VASenateAnalysis

Georges Colbert

6/12/2018

## Loading required package: RColorBrewer

## Loading required package: NLP

##   
## 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

## Warning: Missing column names filled in: 'X1' [1]

## Parsed with column specification:  
## cols(  
## X1 = col\_integer(),  
## text = col\_character(),  
## favorited = col\_logical(),  
## favoriteCount = col\_integer(),  
## replyToSN = col\_character(),  
## created = col\_datetime(format = ""),  
## truncated = col\_logical(),  
## replyToSID = col\_double(),  
## id = col\_double(),  
## replyToUID = col\_integer(),  
## statusSource = col\_character(),  
## screenName = col\_character(),  
## retweetCount = col\_integer(),  
## isRetweet = col\_logical(),  
## retweeted = col\_logical(),  
## longitude = col\_character(),  
## latitude = col\_character()  
## )

## Warning in rbind(names(probs), probs\_f): number of columns of result is not  
## a multiple of vector length (arg 1)

## Warning: 18 parsing failures.  
## row # A tibble: 5 x 5 col row col expected actual expected <int> <chr> <chr> <chr> actual 1 1603 replyToUID an integer 2720054437 file 2 2077 replyToUID an integer 832829346385686529 row 3 2751 replyToUID an integer 2720054437 col 4 3470 replyToUID an integer 2880381305 expected 5 4698 replyToUID an integer 822844948433301504 actual # ... with 1 more variables: file <chr>  
## ... ................. ... ................................................ ........ ................................................ ...... ................................................ .... ................................................ ... ................................................ ... ................................................ ........ ................................................ ...... .......................................  
## See problems(...) for more details.

# Total Number of tweets

nrow(VA.tweetdf)

## [1] 8361

# How many are retweets?

##### How many are retweets?  
   
 table(VA.tweetdf$isRetweet)

##   
## FALSE TRUE   
## 341 8020

### list of top 6 favorite tweets of the period

## text  
## 1 I'm on my way to go vote. \n\nI get so excited because it makes me think of the day we defeated Hillary Clinton. \n#TuesdayThoughts #VAsen  
## 2 Today, I voted in the Virginia Republican Primary! \n\nI voted Republican because I'm happy with the booming economy,… https://t.co/jx3MimabVV  
## 3 The polls are now closed in South Carolina &amp; Virginia. \n\nI'm watching the election results like a hawk. \n#VAsen  
## 4 Virginians, \n\nYou have 1 hour &amp; 15 minutes left to vote. The polls close @ 7 p.m. Once you're in line, stay in line… https://t.co/BXjSR6bHTE  
## 5 The reason I’m running for United States Senate is because I believe in the promise of this country.\n\nVote tomorrow… https://t.co/Cg3qNlytoM  
## 6 Today is the big day. I hope I can earn your support in my candidacy for United States Senate! \n\nPolls are open 6:0… https://t.co/HiDyvHXKxa  
## favoriteCount created screenName  
## 1 4615 2018-06-12 16:04:45 ScottPresler  
## 2 3847 2018-06-12 16:56:02 ScottPresler  
## 3 1708 2018-06-12 23:11:18 ScottPresler  
## 4 1634 2018-06-12 21:46:50 ScottPresler  
## 5 1581 2018-06-12 01:26:14 NickForVA  
## 6 918 2018-06-12 10:11:22 NickForVA

# top 6 retweeted tweets of the period

## text  
## 1 I'm on my way to go vote. \n\nI get so excited because it makes me think of the day we defeated Hillary Clinton. \n#TuesdayThoughts #VAsen  
## 2 Today, I voted in the Virginia Republican Primary! \n\nI voted Republican because I'm happy with the booming economy,… https://t.co/jx3MimabVV  
## 3 Virginians, \n\nYou have 1 hour &amp; 15 minutes left to vote. The polls close @ 7 p.m. Once you're in line, stay in line… https://t.co/BXjSR6bHTE  
## 4 The polls are now closed in South Carolina &amp; Virginia. \n\nI'm watching the election results like a hawk. \n#VAsen  
## 5 Today is the big day. I hope I can earn your support in my candidacy for United States Senate! \n\nPolls are open 6:0… https://t.co/HiDyvHXKxa  
## 6 The reason I’m running for United States Senate is because I believe in the promise of this country.\n\nVote tomorrow… https://t.co/Cg3qNlytoM  
## retweetCount created screenName  
## 1 1389 2018-06-12 16:04:45 ScottPresler  
## 2 1320 2018-06-12 16:56:02 ScottPresler  
## 3 856 2018-06-12 21:46:50 ScottPresler  
## 4 470 2018-06-12 23:11:18 ScottPresler  
## 5 457 2018-06-12 10:11:22 NickForVA  
## 6 440 2018-06-12 01:26:14 NickForVA

### Top ten retweets overall

In this case, there were tweets that were creeated before the week that were still being shared on Twitter and accumulating retweets.

I <- VA.tweetdf[order(VA.tweetdf$retweetCount, decreasing=TRUE),]  
  
  
 I <- I[!duplicated(I$text),]  
  
  
I <- head(select(I,text,retweetCount,created,screenName,isRetweet),10)  
   
 as.data.frame(I)

## text  
## 1 RT @ScottPresler: Tomorrow is Election Day in Virginia &amp; I will be voting Republican. \n\nNothing is going to stop me from voting tomorrow or…  
## 2 RT @ScottPresler: I'm on my way to go vote. \n\nI get so excited because it makes me think of the day we defeated Hillary Clinton. \n#TuesdayT…  
## 3 I'm on my way to go vote. \n\nI get so excited because it makes me think of the day we defeated Hillary Clinton. \n#TuesdayThoughts #VAsen  
## 4 RT @ScottPresler: Today, I voted in the Virginia Republican Primary! \n\nI voted Republican because I'm happy with the booming economy, low u…  
## 5 Today, I voted in the Virginia Republican Primary! \n\nI voted Republican because I'm happy with the booming economy,… https://t.co/jx3MimabVV  
## 6 RT @ScottPresler: Virginians, \n\nYou have 1 hour &amp; 15 minutes left to vote. The polls close @ 7 p.m. Once you're in line, stay in line. It's…  
## 7 Virginians, \n\nYou have 1 hour &amp; 15 minutes left to vote. The polls close @ 7 p.m. Once you're in line, stay in line… https://t.co/BXjSR6bHTE  
## 8 RT @NickForVA: I have an A+ rating from the @NRA and I am the only candidate endorsed by the @NRAPVF in this election. \n\nOur Second Amendme…  
## 9 RT @Stump\_for\_Trump: Green Beret Combat Veteran and member of the Virginia House of Delegates @NickForVA will crush @timkaine this November…  
## 10 RT @NickForVA: Don't let politicians deceive you. Vote Nick Freitas on June 12. #LibertyRising #VASen #VAPol https://t.co/AuA98g9Yua  
## retweetCount created screenName isRetweet  
## 1 3084 2018-06-12 23:51:23 pbprecious7 TRUE  
## 2 1389 2018-06-12 23:57:38 miket4300 TRUE  
## 3 1389 2018-06-12 16:04:45 ScottPresler FALSE  
## 4 1320 2018-06-12 23:59:45 JimPolk TRUE  
## 5 1320 2018-06-12 16:56:02 ScottPresler FALSE  
## 6 856 2018-06-12 23:59:45 svagdis TRUE  
## 7 856 2018-06-12 21:46:50 ScottPresler FALSE  
## 8 611 2018-06-12 22:32:18 TheNeaL TRUE  
## 9 575 2018-06-12 22:47:15 Juicerden TRUE  
## 10 521 2018-06-12 22:31:36 TheNeaL TRUE

## accounts that have the most total Retweets

## screenName x  
## 155 ScottPresler 4148  
## 121 NickForVA 1766  
## 80 JackPosobiec 1146  
## 35 CoreyStewartVA 308  
## 169 Stump\_for\_Trump 297  
## 115 michael\_\_1776 197

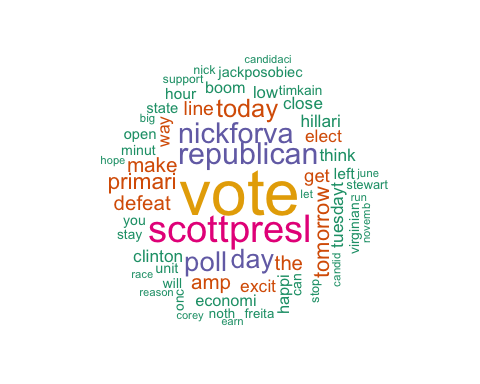
## Most Active Accounts (including Retweets)

## Var1 Freq  
## 3440 michael\_\_1776 76  
## 3431 MFEUZ3 32  
## 4030 Phibear94 32  
## 5193 TheRickCanton 20  
## 2179 ItsChad 18  
## 3910 ParentWayne 18

## Most Active Accounts

## Var1 Freq  
## 80 JackPosobiec 14  
## 77 IanTreyParish 10  
## 183 VCHoos 8  
## 121 NickForVA 7  
## 48 DKElections 6  
## 114 MFEUZ3 6

# WORDCLOUD of most common used words



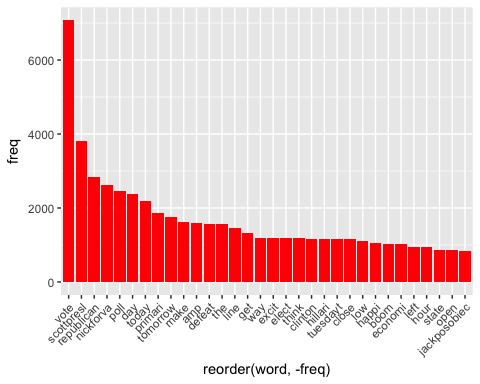
# List of The Top 15 Frequently used Words

## word freq  
## vote vote 7072  
## scottpresl scottpresl 3822  
## republican republican 2840  
## nickforva nickforva 2627  
## poll poll 2461  
## day day 2383  
## today today 2198  
## primari primari 1861  
## tomorrow tomorrow 1766  
## make make 1611  
## amp amp 1586  
## defeat defeat 1575  
## the the 1571  
## line line 1447  
## get get 1325

# Plot of Words that Appear at least 800 times

##   
## Attaching package: 'ggplot2'

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



## Words Associated with

To find out what words are associated with the candidates, I use a function called FindAssocs() on the dataset. For any given word, findAssocs() calculates its correlation with every other word in the dataset. Scores range from 0 to 1. A score of 1 means that two words always appear together, while a score of 0 means that they never appear together.

# which words are associated with "Corey Stewart"  
   
 findAssocs(tdm,"coreystewartva", 0.1)

## $coreystewartva  
## oann studio set talk   
## 0.53 0.53 0.49 0.48   
## also want border billkristol   
## 0.41 0.39 0.34 0.33   
## but korea nkorea protect   
## 0.32 0.32 0.32 0.32   
## troop usmexico novelect boehner   
## 0.32 0.32 0.32 0.26   
## john magazin ole pine   
## 0.26 0.26 0.26 0.26   
## rag they hdown scarletthara   
## 0.26 0.26 0.24 0.23   
## libertarian yesterday daili earlier   
## 0.23 0.22 0.21 0.21   
## maga listen vagop anymor   
## 0.21 0.21 0.21 0.20   
## infect media roanoketim trumpism   
## 0.20 0.20 0.20 0.20   
## former tru mainstream realdonaldtrump   
## 0.20 0.20 0.19 0.19   
## for apolog neocon nevertrump   
## 0.17 0.17 0.16 0.16   
## winegirl jackposobiec corey pleas   
## 0.16 0.15 0.15 0.15   
## antiamerican imm warmong interview   
## 0.15 0.15 0.15 0.14   
## gop shoahara kristol good   
## 0.14 0.14 0.14 0.13   
## thank bill scare oppos   
## 0.13 0.13 0.12 0.12   
## benpersh zacharycohen ladyh rig   
## 0.12 0.12 0.12 0.12   
## seek cathol sheepdog great   
## 0.12 0.12 0.12 0.11   
## say fellow effort futil   
## 0.11 0.11 0.11 0.11   
## goper lastditch read stewart   
## 0.11 0.11 0.11 0.10   
## someth   
## 0.10

findAssocs(tdm,"nickforva", 0.1)

## $nickforva  
## unit can state timkain open   
## 0.47 0.45 0.43 0.41 0.40   
## big candidaci earn novemb reason   
## 0.34 0.33 0.33 0.33 0.33   
## hope believ countri promis imper   
## 0.32 0.32 0.32 0.32 0.30   
## june run let support candid   
## 0.30 0.30 0.29 0.28 0.28   
## stumpfortrump begin take back senat   
## 0.27 0.26 0.26 0.25 0.25   
## poll sure find nearest shut   
## 0.24 0.23 0.22 0.22 0.22   
## two virg libertyris end commonwealth   
## 0.22 0.22 0.22 0.21 0.21   
## less precinct politician with dont   
## 0.21 0.20 0.20 0.19 0.19   
## vapol booth continu even lie   
## 0.19 0.18 0.18 0.18 0.18   
## oppon veteran differ final interac   
## 0.18 0.17 0.16 0.16 0.16   
## posit visit beret green choos   
## 0.16 0.16 0.16 0.16 0.16   
## opportun pray shanethazel morn tina   
## 0.16 0.16 0.16 0.15 0.15   
## nation pro voter record great   
## 0.15 0.15 0.13 0.13 0.12   
## man combat must endors apliberti   
## 0.12 0.12 0.12 0.11 0.11   
## jefferson libertarian iraq join reevesva   
## 0.11 0.11 0.11 0.11 0.11   
## onli interview dele rlibertycaucus   
## 0.11 0.10 0.10 0.10

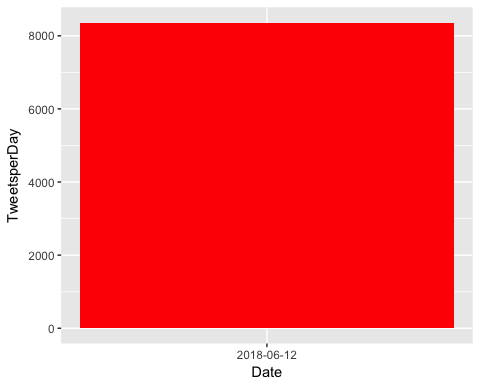
findAssocs(tdm,"kaine",0.1 )

## $kaine  
## numeric(0)

# Number of Tweets per Day

##   
## Attaching package: 'scales'

## The following object is masked from 'package:readr':  
##   
## col\_factor



# Top 6 Negative Tweets

## text  
## 1 lol i bet everyone is freaking out because someone just made a huge repoing error vasen //tco/yjerbydkmz  
## 2 virginia beach definitely made a tabulation error in the vasen primary no way that ewjacksonsr got that many abs //tco/c7ae0wwt5e  
## 3 coreystewava cant beat nick unless he makes up flat out lies he has no honor or integrity virginia needs to r //tco/f6rug3i0dn  
## 4 the vaelect returns have an input error somewhere or ew jackson is about to shock the world in vasen //tco/tvdsiyeecp  
## 5 bad for stewavasen //tco/bbfndtnpoz  
## 6 who's losing more votes to the stanley cup parade nick freitas or corey stewavasen  
## retweetCount created screenName  
## 1 113 2018-06-12 23:38:31 ScottPresler  
## 2 14 2018-06-12 23:50:03 Spires4Office  
## 3 1 2018-06-12 19:43:42 MFEUZ3  
## 4 4 2018-06-12 23:34:14 geoffreyvs  
## 5 0 2018-06-12 19:50:47 michael\_\_1776  
## 6 0 2018-06-12 13:13:25 ChestyFitz

# Top 6 Positive Tweets

## text  
## 1 very very good newsvasen //tco/jqurvseok2  
## 2 please vote for nickforva in the vasen primary he loves occupational licensing reform if he's in senate i will //tco/lxw5wgcino  
## 3 hearing turnout in vasen primary very low today despite good weather for the entire state  
## 4 do you care about reforming veterans affairs vote nickforva vasen vapol libeyrising  
## 5 throwback to rmconservative's interview of nickforva nick freitas' brilliant ability to provide our narrative an //tco/0nodsjax1i  
## 6 happily casted my ballot weeks ago for nickforva a true conservative who values our freedoms and isnt a career //tco/ov886y5qn2  
## retweetCount created screenName  
## 1 1 2018-06-12 23:46:22 MattMackowiak  
## 2 8 2018-06-12 12:32:14 senatorshoshana  
## 3 74 2018-06-12 21:21:47 JackPosobiec  
## 4 2 2018-06-12 14:44:06 IanTreyParish  
## 5 24 2018-06-12 02:34:29 michael\_\_1776  
## 6 3 2018-06-12 16:58:03 ItsChad

## geom\_path: Each group consists of only one observation. Do you need to  
## adjust the group aesthetic?

