1. **Write an algorithm which demonstrates your knowledge about extracting the tweets using TwitterR package**

**PREREQUISITES:**

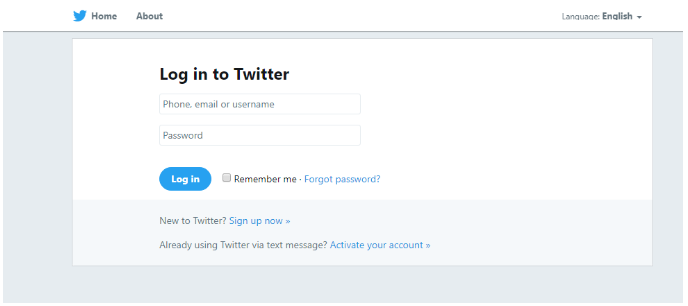
1. Installed [R](http://www.r-project.org/)& [RStudio](https://www.rstudio.com/ide/download/" \t "_blank)

2. A Twitter application and [account](https://twitter.com/) to extract tweets  
3. A [Twitter Developers](https://dev.twitter.com/apps/) login ID & Password

# STEPS:-

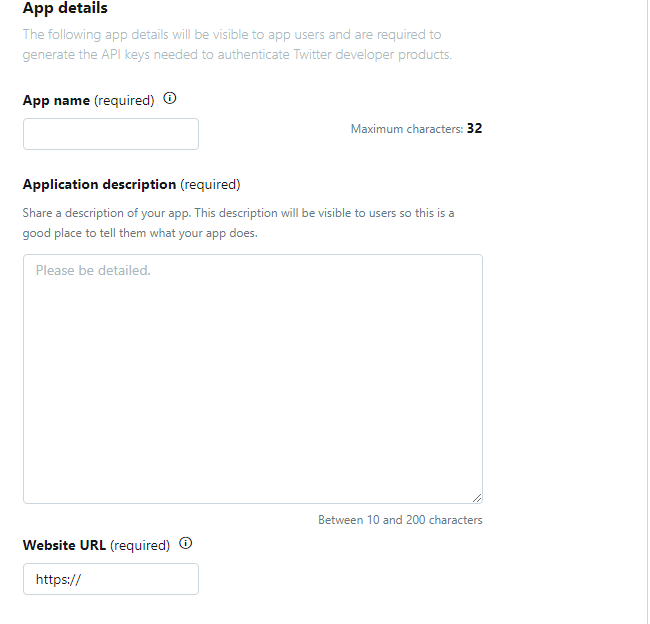
# 1. Create a Twitter Application

Sign in using your Twitter Account

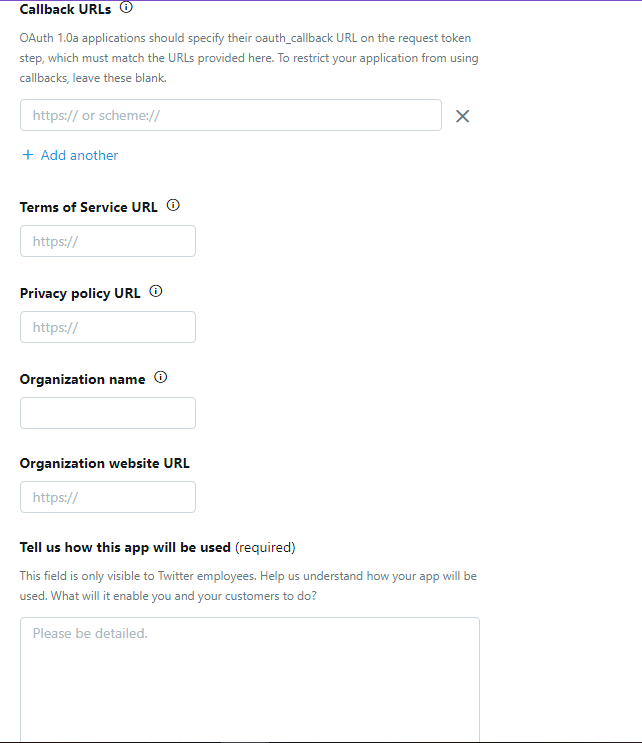


Then we will create our own twitter developer account if not crated already. After creating the twitter developer account, we will go to the dashboard and hit create an app option.

Create an application



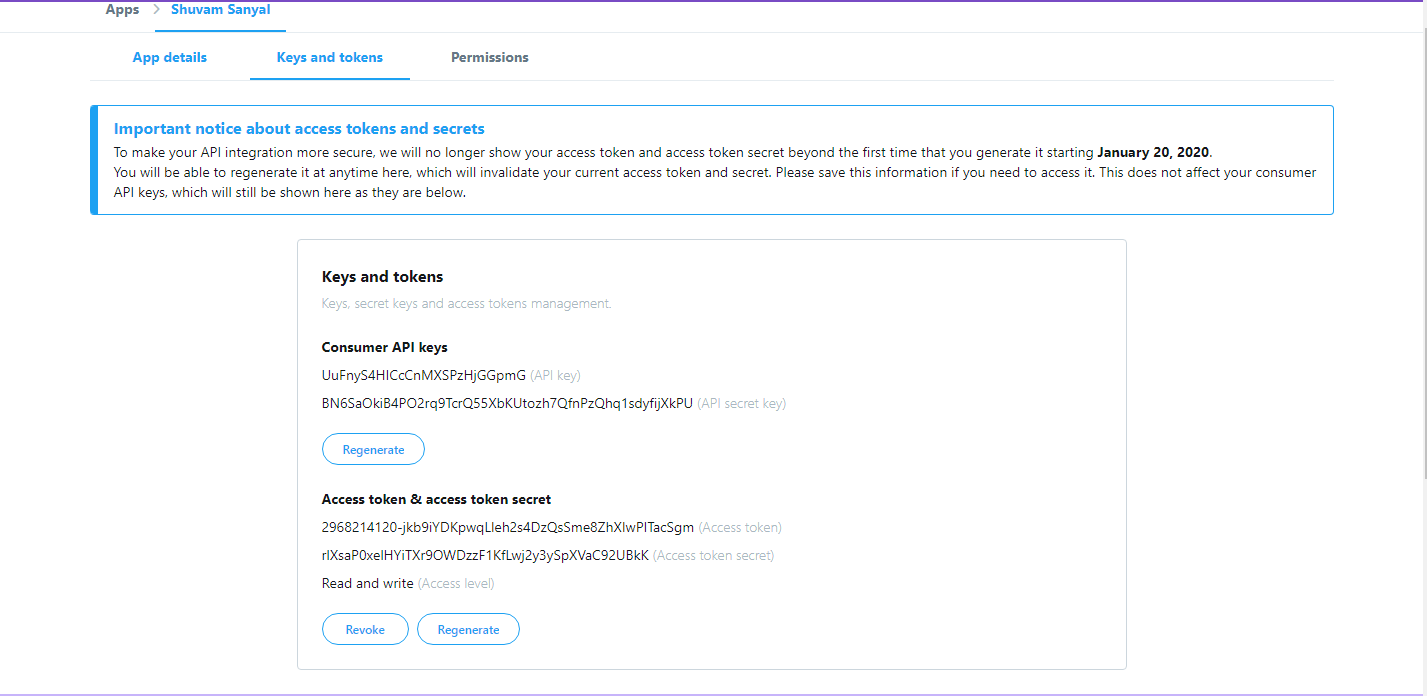
Give an application name as per your choice which you will use later too. Also give a suitable URL which you will use for long term use.



Now here ignore all the fields and write in 100 characters about how this app will be used.

**NOTE:** If you haven’t already, then you may need to add your phone number to your twitter profile. Go to **settings and privacy→ Mobile**

Go to **Keys and Access Tokens**



Write down the **Consumer key (API Key)** and the **Consumer secret (API Secret).** Once the access token is created, write down the **Access Token** and the **Access Token Secret.**

# 2. Install and Load R packages

Open RStudio or whatever IDE you are using for R, and install and load the following two packages.

*twitteR:*Provides an interface to the Twitter web API.

#Install and Load R packages

install.packages("twitteR")  
library("twitteR")

# 3. Setup Twitter Authentication

Use your **consumer key**, **consumer secret**, **access token**, and **access secret**that you wrote down in the earlier steps from the above, and replace ‘xxxxxxxxxxx’ with the appropriate values.

library(twitteR)

library(SnowballC)

library(tm)

library(syuzhet)

library(RJSONIO)

library(RCurl)

library(stringr)

library(wordcloud)

library(httr)

library(rtweet)

library(tidaytext)

library(RColorBrewer)

library(ggthemes)

library(map)

library(gganimate)

library(lubricate)

library(leaflet)

library(rtweet)

library(ggplot2)

#Authentication Key

consumer\_key <- "UuFnyS4HICcCnMXSPzHjGGpmG"

consumer\_secret <-"BN6SaOkiB4PO2rq9TcrQ55XbKUtozh7QfnPzQhq1sdyfijXkPU"

access\_token <- "2968214120-jkb9iYDKpwqLleh2s4DzQsSme8ZhXIwPITacSgm"

access\_secret <- "rlXsaP0xelHYiTXr9OWDzzF1KfLwj2y3ySpXVaC92UBkK"

# 4. Extract Tweets

The SearchTwitter function code below will return tweets about the search item #Environment in english.

# Create Twitter Connection

setup\_twitter\_oauth(consumer\_key, consumer\_secret, access\_token, access\_secret

#Extract Tweets

tweets <- searchTwitter("Environment", n=597, lang="en", Date="2019-10-21")

tweets

tweets.text <- sapply(tweets, function(x) x$getText())

strip\_retweets(tweets)

users\_data(tweets)

lat\_lng(tweets)

tweets.df <- twListToDF(tweets)

head(tweets.df)

head(tweets.df$text)

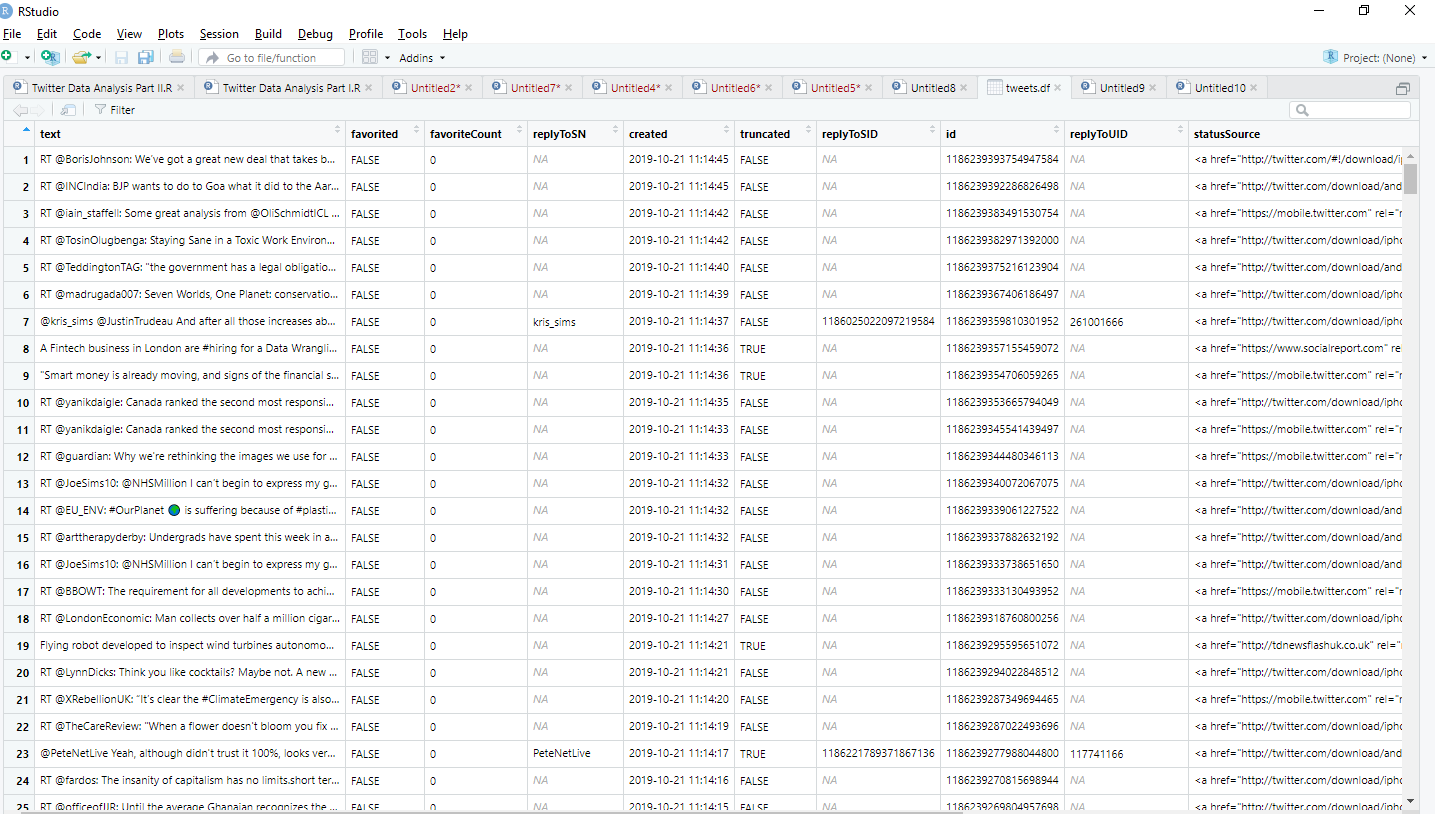
tweets.df <- unique(tweets.df)

View(tweets.df)

head(tweets.df)

Now you get the list of 597 tweets with search term Environment.

Below is a part of whole screenshot of the tweets we fetched.



**b) Convert this algorithm to R Code.**

**c) Who are the people who have used the tweets and what are the times lines of such tweets?**

install.packages("twitteR")

install.packages("SnowballC")

install.packages("tm")

install.packages("syuzhet")

install.packages("RJSONIO")

install.packages("RCurl")

install.packages("stringr")

install.packages("wordcloud")

install.packages("httr")

install.packages("rtweet")

install.packages("tidytext")

install.packages("RColorBrewer")

install.packages("leaflet")

install.packages("gganimate")

install.packages("lubridate")

install.packages("maps")

install.packages("ggthemes")

install.packages("ggplot2")

library(twitteR)

library(SnowballC)

library(tm)

library(syuzhet)

library(RJSONIO)

library(RCurl)

library(stringr)

library(wordcloud)

library(httr)

library(rtweet)

library(tidaytext)

library(RColorBrewer)

library(ggthemes)

library(map)

library(gganimate)

library(lubricate)

library(leaflet)

library(rtweet)

library(ggplot2)

consumer\_key <- "UuFnyS4HICcCnMXSPzHjGGpmG"

consumer\_secret <-"BN6SaOkiB4PO2rq9TcrQ55XbKUtozh7QfnPzQhq1sdyfijXkPU"

access\_token <- "2968214120-jkb9iYDKpwqLleh2s4DzQsSme8ZhXIwPITacSgm"

access\_secret <- "rlXsaP0xelHYiTXr9OWDzzF1KfLwj2y3ySpXVaC92UBkK"

# Create Twitter Connection

setup\_twitter\_oauth(consumer\_key, consumer\_secret, access\_token, access\_secret)

tweets <- searchTwitter("Environment", n=597, lang="en", since="2019-04-20")

tweets

strip\_retweets(tweets)

#Converting the tweets into dataframe

tweets.df <- twListToDF(tweets)

head(tweets.df)

head(tweets.df$text)

tweets.df <- unique(tweets.df)

View(tweets.df)

head(tweets.df)

#convert all text to lower case

tweets.text <- tolower(tweets.text)

# Replace blank space (“rt”)

tweets.text <- gsub("rt", "", tweets.text)

# Replace @UserName

tweets.text <- gsub("@\\w+", "", tweets.text)

# Remove punctuation

tweets.text <- gsub("[[:punct:]]", "", tweets.text)

# Remove links

tweets.text <- gsub("http\\w+", "", tweets.text)

# Remove tabs

tweets.text <- gsub("[ |\t]{2,}", "", tweets.text)

# Remove blank spaces at the beginning

tweets.text <- gsub("^ ", "", tweets.text)

# Remove blank spaces at the end

tweets.text <- gsub(" $", "", tweets.text)

#create corpus

tweets.text.corpus <- Corpus(VectorSource(tweets.text))

#clean up by removing stop words

tweets.text.corpus <- tm\_map(tweets.text.corpus, function(x)removeWords(x,stopwords()))

#Creating Wordcloud

wordcloud(tweets.text.corpus,min.freq = 2, scale=c(7,0.5),colors=brewer.pal(8, "Dark2"), random.color= TRUE, random.order = FALSE, max.words = 150)

#Sentiment Analysis

word.df <- as.vector(tweets.df2)

emotion.df <- get\_nrc\_sentiment(word.df)

emotion.df2 <- cbind(tweets.df2, emotion.df)

head(emotion.df2)

sent.value <- get\_sentiment(word.df)

most.positive <- word.df[sent.value == max(sent.value)]

most.positive

most.negative <- word.df[sent.value <= min(sent.value)]

most.negative

sent.value

positive.tweets <- word.df[sent.value > 0]

head(positive.tweets)

negative.tweets <- word.df[sent.value < 0]

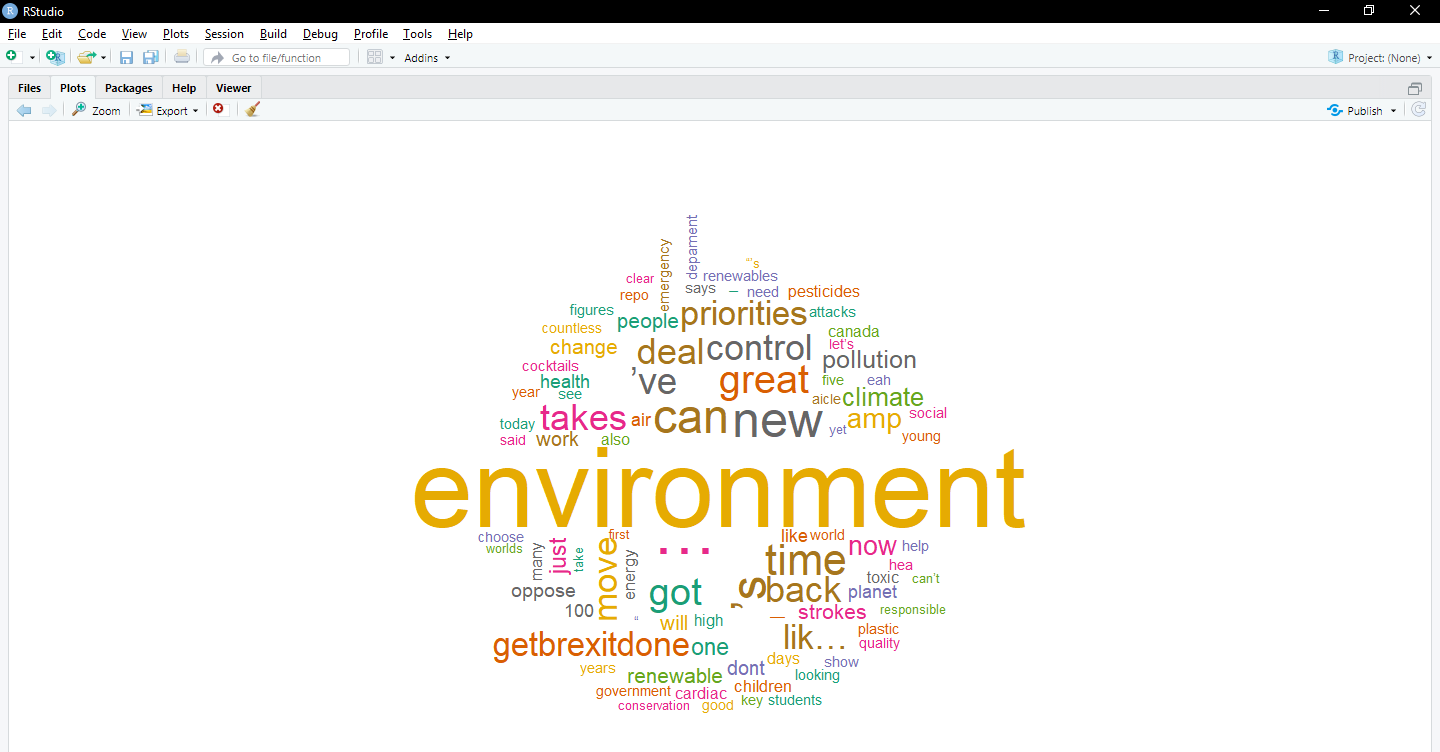
head(negative.tweets)

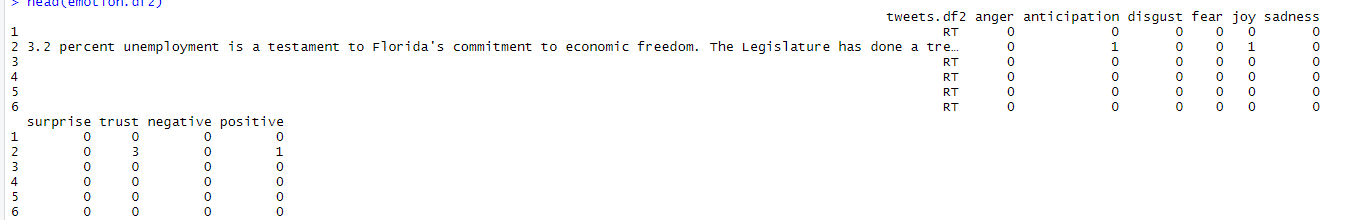
neutral.tweets <- word.df[sent.value == 0]

head(neutral.tweets)

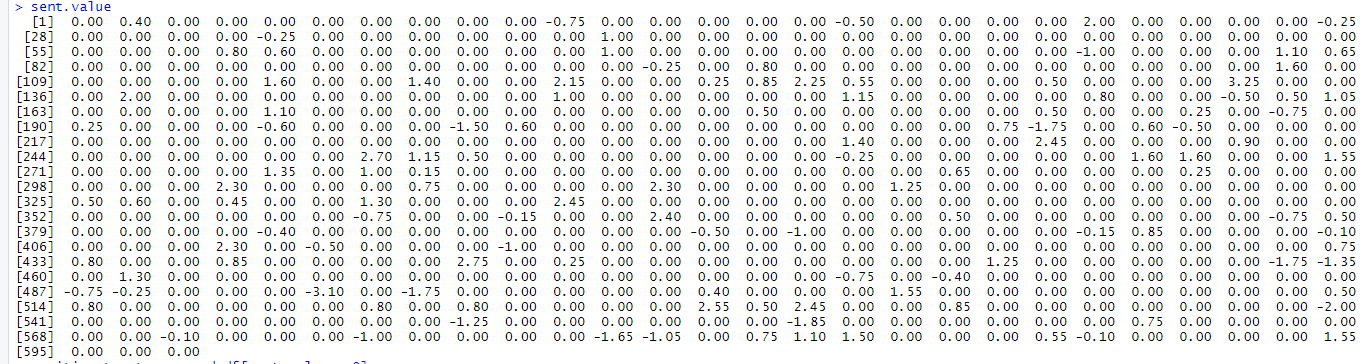
Now from this twitter text mining in R, We got to know who are the users who have tweeted these tweets and for that I have prepared a separate excel file which I will be sharing along with this word document.(Screen Name can be found from the twitter ID of the twitter from where these tweets come from).

Also the timeline, it come from the column created in the excel file. In this case it is 21st of October 2019. We have analyzed the tweets coming on this particular day with search term Environment.

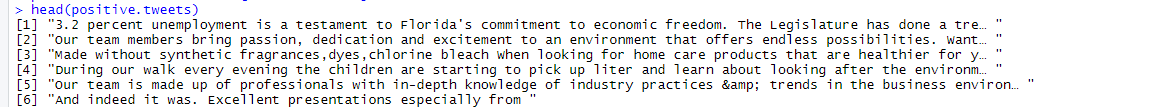




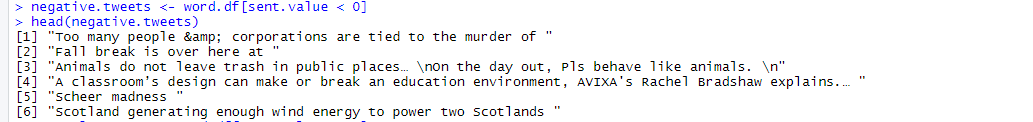
Sentiment Values at a glance:-



Positive Tweets:-



Negative Tweets:-



Neutral Tweets:-



**THE END**

**THANK YOU**

**PREPARED BY-SHUVAM SANYAL**

**MSC DATA SCIENCE**

**PRN-19070243015**