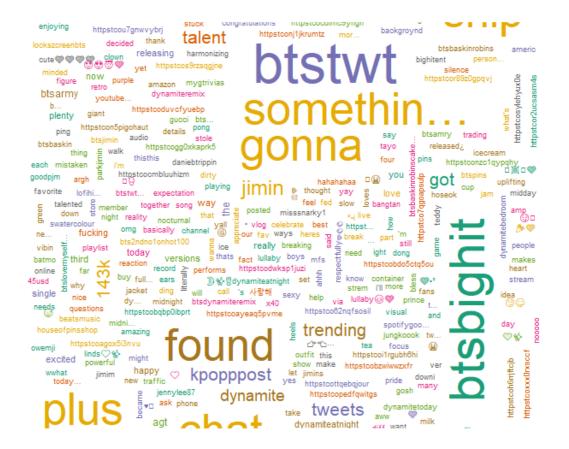
```
> IPL tweets.text.corpus<-IPL text
> Covid tweets.text.corpus<-Covid text
> BTS tweets.text.corpus<-BTS text
> SRK tweets.text.corpus<-SRK text
> wordcloud(IPL tweets.text.corpus,min.freq = 10,colors=brewer.pal(8, "Dark2"),$
Warning messages:
1: In tm map.SimpleCorpus(corpus, tm::removePunctuation) :
  transformation drops documents
2: In tm map.SimpleCorpus(corpus, function(x) tm::removeWords(x, tm::stopwords($
  transformation drops documents
> wordcloud(Covid tweets.text.corpus,min.freq = 10,colors=brewer.pal(8, "Dark2"$
Warning messages:
1: In tm_map.SimpleCorpus(corpus, tm::removePunctuation) :
  transformation drops documents
2: In tm map.SimpleCorpus(corpus, function(x) tm::removeWords(x, tm::stopwords($
  transformation drops documents
> wordcloud(BTS tweets.text.corpus,min.freq = 10,colors=brewer.pal(8, "Dark2"),$
Warning messages:
1: In tm map.SimpleCorpus(corpus, tm::removePunctuation) :
  transformation drops documents
2: In tm map.SimpleCorpus(corpus, function(x) tm::removeWords(x, tm::stopwords($
  transformation drops documents
> wordcloud(BTS tweets.text.corpus,min.freq = 5,colors=brewer.pal(8, "Dark2"),r$
Warning messages:
1: In tm map.SimpleCorpus(corpus, tm::removePunctuation) :
K Console
                                                                      1: In tm map.SimpleCorpus(corpus, tm::removePunctuation) :
  transformation drops documents
2: In tm map.SimpleCorpus(corpus, function(x) tm::removeWords(x, tm::stopwords($
  transformation drops documents
> wordcloud(BTS tweets.text.corpus,min.freq = 2,colors=brewer.pal(8, "Dark2"),r$
Warning messages:
1: In tm_map.SimpleCorpus(corpus, tm::removePunctuation) :
  transformation drops documents
2: In tm map.SimpleCorpus(corpus, function(x) tm::removeWords(x, tm::stopwords($
  transformation drops documents
3: In wordcloud(BTS_tweets.text.corpus, min.freq = 2, colors = brewer.pal(8, :
  btsbighit could not be fit on page. It will not be plotted.
> wordcloud(BTS_tweets.text.corpus,min.freq = 0,colors=brewer.pal(8, "Dark2"),r$
There were 34 warnings (use warnings() to see them)
> wordcloud(SRK tweets.text.corpus,min.freq = 10,colors=brewer.pal(8, "Dark2"),$
Warning messages:
1: In tm map.SimpleCorpus(corpus, tm::removePunctuation) :
  transformation drops documents
2: In tm map.SimpleCorpus(corpus, function(x) tm::removeWords(x, tm::stopwords($
  transformation drops documents
3: In wordcloud(SRK_tweets.text.corpus, min.freq = 10, colors = brewer.pal(8, :
  iamsrk could not be fit on page. It will not be plotted.
> wordcloud(SRK tweets.text.corpus,min.freq = 4,colors=brewer.pal(8, "Dark2"),r$
There were 43 warnings (use warnings() to see them)
>
```



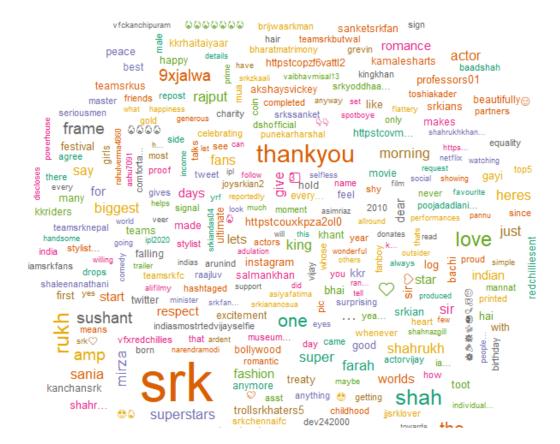
Wordcloud of Covid-19 Trend



Wordcloud of BTS Dynamite Trend



Wordcloud of SRK Trend



Sentiment Analysis

```
mysentiment_IPL<-get_nrc_sentiment((IPL_text))

mysentiment_Covid<-get_nrc_sentiment((Covid_text))

mysentiment_BTS<-get_nrc_sentiment((BTS_text))

mysentiment_SRK<-get_nrc_sentiment((SRK_text))

Sentimentscores_IPL<-data.frame(colSums(mysentiment_IPL[,]))

Sentimentscores_Covid<-data.frame(colSums(mysentiment_Covid[,]))

Sentimentscores_BTS<-data.frame(colSums(mysentiment_BTS[,]))

Sentimentscores_SRK<-data.frame(colSums(mysentiment_SRK[,]))
```

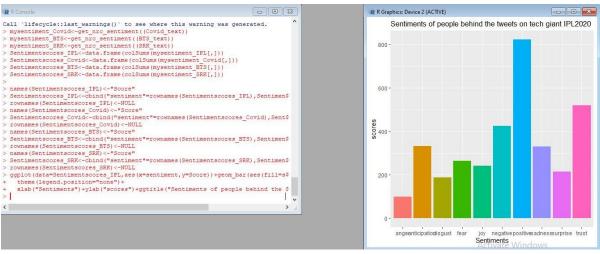
```
names(Sentimentscores_IPL)<-"Score"
Sentimentscores_IPL<-cbind("sentiment"=rownames(Sentimentscores_IPL),Sentimentscores_IPL)
rownames(Sentimentscores_IPL)<-NULL
names(Sentimentscores_Covid)<-"Score"
Sentimentscores_Covid<-
cbind("sentiment"=rownames(Sentimentscores Covid), Sentimentscores Covid)
rownames(Sentimentscores Covid)<-NULL
names(Sentimentscores BTS)<-"Score"
Sentimentscores_BTS<-cbind("sentiment"=rownames(Sentimentscores_BTS),Sentimentscores_BTS)
rownames(Sentimentscores_BTS)<-NULL
names(Sentimentscores_SRK)<-"Score"
Sentimentscores_SRK<-cbind("sentiment"=rownames(Sentimentscores_SRK),Sentimentscores_SRK)
rownames(Sentimentscores_SRK)<-NULL
ggplot(data=Sentimentscores_IPL,aes(x=sentiment,y=Score))+geom_bar(aes(fill=sentiment),stat =
"identity")+
theme(legend.position="none")+
xlab("Sentiments")+ylab("scores")+ggtitle("Sentiments of people behind the tweets on IPL2020")
ggplot(data=Sentimentscores_Covid,aes(x=sentiment,y=Score))+geom_bar(aes(fill=sentiment),stat =
"identity")+
theme(legend.position="none")+
xlab("Sentiments")+ylab("scores")+ggtitle("Sentiments of people behind the tweets on Covid-19")
ggplot(data=Sentimentscores_BTS,aes(x=sentiment,y=Score))+geom_bar(aes(fill=sentiment),stat =
"identity")+
theme(legend.position="none")+
xlab("Sentiments")+ylab("scores")+ggtitle("Sentiments of people behind the tweets on BTS
Dynamite")
```

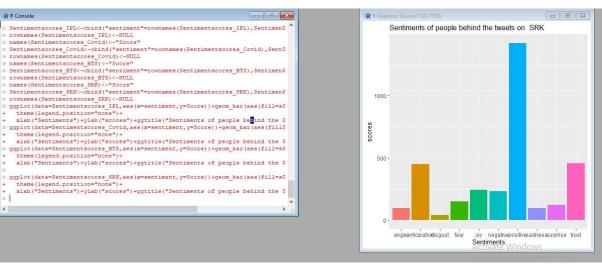
ggplot(data=Sentimentscores_SRK,aes(x=sentiment,y=Score))+geom_bar(aes(fill=sentiment),stat =
"identity")+

theme(legend.position="none")+

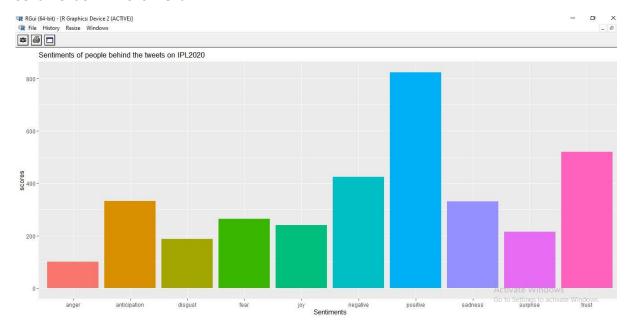
xlab("Sentiments")+ylab("scores")+ggtitle("Sentiments of people behind the tweets on SRK")

```
> mysentiment IPL<-get nrc sentiment((IPL text))
Warning messages:
1: `filter ()` is deprecated as of dplyr 0.7.0.
Please use `filter()` instead.
See vignette('programming') for more help
This warning is displayed once every 8 hours.
Call `lifecycle::last warnings()` to see where this warning was generated.
2: `group by ()` is deprecated as of dplyr 0.7.0.
Please use `group_by()` instead.
See vignette('programming') for more help
This warning is displayed once every 8 hours.
Call `lifecycle::last_warnings()` to see where this warning was generated.
3: `data frame()` is deprecated as of tibble 1.1.0.
Please use `tibble()` instead.
This warning is displayed once every 8 hours.
Call `lifecycle::last warnings()` to see where this warning was generated.
```

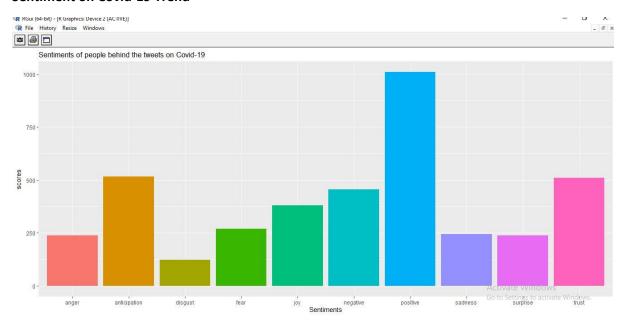




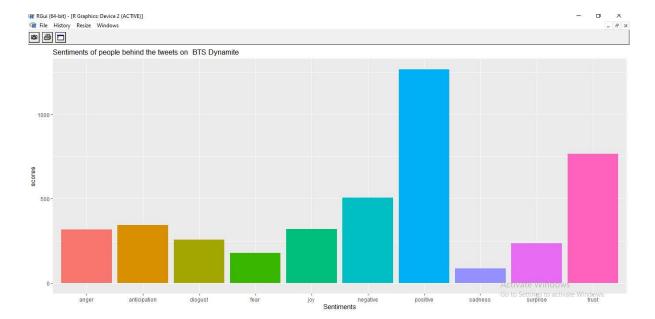
Sentiment on IPL2020 Trend



Sentiment on Covid-19 Trend



Sentiment on BTS Dynamite Trend



Sentiment on SRK Trend

