Sharat\_Sripada\_HW3

requestURL <- 'https://api.twitter.com/oauth/request\_token'  
accessURL <- 'https://api.twitter.com/oauth/access\_token'  
authURL <- 'https://api.twitter.com/oauth/authorize'  
  
# Install the following packages  
# install.packages('twitteR')  
# install.packages('ROAuth')  
# install.packages('rtweet')   
# install.packages('streamR')  
# install.packages('rjson')  
# install.packages('tokenizers')  
# install.packages('tidyverse')  
# install.packages('syuzhet')  
# install.packages('data.table')  
# install.packages('arulesViz')  
# install.packages('stopwords')  
# install.packages('wordcloud')  
# install.packages('tm')  
# install.packages('arules')  
  
library(arules)

## Loading required package: Matrix

##   
## Attaching package: 'arules'

## The following objects are masked from 'package:base':  
##   
## abbreviate, write

library(rtweet)  
library(twitteR)

##   
## Attaching package: 'twitteR'

## The following object is masked from 'package:rtweet':  
##   
## lookup\_statuses

library(ROAuth)  
library(jsonlite)

##   
## Attaching package: 'jsonlite'

## The following object is masked from 'package:rtweet':  
##   
## flatten

library(rjson)

##   
## Attaching package: 'rjson'

## The following objects are masked from 'package:jsonlite':  
##   
## fromJSON, toJSON

library(tokenizers)  
library(tidyverse)

## ── Attaching packages ──────────────────────────────────────────────────────────────────────────────────────────────────────── tidyverse 1.3.0 ──

## ✓ ggplot2 3.3.2 ✓ purrr 0.3.4  
## ✓ tibble 3.0.3 ✓ dplyr 1.0.0  
## ✓ tidyr 1.1.0 ✓ stringr 1.4.0  
## ✓ readr 1.3.1 ✓ forcats 0.5.0

## ── Conflicts ─────────────────────────────────────────────────────────────────────────────────────────────────────────── tidyverse\_conflicts() ──  
## x tidyr::expand() masks Matrix::expand()  
## x dplyr::filter() masks stats::filter()  
## x purrr::flatten() masks jsonlite::flatten(), rtweet::flatten()  
## x rjson::fromJSON() masks jsonlite::fromJSON()  
## x dplyr::id() masks twitteR::id()  
## x dplyr::lag() masks stats::lag()  
## x dplyr::location() masks twitteR::location()  
## x tidyr::pack() masks Matrix::pack()  
## x dplyr::recode() masks arules::recode()  
## x rjson::toJSON() masks jsonlite::toJSON()  
## x tidyr::unpack() masks Matrix::unpack()

library(plyr)

## ------------------------------------------------------------------------------

## You have loaded plyr after dplyr - this is likely to cause problems.  
## If you need functions from both plyr and dplyr, please load plyr first, then dplyr:  
## library(plyr); library(dplyr)

## ------------------------------------------------------------------------------

##   
## Attaching package: 'plyr'

## The following objects are masked from 'package:dplyr':  
##   
## arrange, count, desc, failwith, id, mutate, rename, summarise,  
## summarize

## The following object is masked from 'package:purrr':  
##   
## compact

## The following object is masked from 'package:twitteR':  
##   
## id

library(dplyr)  
library(ggplot2)  
library(syuzhet) #sentiment analysis

##   
## Attaching package: 'syuzhet'

## The following object is masked from 'package:rtweet':  
##   
## get\_tokens

library(stringr)  
#library(arulesViz)  
library(stopwords)  
library(tm)

## Loading required package: NLP

##   
## Attaching package: 'NLP'

## The following object is masked from 'package:ggplot2':  
##   
## annotate

##   
## Attaching package: 'tm'

## The following object is masked from 'package:stopwords':  
##   
## stopwords

## The following object is masked from 'package:arules':  
##   
## inspect

library(RColorBrewer)  
library(wordcloud)

consumerKey <- 'rtOHXdgu2S8SIFctVfF0yhXcY'  
consumerSecret <- 'l8E5AIArXxZvr2idFTCzoLDkjqssLVIeo3TaUSyObqH0tQ7KsE'  
access\_Token <- '185329008-rIYt3Y8HBkgBVSdYcy6iTMkXiUXFF3cSJkjuCZU6'  
access\_Secret <- 'qKIz3VOjOus4mvNNg0JYGbaMHncPgLqajhUxJfiUBTHbL'  
  
# Using twitteR  
setup\_twitter\_oauth(consumerKey, consumerSecret, access\_Token, access\_Secret)

## [1] "Using direct authentication"

Search <- twitteR::searchTwitter("#Trump",n=3000,since="2020-03-01")  
Search\_DF <- twListToDF(Search)  
TransactionTweetsFile = "Trump\_2020.csv"  
Search\_DF$text[1]

## [1] "She and David would be going through a separation for the billionth time and CPS would have full CUstody of everyon… https://t.co/JdToxQ85cS"

## Sentiment analysis

First analysis of the raw tweets associated with #Trump since Mar-2020

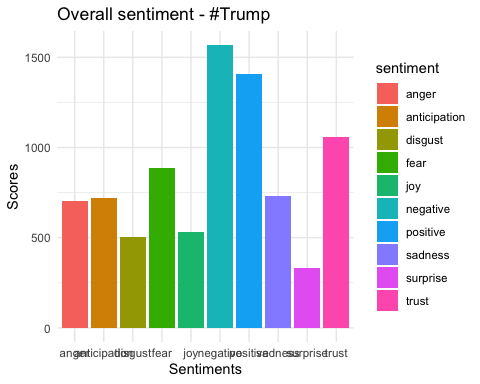
# Converting tweets to ASCII   
tweets <- iconv(Search\_DF$text, from="UTF-8", to="ASCII", sub="")  
  
my\_stop\_words <- c()  
# Clean-up a few-words prior to sentiment analysis  
for(i in stopwords()){  
 pattern <- paste(" ", i, " ")  
 my\_stop\_words <- c(my\_stop\_words, pattern)  
}  
  
my\_stop\_words <- c(my\_stop\_words, '#trump', '#', 'RT.\*: ', 'amp',   
 'trump', 'https')  
  
for(i in my\_stop\_words){  
 tweets <- gsub(i, '', tweets, ignore.case=T)  
}  
  
# Get the sentiment scores  
sentiment <- get\_nrc\_sentiment((tweets))

## Warning: `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.

## Warning: `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.

## Warning: `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\_scores <- data.frame(colSums(sentiment[,]))  
names(sentiment\_scores) <- "Score"  
sentiment\_scores <- cbind("sentiment"=rownames(sentiment\_scores), sentiment\_scores)  
rownames(sentiment\_scores) <- NULL  
  
ggplot(data=sentiment\_scores,aes(x=sentiment, y=Score)) +  
 geom\_bar(aes(fill=sentiment),stat = "identity") +  
 theme(legend.position="none") +  
 xlab("Sentiments")+ylab("Scores") +  
 ggtitle("Overall sentiment - #Trump") +  
 theme\_minimal()



sentiment\_scores <- data.frame(colSums(sentiment[,]))

The overall sentiment seems to be that ‘Negativity’ - Anger, disgust, fear, negative, sadness & possibly (mis-) trust. This was probably a start of a lot of events that caused the overall situation to degrade with time - to cite a few: - socio-economic issues around unemployment rate - BLM protests - Trade wars - DACA and immigration issues - AND perhaps the greatest of them all, the Pandemic itself #COVID-19.

## Wordcloud

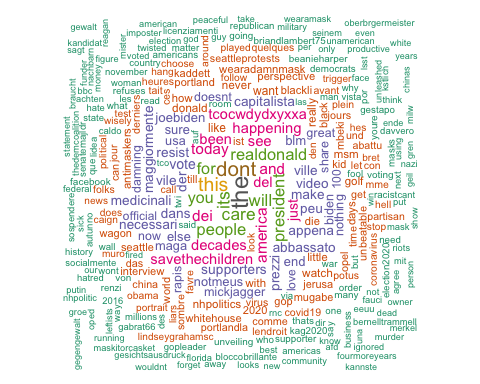
Next, let’s mine the words and corroborate it with the sentiments above.

# library(tm)  
set.seed(1234)  
wordcloud(tweets[1:1500], min.freq=10, scale=c(1.5, .5), random.order=FALSE, rot.per=0.25,   
 colors=brewer.pal(8, "Dark2"))

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

## Warning in tm\_map.SimpleCorpus(corpus, function(x) tm::removeWords(x,  
## tm::stopwords())): transformation drops documents

## Warning in wordcloud(tweets[1:1500], min.freq = 10, scale = c(1.5, 0.5), :  
## vorbild could not be fit on page. It will not be plotted.



NOTE: One of the topics that seems popular is tcocwdydxyxxa (which is essentially a page hosted on twitter). ARM in the coming sections will likely help us understand the association/topic.

## Start the file

# Unload tm since there is a conflict between the inspect method in tm and arules  
# detach(package:tm, unload=TRUE)  
  
Trans <- file(TransactionTweetsFile)

## Tokenize to words

Tokens <- tokenizers::tokenize\_words(Search\_DF$text[1], stopwords = stopwords::stopwords('en'),   
 lowercase = TRUE, strip\_punct = TRUE, strip\_numeric = TRUE,  
 simplify = TRUE)

## Write squished Tokens

cat(unlist(str\_squish(Tokens)), "\n", file=Trans, sep=",")  
close(Trans)

## Append the remaining lists of Tokens into the csv file

Trans <- file(TransactionTweetsFile, open = "a")  
for(i in 2:nrow(Search\_DF)){  
 Tokens <- tokenizers::tokenize\_words(Search\_DF$text[i], stopwords = stopwords::stopwords('en'),   
 lowercase = TRUE, strip\_punct = TRUE, strip\_numeric = TRUE,  
 simplify = TRUE)  
 cat(unlist(str\_squish(Tokens)), "\n", file=Trans, sep=",")  
   
}  
close(Trans)

## Read and inspect transactions

TweetTrans <- read.transactions(TransactionTweetsFile, rm.duplicates = FALSE,  
 format = "basket",   
 sep=",")

## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string  
  
## Warning in scan(text = l, what = "character", sep = sep, quote = quote, : EOF  
## within quoted string

## Warning in asMethod(object): removing duplicated items in transactions

arules::inspect(head(TweetTrans))

## items   
## [1] {billionth,   
## cps,   
## custody,   
## david,   
## everyon,   
## full,   
## going,   
## https,   
## jdtoxq85cs,   
## separation,   
## t.co,   
## time}   
## [2] {blm,   
## deaf,   
## illegal,   
## let,   
## mlb,   
## nyc,   
## occupation,   
## pitch,   
## sundayvibes,   
## threatened,   
## throw,   
## tone,   
## trump,   
## want,   
## yankees}   
## [3] {better,   
## bob,   
## cha,   
## gop,   
## https,   
## idiot,   
## lawmaker,   
## make,   
## president,   
## says,   
## sponge,   
## t.co,   
## told,   
## trump,   
## woulda,   
## zlhxkcv9xv}   
## [4] {artist,   
## blackisking,   
## blacklivesmatter,   
## hidden,   
## https,   
## poverty,   
## pxu1glz6xw,   
## rendition,   
## resistancetaskforce,  
## street,   
## t.co}   
## [5] {alle,   
## dass,   
## der,   
## dieser,   
## es,   
## geschichte,   
## gewinnt,   
## murrraydo,   
## nie,   
## noch,   
## präsident,   
## rt,   
## trump,   
## um,   
## usa,   
## verlieren,   
## verliert,   
## war,   
## wichtig,   
## wir}   
## [6] {5m6fppdlu1,   
## coronavirus,   
## desantis,   
## eru5bwqpv9,   
## florida,   
## followed,   
## https,   
## ravaged,   
## ron,   
## scientists,   
## sidelined,   
## t.co,   
## trump}

Sample\_Trans <- sample(TweetTrans, 50)  
summary(Sample\_Trans)

## transactions as itemMatrix in sparse format with  
## 50 rows (elements/itemsets/transactions) and  
## 8370 columns (items) and a density of 0.001538829   
##   
## most frequent items:  
## rt trump https t.co realdonaldtrump   
## 36 30 19 18 5   
## (Other)   
## 536   
##   
## element (itemset/transaction) length distribution:  
## sizes  
## 4 7 8 9 10 11 12 13 14 15 17 18 20 21 23   
## 1 2 2 3 2 7 8 8 7 2 1 3 1 1 2   
##   
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 4.00 11.00 12.50 12.88 14.00 23.00   
##   
## includes extended item information - examples:  
## labels  
## 1 \_\_huo\_\_  
## 2 \_\_verlaine\_\_  
## 3 \_daniellew\_\_\_

## Clean-up

Remove some frequently appearing words like t.co, rt, https (and it’s nice to see that no one has been posting ‘<http://>’ links)

TweetDF <- read.csv(TransactionTweetsFile, header = FALSE, sep=",")  
head(TweetDF)

## V1 V2 V3 V4 V5 V6  
## 1 david going separation billionth time cps  
## 2 yankees trump threatened nyc illegal occupation  
## 3 sponge bob woulda make better president  
## 4 street artist rendition hidden poverty resistancetaskforce  
## 5 rt murrraydo verliert trump verlieren wir  
## 6 coronavirus ravaged florida ron desantis sidelined  
## V7 V8 V9 V10 V11 V12 V13  
## 1 full custody everyon https t.co jdtoxq85cs   
## 2 want let throw pitch tone deaf mlb  
## 3 idiot trump says gop lawmaker told cha  
## 4 blackisking blacklivesmatter https t.co pxu1glz6xw   
## 5 alle noch nie der geschichte der usa  
## 6 scientists followed trump https t.co 5m6fppdlu1 https  
## V14 V15 V16 V17 V18 V19 V20 V21 V22  
## 1   
## 2 sundayvibes blm   
## 3 https t.co zlhxkcv9xv   
## 4   
## 5 war es wichtig dass dieser präsident gewinnt um   
## 6 t.co eru5bwqpv9

# Convert all the columns to char  
TweetDF <- TweetDF %>%  
 mutate\_all(as.character)  
(str(TweetDF))

## 'data.frame': 3142 obs. of 22 variables:  
## $ V1 : chr "david" "yankees" "sponge" "street" ...  
## $ V2 : chr "going" "trump" "bob" "artist" ...  
## $ V3 : chr "separation" "threatened" "woulda" "rendition" ...  
## $ V4 : chr "billionth" "nyc" "make" "hidden" ...  
## $ V5 : chr "time" "illegal" "better" "poverty" ...  
## $ V6 : chr "cps" "occupation" "president" "resistancetaskforce" ...  
## $ V7 : chr "full" "want" "idiot" "blackisking" ...  
## $ V8 : chr "custody" "let" "trump" "blacklivesmatter" ...  
## $ V9 : chr "everyon" "throw" "says" "https" ...  
## $ V10: chr "https" "pitch" "gop" "t.co" ...  
## $ V11: chr "t.co" "tone" "lawmaker" "pxu1glz6xw" ...  
## $ V12: chr "jdtoxq85cs" "deaf" "told" "" ...  
## $ V13: chr "" "mlb" "cha" "" ...  
## $ V14: chr "" "sundayvibes" "https" "" ...  
## $ V15: chr "" "blm" "t.co" "" ...  
## $ V16: chr "" "" "zlhxkcv9xv" "" ...  
## $ V17: chr "" "" "" "" ...  
## $ V18: chr "" "" "" "" ...  
## $ V19: chr "" "" "" "" ...  
## $ V20: chr "" "" "" "" ...  
## $ V21: chr "" "" "" "" ...  
## $ V22: chr "" "" "" "" ...

## NULL

TweetDF[TweetDF == 't.co'] <- ''  
TweetDF[TweetDF == 'rt'] <- ''  
TweetDF[TweetDF == 'http'] <- ''  
TweetDF[TweetDF == 'https'] <- ''  
TweetDF[TweetDF == 'amp'] <- ''  
  
# Clean-up with grepl  
MyDF <- NULL  
for (i in 1:ncol(TweetDF)) {  
 MyList <- c()  
 MyList <- c(MyList, grepl("[[:digit:]]", TweetDF[[i]]))  
 MyDF <- cbind(MyDF, MyList)  
}  
  
TweetDF[MyDF] <- ""  
# (TweetDF)

## Save the data-frame using the write table command

write.table(TweetDF, file = 'UpdatedTrump\_2020.csv', col.names = FALSE,   
 row.names = FALSE, sep = ',')  
TweetTrans <- read.transactions('UpdatedTrump\_2020.csv', sep=',', format('basket'), rm.duplicates = TRUE)

## distribution of transactions with duplicates:  
## items  
## 1 2 3 4 5 6   
## 663 260 35 13 35 1

arules::inspect(head(TweetTrans))

## items   
## [1] {billionth,   
## cps,   
## custody,   
## david,   
## everyon,   
## full,   
## going,   
## separation,   
## time}   
## [2] {blm,   
## deaf,   
## illegal,   
## let,   
## mlb,   
## nyc,   
## occupation,   
## pitch,   
## sundayvibes,   
## threatened,   
## throw,   
## tone,   
## trump,   
## want,   
## yankees}   
## [3] {better,   
## bob,   
## cha,   
## gop,   
## idiot,   
## lawmaker,   
## make,   
## president,   
## says,   
## sponge,   
## told,   
## trump,   
## woulda}   
## [4] {artist,   
## blackisking,   
## blacklivesmatter,   
## hidden,   
## poverty,   
## rendition,   
## resistancetaskforce,  
## street}   
## [5] {alle,   
## dass,   
## der,   
## dieser,   
## es,   
## geschichte,   
## gewinnt,   
## murrraydo,   
## nie,   
## noch,   
## präsident,   
## trump,   
## um,   
## usa,   
## verlieren,   
## verliert,   
## war,   
## wichtig,   
## wir}   
## [6] {coronavirus,   
## desantis,   
## florida,   
## followed,   
## ravaged,   
## ron,   
## scientists,   
## sidelined,   
## trump}

## Association Rule Mining

Exploring apriori methods of translating transactions in a hierarchical tree-like data-structure and pruning out less popular/frequent paths. For this we will use the following support and confidence thresholds: - support-threshold = 0.025 - confidence-threshold = 0.5

TweetTrans\_rules <- arules::apriori(TweetTrans,   
 parameter = list(support=0.025, confidence=0.5, minlen=3))

## Apriori  
##   
## Parameter specification:  
## confidence minval smax arem aval originalSupport maxtime support minlen  
## 0.5 0.1 1 none FALSE TRUE 5 0.025 3  
## maxlen target ext  
## 10 rules TRUE  
##   
## Algorithmic control:  
## filter tree heap memopt load sort verbose  
## 0.1 TRUE TRUE FALSE TRUE 2 TRUE  
##   
## Absolute minimum support count: 78   
##   
## set item appearances ...[0 item(s)] done [0.00s].  
## set transactions ...[7592 item(s), 3142 transaction(s)] done [0.01s].  
## sorting and recoding items ... [37 item(s)] done [0.00s].  
## creating transaction tree ... done [0.00s].  
## checking subsets of size 1 2 3 4 5 6 7 8 done [0.00s].  
## writing ... [980 rule(s)] done [0.00s].  
## creating S4 object ... done [0.00s].

arules::inspect(head(TweetTrans\_rules))

## lhs rhs support confidence  
## [1] {cwdydxyxxa,savethechildren} => {\_whiterabbitt\_} 0.03723743 1   
## [2] {\_whiterabbitt\_,savethechildren} => {cwdydxyxxa} 0.03723743 1   
## [3] {\_whiterabbitt\_,cwdydxyxxa} => {savethechildren} 0.03723743 1   
## [4] {cwdydxyxxa,savethechildren} => {dont} 0.03723743 1   
## [5] {dont,savethechildren} => {cwdydxyxxa} 0.03723743 1   
## [6] {cwdydxyxxa,dont} => {savethechildren} 0.03723743 1   
## coverage lift count  
## [1] 0.03723743 26.85470 117   
## [2] 0.03723743 26.85470 117   
## [3] 0.03723743 26.85470 117   
## [4] 0.03723743 26.62712 117   
## [5] 0.03723743 26.85470 117   
## [6] 0.03723743 26.85470 117

cwdydxyxxa as seen in the wordcloud was one of the popular items and this seems to be associate with saving the children or a related movement at the time.

## Sorted

# By confidence  
SortedRules\_conf <- sort(TweetTrans\_rules, by='confidence', decreasing=TRUE)  
arules::inspect(head(SortedRules\_conf))

## lhs rhs support confidence  
## [1] {cwdydxyxxa,savethechildren} => {\_whiterabbitt\_} 0.03723743 1   
## [2] {\_whiterabbitt\_,savethechildren} => {cwdydxyxxa} 0.03723743 1   
## [3] {\_whiterabbitt\_,cwdydxyxxa} => {savethechildren} 0.03723743 1   
## [4] {cwdydxyxxa,savethechildren} => {dont} 0.03723743 1   
## [5] {dont,savethechildren} => {cwdydxyxxa} 0.03723743 1   
## [6] {cwdydxyxxa,dont} => {savethechildren} 0.03723743 1   
## coverage lift count  
## [1] 0.03723743 26.85470 117   
## [2] 0.03723743 26.85470 117   
## [3] 0.03723743 26.85470 117   
## [4] 0.03723743 26.62712 117   
## [5] 0.03723743 26.85470 117   
## [6] 0.03723743 26.85470 117

# By support  
SortedRules\_sup <- sort(TweetTrans\_rules, by='support', decreasing=TRUE)  
arules::inspect(head(SortedRules\_sup))

## lhs rhs support confidence  
## [1] {cwdydxyxxa,savethechildren} => {\_whiterabbitt\_} 0.03723743 1   
## [2] {\_whiterabbitt\_,savethechildren} => {cwdydxyxxa} 0.03723743 1   
## [3] {\_whiterabbitt\_,cwdydxyxxa} => {savethechildren} 0.03723743 1   
## [4] {cwdydxyxxa,savethechildren} => {dont} 0.03723743 1   
## [5] {dont,savethechildren} => {cwdydxyxxa} 0.03723743 1   
## [6] {cwdydxyxxa,dont} => {savethechildren} 0.03723743 1   
## coverage lift count  
## [1] 0.03723743 26.85470 117   
## [2] 0.03723743 26.85470 117   
## [3] 0.03723743 26.85470 117   
## [4] 0.03723743 26.62712 117   
## [5] 0.03723743 26.85470 117   
## [6] 0.03723743 26.85470 117

# By Lift  
SortedRules\_lift <- sort(TweetTrans\_rules, by='lift', decreasing=TRUE)  
arules::inspect(head(SortedRules\_lift))

## lhs rhs support confidence  
## [1] {richardangwin,trump} => {resist} 0.02673456 1.0000000   
## [2] {resist,trump} => {richardangwin} 0.02673456 0.8842105   
## [3] {cwdydxyxxa,savethechildren} => {\_whiterabbitt\_} 0.03723743 1.0000000   
## [4] {\_whiterabbitt\_,savethechildren} => {cwdydxyxxa} 0.03723743 1.0000000   
## [5] {\_whiterabbitt\_,cwdydxyxxa} => {savethechildren} 0.03723743 1.0000000   
## [6] {dont,savethechildren} => {cwdydxyxxa} 0.03723743 1.0000000   
## coverage lift count  
## [1] 0.02673456 33.07368 84   
## [2] 0.03023552 31.21561 84   
## [3] 0.03723743 26.85470 117   
## [4] 0.03723743 26.85470 117   
## [5] 0.03723743 26.85470 117   
## [6] 0.03723743 26.85470 117

The two things that I chose to mine in light of the current pandemic (with respect to #Trump): - Anthony S Fauci, Director of NIAID and sentiment of Twitterati in general - Masks have possibly had a big role in the surge of COVID cases and I am curious about the tweets around this in the context - Finally, if there’s anything around the election this year

Case-1:

Using the rhs property in the appearance, let’s get the ARM for keyword ‘fauci’

case\_one\_rules <- arules::apriori(TweetTrans,   
 parameter = list(support=0.001, confidence=0.01, minlen=3),  
 appearance = list(rhs='fauci'))

## Apriori  
##   
## Parameter specification:  
## confidence minval smax arem aval originalSupport maxtime support minlen  
## 0.01 0.1 1 none FALSE TRUE 5 0.001 3  
## maxlen target ext  
## 10 rules TRUE  
##   
## Algorithmic control:  
## filter tree heap memopt load sort verbose  
## 0.1 TRUE TRUE FALSE TRUE 2 TRUE  
##   
## Absolute minimum support count: 3   
##   
## set item appearances ...[1 item(s)] done [0.00s].  
## set transactions ...[7592 item(s), 3142 transaction(s)] done [0.01s].  
## sorting and recoding items ... [1671 item(s)] done [0.00s].  
## creating transaction tree ... done [0.00s].  
## checking subsets of size 1 2 3 4 5 6 7 8 9 10

## Warning in arules::apriori(TweetTrans, parameter = list(support = 0.001, :  
## Mining stopped (maxlen reached). Only patterns up to a length of 10 returned!

## done [7.46s].  
## writing ... [2024 rule(s)] done [3.72s].  
## creating S4 object ... done [1.02s].

arules::inspect(head(case\_one\_rules))

## lhs rhs support confidence coverage   
## [1] {fam,reversed} => {fauci} 0.003500955 1 0.003500955  
## [2] {charlesadler,reversed} => {fauci} 0.003500955 1 0.003500955  
## [3] {damage,reversed} => {fauci} 0.003500955 1 0.003500955  
## [4] {defamation,reversed} => {fauci} 0.003500955 1 0.003500955  
## [5] {reversed,threat} => {fauci} 0.003500955 1 0.003500955  
## [6] {decision,reversed} => {fauci} 0.003500955 1 0.003500955  
## lift count  
## [1] 174.5556 11   
## [2] 174.5556 11   
## [3] 174.5556 11   
## [4] 174.5556 11   
## [5] 174.5556 11   
## [6] 174.5556 11

# By Lift  
case\_one\_lift <- sort(case\_one\_rules, by='lift', decreasing=TRUE)  
arules::inspect(head(case\_one\_lift))

## lhs rhs support confidence coverage   
## [1] {fam,reversed} => {fauci} 0.003500955 1 0.003500955  
## [2] {charlesadler,reversed} => {fauci} 0.003500955 1 0.003500955  
## [3] {damage,reversed} => {fauci} 0.003500955 1 0.003500955  
## [4] {defamation,reversed} => {fauci} 0.003500955 1 0.003500955  
## [5] {reversed,threat} => {fauci} 0.003500955 1 0.003500955  
## [6] {decision,reversed} => {fauci} 0.003500955 1 0.003500955  
## lift count  
## [1] 174.5556 11   
## [2] 174.5556 11   
## [3] 174.5556 11   
## [4] 174.5556 11   
## [5] 174.5556 11   
## [6] 174.5556 11

## Plot wordcloud

# library(tm)  
my\_df1 <- DATAFRAME(case\_one\_rules)  
case\_one\_words <- gsub("\\{|\\}|,"," ", my\_df1$LHS)  
set.seed(1234)  
wordcloud(case\_one\_words[1:1500], min.freq=5, scale=c(1.5, .5), random.order=FALSE, rot.per=0.25,   
 colors=brewer.pal(8, "Dark2"))

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

## Warning in tm\_map.SimpleCorpus(corpus, function(x) tm::removeWords(x,  
## tm::stopwords())): transformation drops documents



Case-2:

Using the rhs property in the appearance, let’s get the ARM for keyword ‘masks’

case\_two\_rules <- arules::apriori(TweetTrans,   
 parameter = list(support=0.001, confidence=0.01, minlen=3),  
 appearance = list(rhs='masks'))

## Apriori  
##   
## Parameter specification:  
## confidence minval smax arem aval originalSupport maxtime support minlen  
## 0.01 0.1 1 none FALSE TRUE 5 0.001 3  
## maxlen target ext  
## 10 rules TRUE  
##   
## Algorithmic control:  
## filter tree heap memopt load sort verbose  
## 0.1 TRUE TRUE FALSE TRUE 2 TRUE  
##   
## Absolute minimum support count: 3   
##   
## set item appearances ...[1 item(s)] done [0.00s].  
## set transactions ...[7592 item(s), 3142 transaction(s)] done [0.01s].  
## sorting and recoding items ... [1671 item(s)] done [0.00s].  
## creating transaction tree ... done [0.00s].  
## checking subsets of size 1 2 3 4 5 6 7 8 9 10

## Warning in arules::apriori(TweetTrans, parameter = list(support = 0.001, :  
## Mining stopped (maxlen reached). Only patterns up to a length of 10 returned!

## done [6.87s].  
## writing ... [14898 rule(s)] done [3.73s].  
## creating S4 object ... done [0.99s].

arules::inspect(head(case\_two\_rules))

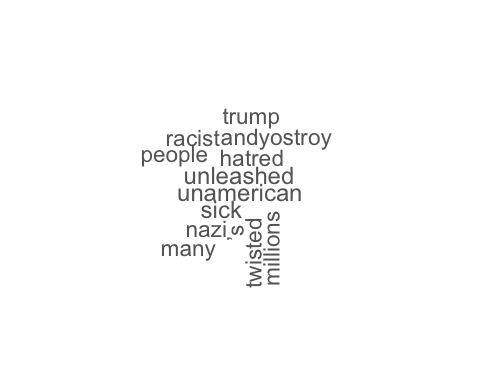
## lhs rhs support confidence coverage   
## [1] {unamerican,unleashed} => {masks} 0.008274984 1 0.008274984  
## [2] {sick,unamerican} => {masks} 0.008274984 1 0.008274984  
## [3] {twisted,unamerican} => {masks} 0.008274984 1 0.008274984  
## [4] {hatred,unamerican} => {masks} 0.008274984 1 0.008274984  
## [5] {millions,unamerican} => {masks} 0.008274984 1 0.008274984  
## [6] {andyostroy,unamerican} => {masks} 0.008274984 1 0.008274984  
## lift count  
## [1] 87.27778 26   
## [2] 87.27778 26   
## [3] 87.27778 26   
## [4] 87.27778 26   
## [5] 87.27778 26   
## [6] 87.27778 26

## Plot wordcloud

# library(tm)  
my\_df2 <- DATAFRAME(case\_two\_rules)  
case\_two\_words <- gsub("\\{|\\}|,"," ", my\_df2$LHS)  
set.seed(1234)  
wordcloud(case\_two\_words[1:1500], min.freq=3, scale=c(1.5, .5), random.order=FALSE, rot.per=0.25,   
 colors=brewer.pal(8, "Dark2"))

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

## Warning in tm\_map.SimpleCorpus(corpus, function(x) tm::removeWords(x,  
## tm::stopwords())): transformation drops documents



Self explanatory - likely topics revolved around Trump’s promoting/not-promoting masks (and possibly thought of as unamerican?).

Case-3:

Using the rhs property in the appearance, let’s get the ARM for keyword ‘covid’

# detach(package:tm, unload=TRUE)  
case\_three\_rules <- arules::apriori(TweetTrans,   
 parameter = list(support=0.001, confidence=0.01, minlen=3),  
 appearance = list(rhs='covid'))

## Apriori  
##   
## Parameter specification:  
## confidence minval smax arem aval originalSupport maxtime support minlen  
## 0.01 0.1 1 none FALSE TRUE 5 0.001 3  
## maxlen target ext  
## 10 rules TRUE  
##   
## Algorithmic control:  
## filter tree heap memopt load sort verbose  
## 0.1 TRUE TRUE FALSE TRUE 2 TRUE  
##   
## Absolute minimum support count: 3   
##   
## set item appearances ...[1 item(s)] done [0.00s].  
## set transactions ...[7592 item(s), 3142 transaction(s)] done [0.01s].  
## sorting and recoding items ... [1671 item(s)] done [0.00s].  
## creating transaction tree ... done [0.00s].  
## checking subsets of size 1 2 3 4 5 6 7 8 9 10

## Warning in arules::apriori(TweetTrans, parameter = list(support = 0.001, :  
## Mining stopped (maxlen reached). Only patterns up to a length of 10 returned!

## done [6.78s].  
## writing ... [7800 rule(s)] done [3.71s].  
## creating S4 object ... done [0.99s].

arules::inspect(head(case\_three\_rules))

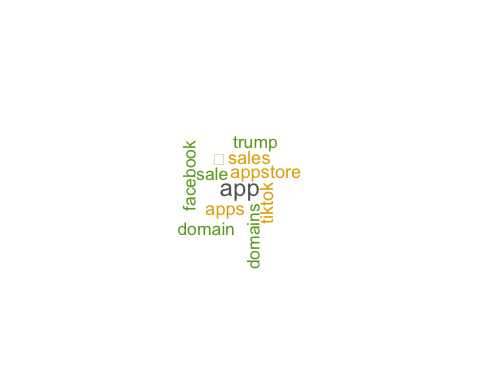
## lhs rhs support confidence coverage lift count  
## [1] {app,apps} => {covid} 0.001273074 1 0.001273074 98.1875 4   
## [2] {app,appstore} => {covid} 0.001273074 1 0.001273074 98.1875 4   
## [3] {app,sales} => {covid} 0.001273074 1 0.001273074 98.1875 4   
## [4] {app,tiktok} => {covid} 0.001273074 1 0.001273074 98.1875 4   
## [5] {ー,app} => {covid} 0.001273074 1 0.001273074 98.1875 4   
## [6] {app,domains} => {covid} 0.001273074 1 0.001273074 98.1875 4

## Plot wordcloud

# library(tm)  
my\_df3 <- DATAFRAME(case\_three\_rules)  
case\_three\_words <- gsub("\\{|\\}|,"," ", my\_df3$LHS)  
set.seed(1234)  
wordcloud(case\_three\_words[1:1500], min.freq=3, scale=c(1.5, .5), random.order=FALSE, rot.per=0.25,   
 colors=brewer.pal(8, "Dark2"))

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

## Warning in tm\_map.SimpleCorpus(corpus, function(x) tm::removeWords(x,  
## tm::stopwords())): transformation drops documents



## Displaying results from Apriori

# plot (SortedRules\_sup[1:50], method='graph', shading='confidence')

NOTE: I was having a lot of trouble with getting the visualize to work. Getting an error related to loading data.tables which is needed for arulesViz. I also spent hours upgrading my MAC to Catalina yet, no luck! Excerpt below: > install.packages(‘data.table’) . .

\*\* testing if installed package can be loaded from temporary location Error: package or namespace load failed for ‘data.table’ in library.dynam(lib, package, package.lib): shared object ‘datatable.so’ not found