rm(list=ls())

install.packages("twitteR", dependencies = T)

install.packages(c('ROAuth','RCurl'))

install.packages("httr")

library(httr)

require('ROAuth')

require('RCurl')

library(twitteR)

library(plyr)

reqURL <- "https://api.twitter.com/oauth/request\_token"

accessURL <- "https://api.twitter.com/oauth/access\_token"

authURL <- "https://api.twitter.com/oauth/authorize"

consumerKey <- "Jy60bQAxQqx1levBVBPW5oZUn"

consumerSecret <- "M3nJWGEGphBnFmB8QvxdcnTYAQ1HDDOyYenMTtZrBZciDy7m5V"

#Class OAuth wraps and handles OAuth handshakes and signatures for the user within R

twitCred <- OAuthFactory$new(consumerKey=consumerKey,consumerSecret=consumerSecret,requestURL=reqURL,accessURL=accessURL,authURL=authURL)

## Bundle of CA Root Certificates

##

## Certificate data from Mozilla as of: Wed Nov 2 04:12:05 2016 GMT

##

## This is a bundle of X.509 certificates of public Certificate Authorities

## (CA). These were automatically extracted from Mozilla's root certificates

## file (certdata.txt).

download.file(url="https://curl.haxx.se/ca/cacert.pem", destfile="cacert.pem")

###destfile – Where the file should be saved (path with a file name)

#CA = Certificate Authorities

twitCred$handshake(cainfo="cacert.pem") #This command would need local SSL certificate

#access Key and #access token

setup\_twitter\_oauth(consumerKey, consumerSecret, "173355826-9AyV0zCKkx3butTUkbJMipzCupubyabouxUt3K4B", "cb7Ys3jaOjZiwvHmA2Hf2dY8wh79xxcfhdpad2s849Jqn")

#Now we have our tweets

tweets <- searchTwitter("Kejriwal",n=1500, lang = "en")

#The Analysis:

df <- twListToDF(tweets)

#To be able to analyze our tweets, we have to extract their text and

#save it into the variable tweets.text by typing:

Tweets.text = laply(tweets,function(t)t$getText())

#We now have to load the words in variables to use them by typing:

#ScanFunction: Read data into a vector or list from the console or file.

#pos = scan('positive-words.txt', what='character', comment.char=';')

#neg = scan('negative-words.txt', what='character', comment.char=';')

#----------- CLEAN UP TEXT-------------#

#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)

#--------------- REMOVE STOP WORDS-------------------#

#In the next step we will use the text mining package tm to remove stop words.

#A stop word is a commonly used word such as “the”. Stop words should not be

#included in the analysis. If tm is not already installed you will need to

#install it (available from the Comprehensive R Archive Network).

#install tm – if not already installed

install.packages("tm", dependencies = T) #Text mining package to remove stop words

library("tm")

sessionInfo()

#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()))

#-------------------------- GENERATE WORD CLOUD-------------------------------#

#Now we’ll generate the word cloud using the wordcloud package. For this example we

#are concerned with plotting no more than 150 words that occur more than once with

#random color, order, and position

#install wordcloud if not already installed

install.packages("wordcloud")

library(wordcloud)

#generate wordcloud

wordcloud(Tweets.text.corpus, min.freq = 3, scale=c(7,0.5),colors=brewer.pal(8, "Dark2"), random.color= TRUE, random.order = FALSE, max.words = 150)