Web Analytics - Assignment

**Group 4 – Jai, Sanjeev, Anu, Amol**

Problem 1- We have selected Yes Bank for finding the current sentiments fo the people. Further getting the data from twitter API and importing it.

#Connect to twitter with API

library(twitteR)

## Warning: package 'twitteR' was built under R version 3.6.2

library(rtweet)

## Warning: package 'rtweet' was built under R version 3.6.2

##   
## Attaching package: 'rtweet'

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

library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:twitteR':  
##   
## id, location

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(tidyr)  
library(tidytext)

## Warning: package 'tidytext' was built under R version 3.6.2

library(SnowballC)  
library(tm)

## Warning: package 'tm' was built under R version 3.6.2

## Loading required package: NLP

library(ggplot2)

##   
## Attaching package: 'ggplot2'

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

library(RColorBrewer)  
library(wordcloud)

## Warning: package 'wordcloud' was built under R version 3.6.2

library(topicmodels)

## Warning: package 'topicmodels' was built under R version 3.6.2

library(data.table)

##   
## Attaching package: 'data.table'

## The following objects are masked from 'package:dplyr':  
##   
## between, first, last

library(stringi)  
library(syuzhet)

## Warning: package 'syuzhet' was built under R version 3.6.2

##   
## Attaching package: 'syuzhet'

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

library(dplyr)  
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:twitteR':  
##   
## id

library(grid)  
library(gridExtra)

##   
## Attaching package: 'gridExtra'

## The following object is masked from 'package:dplyr':  
##   
## combine

consumer\_key <- "5IjiF7IZgkq3kHQxpLOBP4wxS"  
consumer\_secret<-"0UcaDU4A5BnenygXVNi09VNBw6bjp2pNVPH8fS5vWLkV8lvqNP"  
access\_token<-"1231460696300367872-K7r6aI0g5g0noA3aLGgNokU9MbBVY3"  
access\_token\_secret<-"sYqJKlpxNGuuyXqXpS5EvNqHJnICtZX5Q8hkgcxAzVsnM"

#Importing data

account<-"yes\_bank"

setwd("C:/Users/admin/Desktop/web analytics")

account.timeline<-searchTwitter(searchString = "YesBank", n=10000, lang="en")

trialdf<-twListToDF(account.timeline)

file.timeline<-paste(account, "tweets.csv", sep = "\_")

write.csv(trialdf, file.timeline)

#Read data

tweets.df <- read.csv("yes\_bank\_2tweets.csv")  
str(tweets.df)

## 'data.frame': 10000 obs. of 17 variables:  
## $ X : int 1 2 3 4 5 6 7 8 9 10 ...  
## $ text : Factor w/ 2745 levels "'Bankers make money, banks don't'\n#Yesbank #Banking \n\nI'm extremely sorry for the shareholders of Yesbank. \"| \_\_truncated\_\_,..: 2239 2051 1633 983 2257 2088 655 1503 1734 2420 ...  
## $ favorited : logi FALSE FALSE FALSE FALSE FALSE FALSE ...  
## $ favoriteCount: int 0 0 0 0 0 0 0 0 0 0 ...  
## $ replyToSN : Factor w/ 431 levels "\_\_soumitra","\_anujsinghal",..: NA NA NA 428 NA NA 339 NA NA NA ...  
## $ created : Factor w/ 8150 levels "2020-03-07 12:00:32",..: 8150 8149 8148 8148 8147 8146 8145 8144 8143 8142 ...  
## $ truncated : logi FALSE FALSE FALSE TRUE FALSE FALSE ...  
## $ replyToSID : num NA NA NA NA NA ...  
## $ id : num 1.24e+18 1.24e+18 1.24e+18 1.24e+18 1.24e+18 ...  
## $ replyToUID : num NA NA NA 4.2e+07 NA ...  
## $ statusSource : Factor w/ 47 levels "<a href=\"http://127.0.0.1:3000/\" rel=\"nofollow\">Twitter tweets 111</a>",..: 18 17 17 17 17 18 18 17 33 18 ...  
## $ screenName : Factor w/ 7444 levels "\_\_arpana\_\_","\_\_MunnaJi",..: 3180 2608 3807 2622 5769 6053 2588 5230 1181 5503 ...  
## $ retweetCount : int 2 1889 27 0 572 1750 0 1624 16 558 ...  
## $ isRetweet : logi TRUE TRUE TRUE FALSE TRUE TRUE ...  
## $ retweeted : logi FALSE FALSE FALSE FALSE FALSE FALSE ...  
## $ longitude : logi NA NA NA NA NA NA ...  
## $ latitude : logi NA NA NA NA NA NA ...

summary(tweets.df)

## X   
## Min. : 1   
## 1st Qu.: 2501   
## Median : 5000   
## Mean : 5000   
## 3rd Qu.: 7500   
## Max. :10000   
##   
## text   
## RT @pbhushan1: Anil Ambani (of Rafale fame) &amp; Subhash Chandra (of Sudhir Chowdhury fame) were among the biggest dubious groups to whom larg… :1633   
## RT @Sanju\_Verma\_: BIG--Look out circular issued against #RanaKapoor&amp;his wife #BinduKapoor,to prevent them from leaving the country\n\nED,unra…: 474   
## RT @sanjeev\_goyal: Adani Closed his accounts with YES bank last month.\nGujrat Govt bodies took out thier money from #YesBank just 2 days ag… : 399   
## RT @pbhushan1: YesBank's huge loans were to Cos that had already gamed the PSU banking system&amp;run up a massive debt which they were struggl… : 312   
## RT @Sanju\_Verma\_: #SBI to pick 49% stake in #YesBank for just Rs 2450 Cr\n\n48 hours back,49% stake would have cost Rs11760 Cr---Great for SB… : 293   
## RT @JayasreeVijayan: "History will be kind to me" said a prime minister \n\nUnder his PM ship there were NO \n\n\* Hindu Muslim riots \n\* Econom… : 213   
## (Other) :6676   
## favorited favoriteCount replyToSN   
## Mode :logical Min. : 0.000 YESBANK : 291   
## FALSE:10000 1st Qu.: 0.000 RBI : 18   
## Median : 0.000 ikaveri : 14   
## Mean : 1.155 nsitharaman : 14   
## 3rd Qu.: 0.000 sandipsabharwal: 13   
## Max. :3560.000 (Other) : 625   
## NA's :9025   
## created truncated replyToSID   
## 2020-03-07 13:02:34: 5 Mode :logical Min. :5.408e+17   
## 2020-03-07 17:39:14: 5 FALSE:9007 1st Qu.:1.236e+18   
## 2020-03-07 12:12:00: 4 TRUE :993 Median :1.236e+18   
## 2020-03-07 12:15:56: 4 Mean :1.232e+18   
## 2020-03-07 12:27:47: 4 3rd Qu.:1.236e+18   
## 2020-03-07 12:28:01: 4 Max. :1.236e+18   
## (Other) :9974 NA's :9228   
## id replyToUID   
## Min. :1.236e+18 Min. :8.167e+06   
## 1st Qu.:1.236e+18 1st Qu.:4.199e+07   
## Median :1.236e+18 Median :1.264e+08   
## Mean :1.236e+18 Mean :1.612e+17   
## 3rd Qu.:1.236e+18 3rd Qu.:2.362e+09   
## Max. :1.236e+18 Max. :1.236e+18   
## NA's :9025   
## statusSource   
## <a href="http://twitter.com/download/android" rel="nofollow">Twitter for Android</a>:6735   
## <a href="https://mobile.twitter.com" rel="nofollow">Twitter Web App</a> :1494   
## <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a> :1460   
## <a href="http://twitter.com/#!/download/ipad" rel="nofollow">Twitter for iPad</a> : 88   
## <a href="https://www.locobuzz.com" rel="nofollow">LocoBuzz India</a> : 59   
## <a href="http://twitter.com" rel="nofollow">Twitter Web Client</a> : 26   
## (Other) : 138   
## screenName retweetCount isRetweet retweeted   
## lovgill : 75 Min. : 0.0 Mode :logical Mode :logical   
## YESBANK : 60 1st Qu.: 8.0 FALSE:1685 FALSE:10000   
## SrikanthIyer10 : 44 Median : 163.0 TRUE :8315   
## reetej : 35 Mean : 611.9   
## Siddhes54664305: 34 3rd Qu.:1231.0   
## Mohamma67101207: 24 Max. :4498.0   
## (Other) :9728   
## longitude latitude   
## Mode:logical Mode:logical   
## NA's:10000 NA's:10000   
##   
##   
##   
##   
##

Problem 2 - EDA

# Remove retweets and replies

yes\_bank\_2tweets\_unique <- tweets.df[tweets.df$isRetweet==FALSE, ]  
yes\_bank\_2tweets\_unique\_nreplies <- subset(yes\_bank\_2tweets\_unique, is.na(yes\_bank\_2tweets\_unique$replyToSID))   
  
nrow(yes\_bank\_2tweets\_unique)

## [1] 1685

ncol(yes\_bank\_2tweets\_unique)

## [1] 17

#Exploring data for further analysis

# Remove retweets  
yes\_bank\_2tweets\_unique <- tweets.df[tweets.df$isRetweet==FALSE, ]   
nrow(yes\_bank\_2tweets\_unique)

## [1] 1685

# Remove replies  
yes\_bank\_2tweets\_unique <- subset(yes\_bank\_2tweets\_unique, is.na(yes\_bank\_2tweets\_unique$replyToSID))   
nrow(yes\_bank\_2tweets\_unique)

## [1] 913

# Analyse engagement by looking at the variables: favoriteCount (i.e. the number of likes)  
# or retweetCount (i.e. the number of retweets). Simply arrange them in descending order (with a minus “-” before the variable)   
# to find the one with the highest number of likes or retweets or ascending order (without the minus) to find the one with lowest   
# number of engagements.  
  
# Msst favourite  
yes\_bank\_2tweets\_unique <- yes\_bank\_2tweets\_unique %>% arrange(-favoriteCount)  
yes\_bank\_2tweets\_unique[1,5]

## [1] <NA>  
## 431 Levels: \_\_soumitra \_anujsinghal \_dineshkp \_Mukul ... Yuyudhana

# Highest retweets  
yes\_bank\_2tweets\_unique <- yes\_bank\_2tweets\_unique %>% arrange(-retweetCount)  
yes\_bank\_2tweets\_unique[1,5]

## [1] <NA>  
## 431 Levels: \_\_soumitra \_anujsinghal \_dineshkp \_Mukul ... Yuyudhana

# Least favourite  
yes\_bank\_2tweets\_unique <- yes\_bank\_2tweets\_unique %>% arrange(favoriteCount)  
yes\_bank\_2tweets\_unique[1,5]

## [1] <NA>  
## 431 Levels: \_\_soumitra \_anujsinghal \_dineshkp \_Mukul ... Yuyudhana

Problem 3 - Data cleaning # Create document corpus with tweet text

myCorpus<- Corpus(VectorSource(yes\_bank\_2tweets\_unique$text))   
myCorpuscopy<-Corpus(VectorSource(yes\_bank\_2tweets\_unique$text))   
myCorpus <- tm\_map(myCorpus, content\_transformer(stri\_trans\_tolower))

## Warning in tm\_map.SimpleCorpus(myCorpus,  
## content\_transformer(stri\_trans\_tolower)): transformation drops documents

writeLines(strwrap(myCorpus[[750]]$content,60))

## don't we learn anything from history, from our past  
## mistakes, from other people's mistakes. is anyone  
## accountable a… https://t.co/xbsjade9ol

#removing URLs

removeURL <- function(x) gsub("http[^[:space:]]\*", "", x)   
myCorpus <- tm\_map(myCorpus, content\_transformer(removeURL))

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

writeLines(strwrap(myCorpus[[750]]$content,60))

## don't we learn anything from history, from our past  
## mistakes, from other people's mistakes. is anyone  
## accountable a…

#Removing user names

removeUsername <- function(x) gsub("@[^[:space:]]\*", "", x)   
myCorpus <- tm\_map(myCorpus, content\_transformer(removeUsername))

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

writeLines(strwrap(myCorpus[[751]]$content,60))

## adani knew it before rbi <u+0001f928>

#Removing stopwords

myStopWords<- c((stopwords('english')),c("rt","and","the","yesbank","yes", "bank","whom", "into", "via","rs","will","can", "now", "hi","get","cr","sb","ag","pick","took","larg", "groups", "ufufaf","uf", "ufufuf", "sir","said","ub","th"))  
myCorpus<- tm\_map(myCorpus,removeWords , myStopWords)

## Warning in tm\_map.SimpleCorpus(myCorpus, removeWords, myStopWords):  
## transformation drops documents

writeLines(strwrap(myCorpus[[755]]$content,60))

## #<u+092a><u+0940><u+090f><u+092e>\_<u+092a><u+0928><u+094c><u+0924><u+0940>  
## #yes\_bank #yesbankscam # #yesbankcrises  
## #\_suspend\_delhi\_govt  
##   
## follow <u+0001f447><u+0001f4af> followback…

#Removing panctuation

removeNumPunct <- function(x) gsub("[^[:alpha:][:space:]]\*", "", x)   
myCorpus <- tm\_map(myCorpus, content\_transformer(removeNumPunct))

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

writeLines(strwrap(myCorpus[[755]]$content,60))

## uauufueuauucuu yesbank yesbankscam  
##   
## yesbankcrises suspenddelhigovt  
##   
## follow ufufaf followback

#Removing single letter word

removeSingle <- function(x) gsub(" . ", " ", x)   
myCorpus <- tm\_map(myCorpus, content\_transformer(removeSingle))

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

writeLines(strwrap(myCorpus[[758]]$content,60))

## at bonds write help absorb losses tune approx sbi got sweet  
## deal many inves

#Removing extra whitespaces

myCorpus<- tm\_map(myCorpus, stripWhitespace)

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

writeLines(strwrap(myCorpus[[758]]$content,60))

## at bonds write help absorb losses tune approx sbi got sweet  
## deal many inves

#Creting document term matrix

tdm<- TermDocumentMatrix(myCorpus, control= list(wordLengths= c(1, Inf)))  
tdm

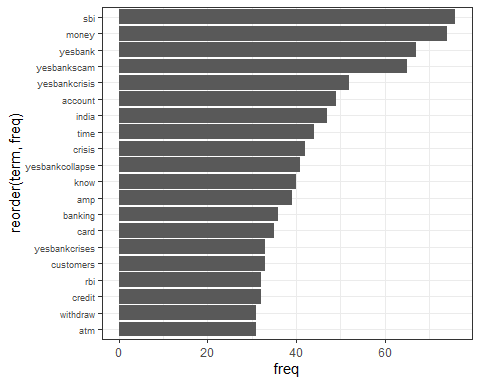
## <<TermDocumentMatrix (terms: 2761, documents: 913)>>  
## Non-/sparse entries: 7170/2513623  
## Sparsity : 100%  
## Maximal term length: 68  
## Weighting : term frequency (tf)

#Problem 4 - Ananlysing text frequency

(freq.terms <- findFreqTerms(tdm, lowfreq = 30))

## [1] "account" "know" "money" "time"   
## [5] "yesbank" "yesbankscam" "yesbankcrisis" "india"   
## [9] "yesbankcrises" "customers" "yesbankcollapse" "amp"   
## [13] "crisis" "rbi" "withdraw" "sbi"   
## [17] "card" "banking" "atm" "credit"

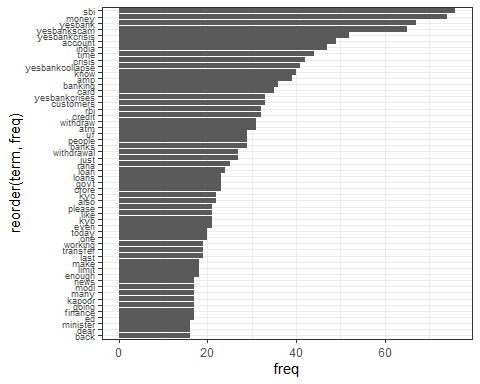
term.freq <- rowSums(as.matrix(tdm))  
term.freq <- subset(term.freq, term.freq > 30)  
df <- data.frame(term = names(term.freq), freq= term.freq)  
p1=ggplot(df, aes(reorder(term, freq),freq)) + theme\_bw() + geom\_bar(stat = "identity") + coord\_flip() +labs(list(title="@10", x="Terms", y="Term Counts")) + theme(axis.text.y = element\_text(size=7))  
p1



(freq.terms <- findFreqTerms(tdm, lowfreq = 15))

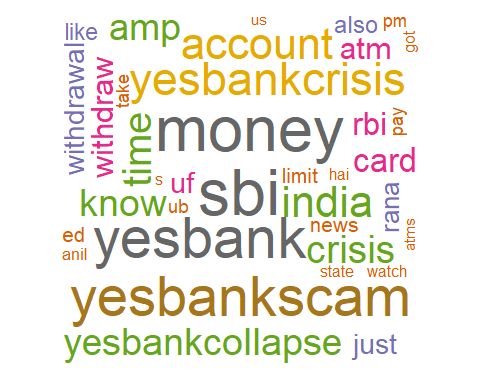
## [1] "account" "just" "know" "kyc"   
## [5] "many" "money" "time" "loan"   
## [9] "back" "please" "yesbank" "yesbankscam"   
## [13] "yesbankcrisis" "last" "dear" "enough"   
## [17] "india" "kyb" "uf" "yesbankcrises"   
## [21] "cash" "customers" "withdrawal" "yesbankcollapse"  
## [25] "going" "amp" "crisis" "rbi"   
## [29] "banks" "people" "able" "limit"   
## [33] "withdraw" "ub" "open" "kapoor"   
## [37] "rana" "sbi" "make" "even"   
## [41] "day" "crore" "card" "news"   
## [45] "banking" "govt" "atm" "today"   
## [49] "one" "like" "transfer" "modi"   
## [53] "credit" "must" "working" "also"   
## [57] "loans" "finance" "minister" "founder"   
## [61] "financial" "ed"

term.freq <- rowSums(as.matrix(tdm))  
term.freq <- subset(term.freq, term.freq > 15)  
df1 <- data.frame(term = names(term.freq), freq= term.freq)  
  
p2=ggplot(df1, aes(reorder(term, freq),freq)) + theme\_bw() + geom\_bar(stat = "identity") + coord\_flip() +labs(list(title="@10", x="Terms", y="Term Counts")) + theme(axis.text.y = element\_text(size=7))  
p2

 ###Considering low frequency as 30.

#Problem 5 - Creating word cloud

#Overall word cloud  
word.freq <-sort(rowSums(as.matrix(tdm)), decreasing= F)  
pal<- brewer.pal(8, "Dark2")  
wordcloud(words = names(word.freq), freq = word.freq, min.freq = 2, random.order = F, colors = pal, max.words = 150)



dtm <- as.DocumentTermMatrix(tdm)  
  
#Calculating Sentiments for word cloud  
freq\_up <- colSums(as.matrix(dtm))  
##install.packages('RSentiment')  
library('RSentiment')

## Warning: package 'RSentiment' was built under R version 3.6.3

sentiments\_up = calculate\_sentiment(names(freq\_up))

sentiments\_up = cbind(sentiments\_up, as.data.frame(freq\_up))  
sent\_pos\_up = sentiments\_up[sentiments\_up$sentiment == 'Positive',]  
sent\_neg\_up = sentiments\_up[sentiments\_up$sentiment == 'Negative',]  
  
cat("We have relatively lower negative Sentiments: ",sum(sent\_neg\_up$freq\_up)," than positive: ",sum(sent\_pos\_up$freq\_up))

## We have relatively lower negative Sentiments: 690 than positive: 948

#Create Positive WordCloud  
  
layout(matrix(c(1, 2), nrow=2), heights=c(1, 4))  
par(mar=rep(0, 4))  
plot.new()  
set.seed(100)  
wordcloud(sent\_pos\_up$text,sent\_pos\_up$freq,min.freq=10,colors=brewer.pal(6,"Dark2"))

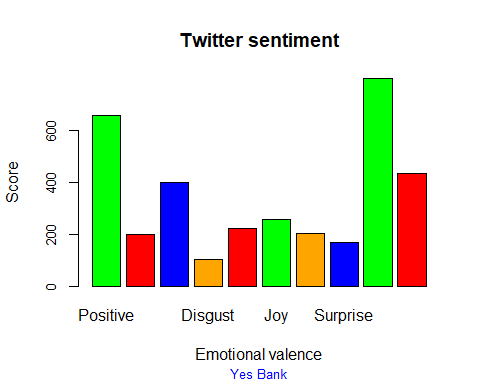


#Create Negative WordCloud  
  
plot.new()  
set.seed(100)  
wordcloud(sent\_neg\_up$text,sent\_neg\_up$freq, min.freq=10,colors=brewer.pal(6,"Dark2"))



#Problem 6 - Obtaining sentiment scores

yes\_bank\_2tweets\_unique$text<-as.character(yes\_bank\_2tweets\_unique$text)  
mysentiment<- get\_nrc\_sentiment(yes\_bank\_2tweets\_unique$text)  
  
# Get the sentiment score for each emotion  
mysentiment.positive =sum(mysentiment$positive)  
mysentiment.anger =sum(mysentiment$anger)  
mysentiment.anticipation =sum(mysentiment$anticipation)  
mysentiment.disgust =sum(mysentiment$disgust)  
mysentiment.fear =sum(mysentiment$fear)  
mysentiment.joy =sum(mysentiment$joy)  
mysentiment.sadness =sum(mysentiment$sadness)  
mysentiment.surprise =sum(mysentiment$surprise)  
mysentiment.trust =sum(mysentiment$trust)  
mysentiment.negative =sum(mysentiment$negative)  
  
# Create the bar chart  
yAxis <- c(mysentiment.positive,  
 + mysentiment.anger,  
 + mysentiment.anticipation,  
 + mysentiment.disgust,  
 + mysentiment.fear,  
 + mysentiment.joy,  
 + mysentiment.sadness,  
 + mysentiment.surprise,  
 + mysentiment.trust,  
 + mysentiment.negative)  
  
xAxis <- c("Positive","Anger","Anticipation","Disgust","Fear","Joy","Sadness","Surprise","Trust","Negative")  
colors <- c("green","red","blue","orange","red","green","orange","blue","green","red")  
yRange <- range(0,yAxis)  
barplot(yAxis, names.arg = xAxis,   
 xlab = "Emotional valence", ylab = "Score", main = "Twitter sentiment",   
 sub = "Yes Bank", col = colors, border = "black", xpd = F, ylim = yRange,  
 axisnames = T, cex.axis = 0.8, cex.sub = 0.8, col.sub = "blue")



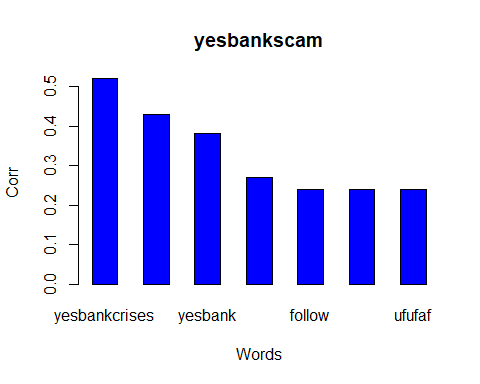
#Considering the graph above we can conclude that, due to recent turbulance there are some negative sentiments for Yes Bank but over all people seems to have trust and positive attitude towards the bank.

Problem 7 - Correlation #Correlation with top key words and assiciation

list1<- findAssocs(tdm, "yesbankscam", 0.2)  
corrdf1 <- t(data.frame(t(sapply(list1,c))))  
corrdf1

## yesbankscam  
## yesbankcrises 0.52  
## yesbankcollapse 0.43  
## yesbank 0.38  
## suspenddelhigovt 0.27  
## follow 0.24  
## followback 0.24  
## ufufaf 0.24

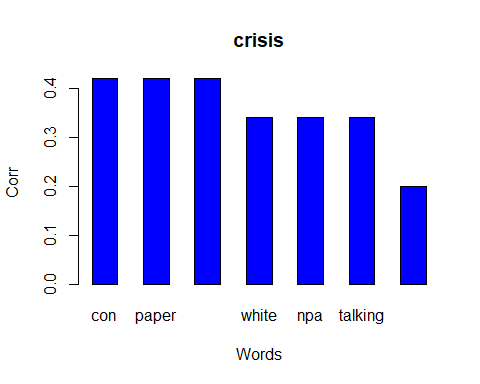
barplot(t(as.matrix(corrdf1)), beside=TRUE,xlab = "Words",ylab = "Corr",col = "blue",main = "yesbankscam",border = "black")



list2<- findAssocs(tdm, "crisis", 0.2)  
corrdf2 <- t(data.frame(t(sapply(list2,c))))  
corrdf2

## crisis  
## con 0.42  
## paper 0.42  
## sharing 0.42  
## white 0.34  
## npa 0.34  
## talking 0.34  
## speaks 0.20

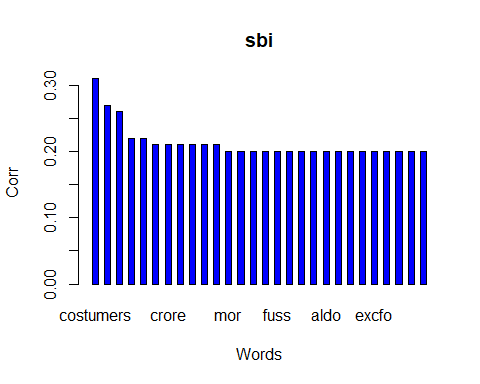
barplot(t(as.matrix(corrdf2)), beside=TRUE,xlab = "Words",ylab = "Corr",col = "blue",main = "crisis",border = "black")



list3<- findAssocs(tdm, "sbi", 0.2)  
corrdf3 <- t(data.frame(t(sapply(list3,c))))  
corrdf3

## sbi  
## costumers 0.31  
## bail 0.27  
## investment 0.26  
## kumar 0.22  
## rescue 0.22  
## acquiring 0.21  
## crore 0.21  
## zero 0.21  
## chief 0.21  
## rajnishkumar 0.21  
## prashant 0.21  
## mor 0.20  
## altogether 0.20  
## bailed 0.20  
## eventuallymerge 0.20  
## fuss 0.20  
## precedents 0.20  
## warranted 0.20  
## upper 0.20  
## aldo 0.20  
## invests 0.20  
## administrator 0.20  
## de 0.20  
## excfo 0.20  
## release 0.20  
## ufmr 0.20  
## forcing 0.20  
## pushes 0.20

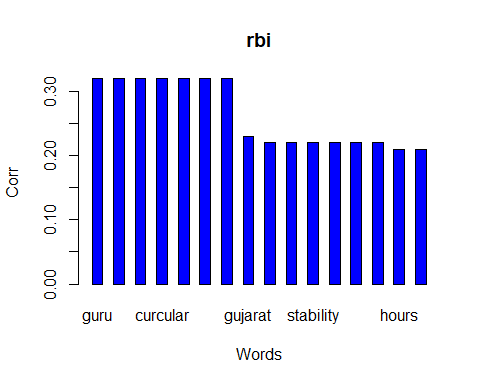
barplot(t(as.matrix(corrdf3)), beside=TRUE,xlab = "Words",ylab = "Corr",col = "blue",main = "sbi",border = "black")



list4<- findAssocs(tdm, "rbi", 0.2)  
corrdf4 <- t(data.frame(t(sapply(list4,c))))  
corrdf4

## rbi  
## guru 0.32  
## raghavendra 0.32  
## achieved 0.32  
## curcular 0.32  
## maintain 0.32  
## within 0.32  
## draftscheme 0.32  
## gujarat 0.23  
## bangalore 0.22  
## helping 0.22  
## stability 0.22  
## weeks 0.22  
## reconstruction 0.22  
## approve 0.22  
## hours 0.21  
## draft 0.21

barplot(t(as.matrix(corrdf4)), beside=TRUE,xlab = "Words",ylab = "Corr",col = "blue",main = "rbi",border = "black")



list5<- findAssocs(tdm, "cash", 0.2)  
corrdf5 <- t(data.frame(t(sapply(list5,c))))  
corrdf5

## cash  
## chennai 0.36  
## dd 0.36  
## mogappair 0.36  
## branches 0.31  
## visited 0.29  
## facility 0.29  
## mantras 0.26  
## unfortunately 0.26  
## handing 0.26  
## madhapur 0.26  
## notch 0.26  
## staying 0.26  
## bran 0.26  
## crowded 0.26  
## petty 0.26  
## didu 0.26  
## wasted 0.26  
## countries 0.26  
## stranded 0.26  
## assisting 0.26  
## coupons 0.26  
## pigs 0.26  
## pity 0.26  
## rush 0.26  
## transfernet 0.26  
## bhopal 0.26  
## dry 0.26  
## scramble 0.26  
## hand 0.25  
## branch 0.23

barplot(t(as.matrix(corrdf5)), beside=TRUE,xlab = "Words",ylab = "Corr",col = "blue",main = "cash",border = "black")

