Tim Kaine Mentions Analysis

Georges Colbert

07/15/2018

# Period of July 08 to July 14

Our goal is to cover the conversations being had about the candidates across multiple platforms, interpret it and visualize the information in a way that can help drive decision making. The Graphs shown and metrics calculated are best estimation of what can be most useful to the candidate. Through a process of constant feedback between us and the campaign, the most relevant information will be identified and facilitate the creation of a personalized dashboard that the campaign can access 24/7. For now, multiple visualizations of the same information could be provided as we try to identify the best way for the campaign to retain the information and identify the KPIs amongst the myriad of metrics available.

In the case of Twitter, it is really fortunate that twitter provides an API on which developers can request an acquire a substantial amount of tweets from the previous ten days before the API call.

For more information on the twitter API the link is:

<https://developer.twitter.com/en/docs/tweets/search/api-reference/get-search-tweets.html>

To facilitate the process of Data mining the tweets, here’s a list of packages used:

* [TwitteR Package](https://cran.r-project.org/web/packages/twitteR/twitteR.pdf)
* [TM package](https://cran.r-project.org/web/packages/tm/index.html)
* [ggplot2 package](https://cran.r-project.org/web/packages/ggplot2/index.html)
* [wordcloud package](https://cran.r-project.org/web/packages/wordcloud/wordcloud.pdf)
* [SnowballC package](https://cran.r-project.org/web/packages/SnowballC/SnowballC.pdf)
* [SentimentR package](https://cran.r-project.org/web/packages/sentimentr/sentimentr.pdf)

For this analysis, a request was made to the API for up to 30,000 tweets that mentioned either "Tim Kaine" or "@timkaine" covering the period of July 08-July 14. All the tweets returned by the API are in the zip folder.

**OVERVIEW**

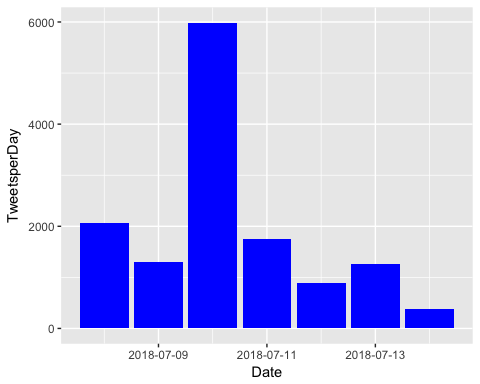
# Total Number of tweets

## [1] 13624

# Time-Series of Tweets Per Day

##   
## Attaching package: 'scales'

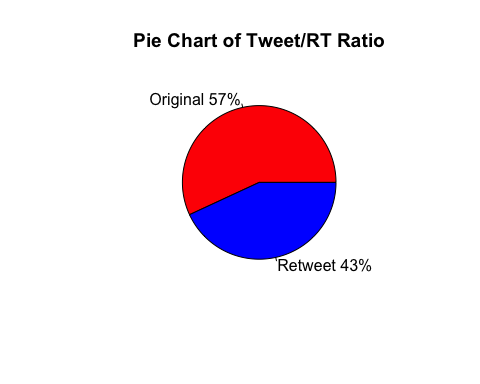
## The following object is masked from 'package:readr':  
##   
## col\_factor



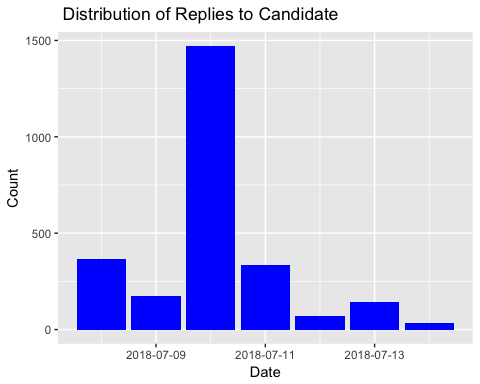
# How many are retweets?

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

##   
## FALSE TRUE   
## 7753 5871

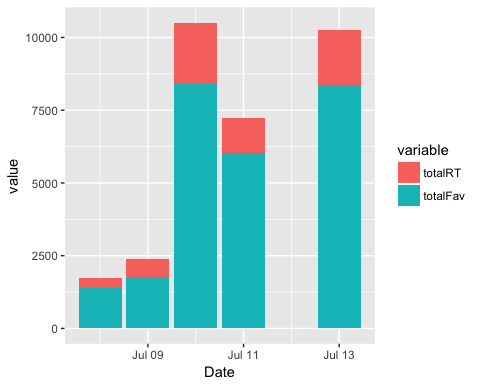
**ENGAGEMENT**

Engagement is a measurement of how people on the platform are interacting with the candidate's profile. Tracking engagement allows the campaign to assess how intensively are potential voters discussing or following the candidate, and identify which topics/tweets are "riling the base" by comparing popular tweets to ones that are less so. As a Measure of direct engagement, here is a time series graph created with the ggplot2 package that shows, per day, how often people tweeted direct replies to one of the Candidates tweets.



As another form of engagement, here is a breakdown of how many Favorable engagements(RT and Favs) did the candidate accumulate per day for every day that he/she tweeted.

# Fav/RT per Day



It helps to track which tweets are the most popular, both from the candidate's account and from third party accounts, as a way to understand which points/topics people are engaging the most favorably with. When printed, the tweets may be truncated. To see the full tweet, follow the (https://...) link at the end of the tweet.

Note: It might be helpful to look up these third-party accounts that sent out such popular tweets, as a way to identify potential allies that motivate other voters or share links for fundraising, or as a way to monitor detractors that may be spreading false/malicious points about the candidate.

### list of top 6 favorite tweets by the candidate

## text  
## 1 Here's what I'm wondering about the #SCOTUSpick:\n\nWill Judge Kavanaugh respect rulings to uphold the ACA? Will he s… https://t.co/Qy3LBXeJIH  
## 2 Projecting much? https://t.co/sRNsczSYQG  
## 3 More indictments prove, yet again, that hostile Russian actors targeted Americans with sophisticated crimes in 2016… https://t.co/KvF3qorBVs  
## 4 More indictments prove, yet again, that hostile Russian actors targeted Americans with sophisticated crimes in 2016… https://t.co/KvF3qorBVs  
## 5 A shocking example of this Administration weakening America’s global reputation.\n\nhttps://t.co/iMDYhRfXZ2  
## 6 This nominee also must be scrutinized for whether they will be a sufficiently independent Justice. Someone too loya… https://t.co/7SHLF1i4B2  
## favoriteCount created screenName  
## 1 8425 2018-07-10 01:31:10 timkaine  
## 2 6014 2018-07-11 13:56:00 timkaine  
## 3 4114 2018-07-13 16:35:00 timkaine  
## 4 4111 2018-07-13 16:35:00 timkaine  
## 5 1780 2018-07-09 12:14:18 timkaine  
## 6 580 2018-07-08 18:32:55 timkaine

## other than from candidate's account, most favorited

## text  
## 1 @watchingdaisies @Robyn\_Resists @SenateMajLdr @POTUS @SenateDems @timkaine @MarkWarner It's bad enough that they sa… https://t.co/I8vhMHiGxH  
## 2 @watchingdaisies @Astros290 @Robyn\_Resists @SenateMajLdr @POTUS @SenateDems @timkaine @MarkWarner The Senate has to… https://t.co/2OHIkbiB2W  
## 3 @watchingdaisies @DaniGraychik @Robyn\_Resists @SenateMajLdr @POTUS @SenateDems @timkaine @MarkWarner That's when we… https://t.co/XBvms0v5hB  
## 4 @watchingdaisies @DaniGraychik @Robyn\_Resists @SenateMajLdr @POTUS @SenateDems @timkaine @MarkWarner They keep sayi… https://t.co/u52reUH0k4  
## 5 We already know the answers to these questions, Tim Kaine. \nStop playing political games and help us #StopKavanaugh \nhttps://t.co/JCzzRyZH4u  
## 6 Yes @timkaine he probably did but it’s your name on the card, bc you are the idiot. Still pining over losing, chin… https://t.co/6fOTnQY11g  
## favoriteCount created screenName  
## 1 3995 2018-07-09 23:56:02 Astros290  
## 2 2362 2018-07-10 00:27:12 DaniGraychik  
## 3 2314 2018-07-10 00:55:28 Astros290  
## 4 1712 2018-07-10 00:57:16 Astros290  
## 5 1382 2018-07-10 02:39:06 brianefallon  
## 6 929 2018-07-08 00:01:25 PatrioticVa

# top 6 retweeted tweets

## text  
## 1 Here's what I'm wondering about the #SCOTUSpick:\n\nWill Judge Kavanaugh respect rulings to uphold the ACA? Will he s… https://t.co/Qy3LBXeJIH  
## 2 Projecting much? https://t.co/sRNsczSYQG  
## 3 More indictments prove, yet again, that hostile Russian actors targeted Americans with sophisticated crimes in 2016… https://t.co/KvF3qorBVs  
## 4 More indictments prove, yet again, that hostile Russian actors targeted Americans with sophisticated crimes in 2016… https://t.co/KvF3qorBVs  
## 5 A shocking example of this Administration weakening America’s global reputation.\n\nhttps://t.co/iMDYhRfXZ2  
## 6 This nominee also must be scrutinized for whether they will be a sufficiently independent Justice. Someone too loya… https://t.co/7SHLF1i4B2  
## retweetCount created screenName  
## 1 2086 2018-07-10 01:31:10 timkaine  
## 2 1203 2018-07-11 13:56:00 timkaine  
## 3 939 2018-07-13 16:35:00 timkaine  
## 4 938 2018-07-13 16:35:00 timkaine  
## 5 600 2018-07-09 12:14:18 timkaine  
## 6 130 2018-07-08 18:32:55 timkaine

# top 6 retweeted tweet Not including candidate

## text  
## 1 Yes @timkaine he probably did but it’s your name on the card, bc you are the idiot. Still pining over losing, chin… https://t.co/6fOTnQY11g  
## 2 We already know the answers to these questions, Tim Kaine. \nStop playing political games and help us #StopKavanaugh \nhttps://t.co/JCzzRyZH4u  
## 3 @watchingdaisies @Robyn\_Resists @SenateMajLdr @POTUS @SenateDems @timkaine @MarkWarner It's bad enough that they sa… https://t.co/I8vhMHiGxH  
## 4 Tim Kaine called for fighting in the streets, and former President Hussein Obama praised the pro-Isis pro-Syrian pr… https://t.co/01s2kbnx86  
## 5 @watchingdaisies @Astros290 @Robyn\_Resists @SenateMajLdr @POTUS @SenateDems @timkaine @MarkWarner The Senate has to… https://t.co/2OHIkbiB2W  
## 6 @watchingdaisies @DaniGraychik @Robyn\_Resists @SenateMajLdr @POTUS @SenateDems @timkaine @MarkWarner That's when we… https://t.co/XBvms0v5hB  
## retweetCount created screenName  
## 1 570 2018-07-08 00:01:25 PatrioticVa  
## 2 360 2018-07-10 02:39:06 brianefallon  
## 3 323 2018-07-09 23:56:02 Astros290  
## 4 309 2018-07-09 08:30:33 Combat\_Vet\_  
## 5 164 2018-07-10 00:27:12 DaniGraychik  
## 6 129 2018-07-10 00:55:28 Astros290

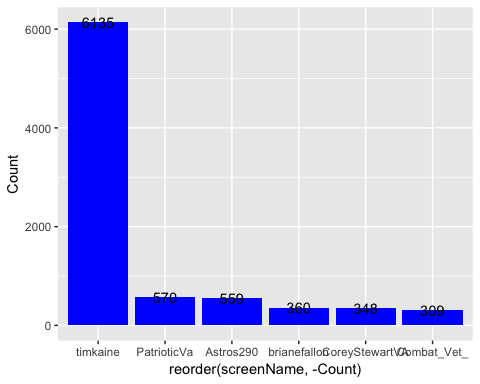
### Top ten retweets overall

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

## text  
## 1 RT @timkaine: The real Trump Hotel. https://t.co/PP6nIbzNQR  
## 2 RT @realDonaldTrump: Congratulations to Corey Stewart for his great victory for Senator from Virginia. Now he runs against a total stiff, T…  
## 3 RT @ScottPresler: Senator @timkaine (D-VA) said democrats should fight Republicans in the streets.\n\nDid you get what you wanted?\n#SteveScal…  
## 4 RT @Camila\_Cabello: i am at a loss for words ...... https://t.co/ginoXLsEqS  
## 5 RT @CoreyStewartVA: I am running for U.S. Senate to defeat @timkaine - Join the movement. https://t.co/HXPdTRFsHM #DefeatKaine  
## 6 RT @nowthisnews: Liviu Librescu survived the Holocaust as a young boy. \n\nDecades later, he lost his life to senseless gun violence as a pro…  
## 7 RT @DisavowTrump20: BREAKING: Far-right Trump supporter and champion of the confederate flag Corey Stewart has won the GOP nomination for t…  
## 8 Here's what I'm wondering about the #SCOTUSpick:\n\nWill Judge Kavanaugh respect rulings to uphold the ACA? Will he s… https://t.co/Qy3LBXeJIH  
## 9 RT @funder: BREAKING: Senator Tim Kaine’s opponent Corey Stewart is now under investigation @TheDemCoalition. Help fund our research team b…  
## 10 RT @BillClinton: Day 2 of @HillaryClinton and @TimKaine's bus tour. Here we go! https://t.co/e6I5Ye2oGD  
## retweetCount created screenName isRetweet  
## 1 41140 2018-07-13 12:34:36 Lika\_Way\_ TRUE  
## 2 16304 2018-07-13 03:23:28 iaalrogmom TRUE  
## 3 10505 2018-07-09 22:24:38 Rosemarie910 TRUE  
## 4 8147 2018-07-13 02:46:20 lmjazyst TRUE  
## 5 6221 2018-07-14 03:31:30 Jamieph41382 TRUE  
## 6 4002 2018-07-14 00:59:45 rayntpa TRUE  
## 7 3812 2018-07-14 05:18:13 AntibanterKraky TRUE  
## 8 2086 2018-07-10 01:31:10 timkaine FALSE  
## 9 1299 2018-07-10 01:35:53 JimmyV08070562 TRUE  
## 10 1211 2018-07-09 02:32:51 sarahtha TRUE

## accounts that have the most total Retweets

## screenName x  
## 5104 timkaine 6135  
## 3854 PatrioticVa 570  
## 372 Astros290 559  
## 687 brianefallon 360  
## 1064 CoreyStewartVA 348  
## 1032 Combat\_Vet\_ 309



## Most Active Accounts (including Retweets)

## Var1 Freq  
## 6520 openletterbot 123  
## 2548 Dufenough 30  
## 9027 tryan5786 29  
## 572 aquiroga72 27  
## 9332 watchingdaisies 26  
## 1372 CarlosV64910815 25

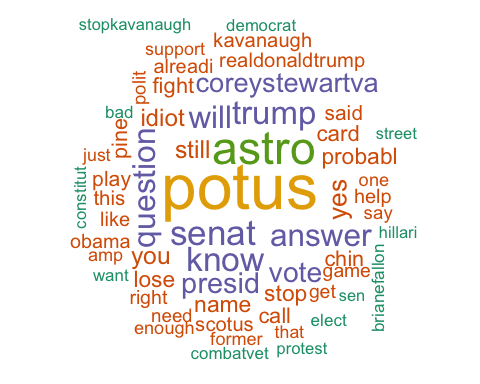
## Most Active Accounts

## Var1 Freq  
## 3778 openletterbot 123  
## 1499 Dufenough 30  
## 5236 tryan5786 29  
## 331 aquiroga72 27  
## 5407 watchingdaisies 26  
## 3363 MikeRitt102 24

**TEXT MINING**

What are people saying? The goal is to capture what words are being used to describe the candidate. We expect this section to be the one of the most interesting, as the team's NLP capabilities get better we expect this section to expand.

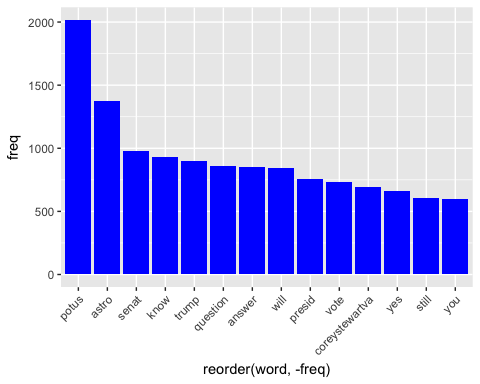
# WORDCLOUD of most common used words (Top 500)



# List of The Top 15 Frequently used Words

## word freq  
## potus potus 2016  
## astro astro 1378  
## senat senat 979  
## know know 928  
## trump trump 897  
## question question 860  
## answer answer 849  
## will will 844  
## presid presid 753  
## vote vote 732  
## coreystewartva coreystewartva 693  
## yes yes 662  
## still still 607  
## you you 602  
## idiot idiot 585

# Plot of Words that Appear at least 1500 times



## Words Association

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 "tim", "kaine", or "timkaine", "corey","stewart" or "coreystewartva"

## $tim  
## kain card patrioticva chin pine   
## 0.76 0.42 0.41 0.41 0.41   
## lose idiot name probabl still   
## 0.40 0.39 0.39 0.39 0.37   
## former yes combatvet street prais   
## 0.34 0.33 0.32 0.31 0.31   
## brianefallon stopkavanaugh hussein proisi prosyrian   
## 0.31 0.31 0.31 0.31 0.31   
## protest polit call help obama   
## 0.30 0.28 0.27 0.25 0.25   
## alreadi game fight play stop   
## 0.25 0.25 0.22 0.22 0.21   
## dva son presid hillari sen   
## 0.19 0.17 0.17 0.16 0.15   
## know hurl stewart via clinton   
## 0.15 0.15 0.14 0.14 0.14   
## abolish arrest sawmilltat smoke elton   
## 0.14 0.14 0.14 0.14 0.14   
## answer question aid trash jong   
## 0.13 0.13 0.13 0.13 0.13   
## kim gift bomb bigleaguepol ice   
## 0.13 0.13 0.13 0.12 0.12   
## capitol dailycal over march minnesota   
## 0.12 0.12 0.12 0.12 0.12   
## helenh john korean north rocket   
## 0.12 0.12 0.12 0.12 0.12   
## man bash copi respond too   
## 0.11 0.11 0.11 0.11 0.11   
## safetypindaili dutyofapatriot corey attend funni   
## 0.10 0.10 0.10 0.10 0.10

## $kain  
## tim former combatvet street prais   
## 0.76 0.39 0.37 0.35 0.35   
## brianefallon stopkavanaugh hussein proisi prosyrian   
## 0.35 0.35 0.35 0.35 0.35   
## protest call polit help obama   
## 0.34 0.32 0.31 0.29 0.29   
## game alreadi play fight sen   
## 0.29 0.28 0.26 0.25 0.24   
## stop dva presid son hillari   
## 0.24 0.22 0.20 0.19 0.19   
## know jong kim hurl stewart   
## 0.18 0.18 0.17 0.17 0.16   
## via clinton abolish ice copi   
## 0.16 0.16 0.16 0.16 0.16   
## sawmilltat smoke elton answer question   
## 0.16 0.16 0.16 0.15 0.15   
## arrest aid trash bomb bigleaguepol   
## 0.15 0.15 0.15 0.15 0.14   
## capitol over gift march helenh   
## 0.14 0.14 0.14 0.14 0.14   
## john korean rocket corey respond   
## 0.14 0.14 0.14 0.13 0.13   
## dailycal minnesota north safetypindaili dutyofapatriot   
## 0.13 0.13 0.13 0.12 0.12   
## man attend bash funni too   
## 0.12 0.12 0.12 0.12 0.12   
## terrorist sarahpeckva norml vanorml nomine   
## 0.11 0.11 0.11 0.11 0.11   
## longtim mikala less race exam   
## 0.11 0.11 0.10 0.10 0.10   
## vas feder garrett taylor   
## 0.10 0.10 0.10 0.10

## $timkain  
## markwarn senatemajldr potus senatedem robynresist   
## 0.30 0.26 0.25 0.25 0.25   
## watchingdaisi astro patrioticva danigraychik dutyofapatriot   
## 0.24 0.20 0.17 0.15 0.14   
## terrorist lose chin pine less   
## 0.13 0.13 0.13 0.13 0.12   
## yes card still name probabl   
## 0.12 0.12 0.12 0.12 0.12   
## expect idiot enough   
## 0.10 0.10 0.10

## $corey  
## stewart bash   
## 0.79 0.50   
## attend abolish   
## 0.46 0.41   
## bigleaguepol ice   
## 0.38 0.38   
## ralli race   
## 0.28 0.25   
## congressman amayalindsey   
## 0.25 0.22   
## pant antif   
## 0.21 0.19   
## chrismo maganinajo   
## 0.19 0.19   
## rate crackdown   
## 0.19 0.19   
## parscal zenonichol   
## 0.18 0.17   
## countycounti map   
## 0.17 0.17   
## antitrumpmvmt quot   
## 0.17 0.17   
## massd stiff   
## 0.17 0.15   
## chaznuttycomb scotttaylorva   
## 0.15 0.15   
## kain scottpresl   
## 0.13 0.13   
## politifact champion   
## 0.13 0.12   
## disavowtrump farright   
## 0.12 0.12   
## support deplorablemo   
## 0.12 0.12   
## iwillredpillu trumpback   
## 0.12 0.12   
## result poll   
## 0.12 0.12   
## agress ssssshhhhhhhhhoooooooooshhhhhhhhh   
## 0.12 0.12   
## stupid rid   
## 0.11 0.11   
## unequivoc tim   
## 0.11 0.10   
## updat   
## 0.10

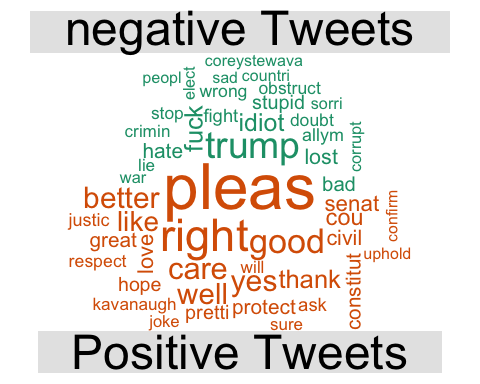
## $stewart  
## corey bash attend abolish bigleaguepol   
## 0.79 0.51 0.47 0.40 0.38   
## ice wannab redstepforward unseat cori   
## 0.37 0.35 0.34 0.34 0.33   
## killari ralli chosen race congressman   
## 0.33 0.28 0.26 0.25 0.23   
## amayalindsey pant rate must antif   
## 0.22 0.21 0.20 0.19 0.19   
## chrismo maganinajo crackdown parscal zenonichol   
## 0.19 0.19 0.19 0.18 0.17   
## countycounti map irepdcsport scotttaylorva antitrumpmvmt   
## 0.17 0.17 0.17 0.17 0.17   
## massd kain quot stiff chaznuttycomb   
## 0.17 0.16 0.16 0.15 0.15   
## tim hell scottpresl poll politifact   
## 0.14 0.13 0.13 0.13 0.13   
## champion disavowtrump farright clinton deplorablemo   
## 0.12 0.12 0.12 0.12 0.12   
## neonazi iwillredpillu trumpback result agress   
## 0.12 0.12 0.12 0.12 0.12   
## repudi support rid updat unequivoc   
## 0.12 0.11 0.11 0.11 0.11   
## stupid congratul   
## 0.10 0.10

## $coreystewartva  
## economi growth brought gtecon   
## 0.38 0.37 0.36 0.36   
## gtunemploy lowest bold restor   
## 0.36 0.36 0.35 0.35   
## govt took tonight toward   
## 0.34 0.33 0.32 0.32   
## select step realdonaldtrump constitut   
## 0.32 0.32 0.31 0.30   
## perceiv slightest tantrum throw   
## 0.28 0.28 0.28 0.28   
## violentleft agend leftw nati   
## 0.28 0.28 0.28 0.27   
## offens radic back advanc   
## 0.27 0.27 0.26 0.26   
## defend defeatkain touch brett   
## 0.26 0.26 0.26 0.25   
## movement defeat america anoth   
## 0.25 0.24 0.23 0.21   
## commonwealth run justinhendrix brerva   
## 0.20 0.18 0.18 0.18   
## good johnreason denverva racist   
## 0.16 0.16 0.16 0.15   
## american esquir robertcd join   
## 0.14 0.14 0.14 0.13   
## kavanaugh indivisibleva want queenboo   
## 0.13 0.13 0.12 0.12   
## scotttaylorva reptomgarrett joshuarosen karenshavoc   
## 0.12 0.12 0.12 0.11   
## shall vabeachdem jimbeaumtb   
## 0.10 0.10 0.10

**Sentiment Analysis**

Using the (SentimentR) package, every tweet is scored to assess whether they are negative, positive or neutral. The algorithm estimates how negative or positive a tweet is based on the words and punctuation.

# WorldCloud of words used in top 1000 most negative and top 1000 most positive tweets



**Note: The WordCloud indicates which words were most common amongst the positive and negative tweets, it’s not stating if these words themselves are positive or negative.**

