Tim Kaine Mentions Analysis

Georges Colbert

07/10/2018

# Period of June 30 to July 07

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 June 30-July 07. All the tweets returned by the API are in the zip folder

**OVERVIEW**

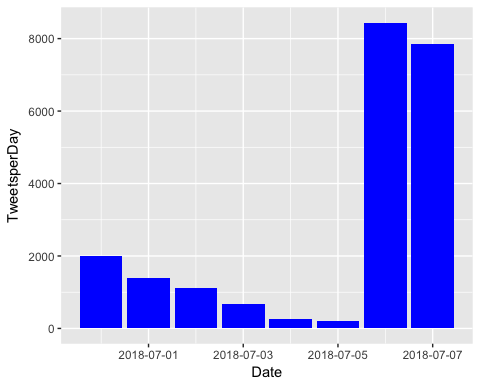
# Total Number of tweets

## [1] 21925

# Time-Series of Tweets Per Day

##   
## Attaching package: 'scales'

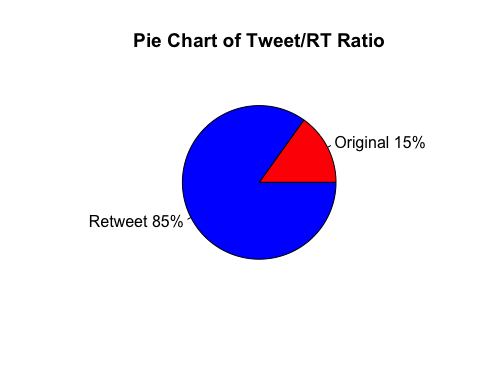
## The following object is masked from 'package:readr':  
##   
## col\_factor



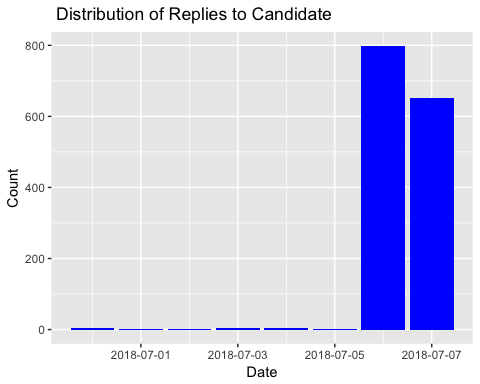
# How many are retweets?

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

##   
## FALSE TRUE   
## 3314 18611

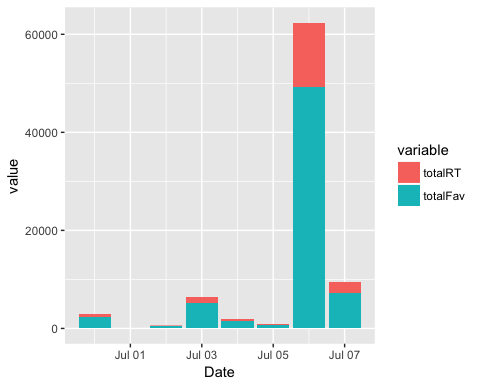
**ENGAGEMENT**

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://](NULL)...) 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 And Kim Jong Un sends back a copy of Green Day’s “American Idiot.” https://t.co/YCGjR2QcgU  
## 2 When you mock John McCain...\nAnd George H.W. Bush...\nAnd Senator Tester...\nAnd Senator Warren...\nAnd the #MeToo mov… https://t.co/iy7wS6Aht8  
## 3 Hey EPA director—how about protecting the environment rather than your bank account?\n\nhttps://t.co/dRj6YcoSYt  
## 4 Good piece about Trump's stupid trade war. His steel &amp; aluminum tariffs will likely add 30k jobs in certain industr… https://t.co/rz3YG1SY7F  
## 5 Today I rallied with hundreds of Virginians in Richmond to voice a simple message: #FamiliesBelongTogether.\n\nThe Tr… https://t.co/qfEzTojVZV  
## 6 On Independence Day, I can’t help but think about the founding values of our country - the ones we pass along to ou… https://t.co/HsDKC0reKO  
## favoriteCount created screenName  
## 1 45738 2018-07-06 15:27:49 timkaine  
## 2 7162 2018-07-07 18:20:51 timkaine  
## 3 4182 2018-07-03 14:07:05 timkaine  
## 4 2525 2018-07-06 13:34:57 timkaine  
## 5 2399 2018-06-30 19:15:09 timkaine  
## 6 1373 2018-07-04 18:06:31 timkaine

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

## text  
## 1 WHERE WAS THIS TIM KAINE IN 2016? https://t.co/u8BaXZCkpL  
## 2 Rather than support racist, Trump-loving pro-Confederate candidate Corey Stewart, who promises to run a 'vicious ca… https://t.co/u2TWIeoSpI  
## 3 Tim Kaine should rule out supporting Barrett in advance of Monday. Take the issue off the table now so WH cant try… https://t.co/WUSBt29BhO  
## 4 \xed\xa0\xbd\xed\xb1\x8f\xed\xa0\xbd\xed\xb1\x8fPro-Trump Senate Candidate for Virginia, Corey Stewart Holds Rally at Red Hen Restaurant. \xed\xa0\xbd\xed\xb1\x8a\xed\xa0\xbd\xed\xb1\x8a\n\nCorey Stewart is ru… https://t.co/k14zb1G8qK  
## 5 Tim Kaine jokes that Kim Jong Un will send Trump copy of Green Day's "American Idiot" https://t.co/hsHYVk9UvI https://t.co/dXB2zHd4xj  
## 6 Yooooo I was at an immigration rally and I went up to the mic to point out that Dems deport people too, said Tim Ka… https://t.co/tfXwmrmtDO  
## favoriteCount created screenName  
## 1 1161 2018-07-06 20:34:14 samswey  
## 2 878 2018-06-30 23:38:59 activist360  
## 3 727 2018-07-03 04:00:00 brianefallon  
## 4 717 2018-07-02 00:50:06 cs0058sc  
## 5 607 2018-07-06 16:55:08 thehill  
## 6 523 2018-07-01 18:26:49 amal\_monstah

# top 6 retweeted tweets

## text  
## 1 And Kim Jong Un sends back a copy of Green Day’s “American Idiot.” https://t.co/YCGjR2QcgU  
## 2 When you mock John McCain...\nAnd George H.W. Bush...\nAnd Senator Tester...\nAnd Senator Warren...\nAnd the #MeToo mov… https://t.co/iy7wS6Aht8  
## 3 Hey EPA director—how about protecting the environment rather than your bank account?\n\nhttps://t.co/dRj6YcoSYt  
## 4 Good piece about Trump's stupid trade war. His steel &amp; aluminum tariffs will likely add 30k jobs in certain industr… https://t.co/rz3YG1SY7F  
## 5 Today I rallied with hundreds of Virginians in Richmond to voice a simple message: #FamiliesBelongTogether.\n\nThe Tr… https://t.co/qfEzTojVZV  
## 6 On Independence Day, I can’t help but think about the founding values of our country - the ones we pass along to ou… https://t.co/HsDKC0reKO  
## retweetCount created screenName  
## 1 11897 2018-07-06 15:27:49 timkaine  
## 2 2283 2018-07-07 18:20:51 timkaine  
## 3 1152 2018-07-03 14:07:05 timkaine  
## 4 1066 2018-07-06 13:34:57 timkaine  
## 5 507 2018-06-30 19:15:09 timkaine  
## 6 233 2018-07-04 18:06:31 timkaine

# top 6 retweeted tweet Not including candidate

## text  
## 1 Rather than support racist, Trump-loving pro-Confederate candidate Corey Stewart, who promises to run a 'vicious ca… https://t.co/u2TWIeoSpI  
## 2 \xed\xa0\xbd\xed\xb1\x8f\xed\xa0\xbd\xed\xb1\x8fPro-Trump Senate Candidate for Virginia, Corey Stewart Holds Rally at Red Hen Restaurant. \xed\xa0\xbd\xed\xb1\x8a\xed\xa0\xbd\xed\xb1\x8a\n\nCorey Stewart is ru… https://t.co/k14zb1G8qK  
## 3 Tim Kaine should rule out supporting Barrett in advance of Monday. Take the issue off the table now so WH cant try… https://t.co/WUSBt29BhO  
## 4 I have NEVER seen this photo before…….Marc Mezvinsky, Bill Clinton, and Tim Kaine on election night!! THEIR ANGUISH… https://t.co/6uPpX55TWX  
## 5 Democrats who voted to confirm Amy Coney Barrett in 2017 to be U.S. Circuit Judge for the Seventh Circuit:\n\nJoe Don… https://t.co/q0W4lM1PVu  
## 6 Tim Kaine Call Trump an “Idiot” After Pompeo Gives Kim Jong Un An Elton John CD. Forgets his loser son was arrested… https://t.co/cY2ZCbCyBe  
## retweetCount created screenName  
## 1 548 2018-06-30 23:38:59 activist360  
## 2 517 2018-07-02 00:50:06 cs0058sc  
## 3 247 2018-07-03 04:00:00 brianefallon  
## 4 247 2018-06-30 03:39:21 Bud\_Doggin  
## 5 213 2018-07-02 22:51:57 RealSaavedra  
## 6 194 2018-07-06 20:50:15 trustrestored

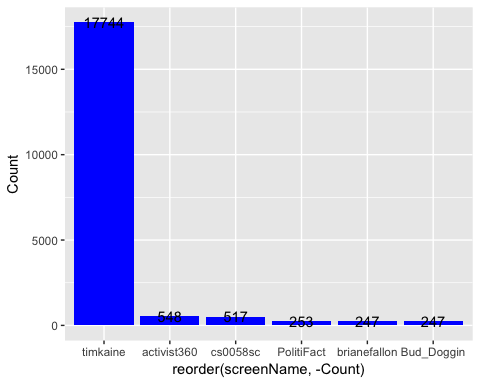
### Top ten retweets overall

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 @timkaine: I wanted to be first at my polling place, but 99-year-old Minerva Turpin beat me to it. Looks like I need to get used to bein…  
## 3 RT @timkaine: People often forget that we \*already had\* an assault weapons ban. For a decade! So don't let the gun lobby cry "unconstitutio…  
## 4 RT @realDonaldTrump: Congratulations to Corey Stewart for his great victory for Senator from Virginia. Now he runs against a total stiff, T…  
## 5 And Kim Jong Un sends back a copy of Green Day’s “American Idiot.” https://t.co/YCGjR2QcgU  
## 6 RT @timkaine: And Kim Jong Un sends back a copy of Green Day’s “American Idiot.” https://t.co/YCGjR2QcgU  
## 7 RT @CoreyStewartVA: I am running for U.S. Senate to defeat @timkaine - Join the movement. https://t.co/HXPdTRFsHM #DefeatKaine  
## 8 RT @timkaine: The Constitution is clear: With SCOTUS, both the President and Senate have roles in this process. \n\nThe GOP was clear: In an…  
## 9 RT @timkaine: The Old Testament instructs the faithful: “When a foreigner resides among you in your land, do not mistreat them... Love them…  
## 10 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…  
## retweetCount created screenName isRetweet  
## 1 41182 2018-07-07 18:48:17 DCBill5 TRUE  
## 2 25295 2018-07-07 16:01:13 mary\_swilling TRUE  
## 3 19481 2018-07-07 04:38:44 organikgrl TRUE  
## 4 16332 2018-07-06 21:56:55 tloerjtoker57 TRUE  
## 5 11897 2018-07-06 15:27:49 timkaine FALSE  
## 6 11893 2018-07-07 23:59:34 jimspeese TRUE  
## 7 6185 2018-07-07 00:04:05 MotherVoorhees TRUE  
## 8 5349 2018-07-07 01:56:54 sdhetman1 TRUE  
## 9 4750 2018-07-07 14:15:49 mirthandmatter1 TRUE  
## 10 4000 2018-06-30 18:55:15 Arn68353558Arn TRUE

## accounts that have the most total Retweets

## screenName x  
## 2469 timkaine 17744  
## 55 activist360 548  
## 566 cs0058sc 517  
## 1935 PolitiFact 253  
## 341 brianefallon 247  
## 372 Bud\_Doggin 247



## Most Active Accounts (including Retweets)

## Var1 Freq  
## 13286 Oooooo\_Donna 35  
## 18348 USA\_First\_2016 29  
## 7099 HillaryisGod 24  
## 2807 CarlosV64910815 18  
## 15940 ShawnaOhm 18  
## 15345 SackComedian 17

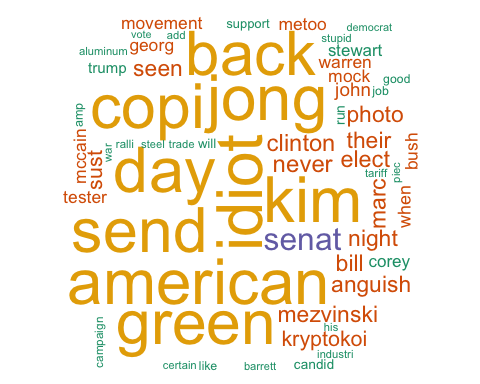
## Most Active Accounts

## Var1 Freq  
## 1828 Oooooo\_Donna 32  
## 2551 USA\_First\_2016 29  
## 1006 HillaryisGod 24  
## 2212 ShawnaOhm 18  
## 1089 jamesmpope 14  
## 864 FuzzeePundit 13

**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.

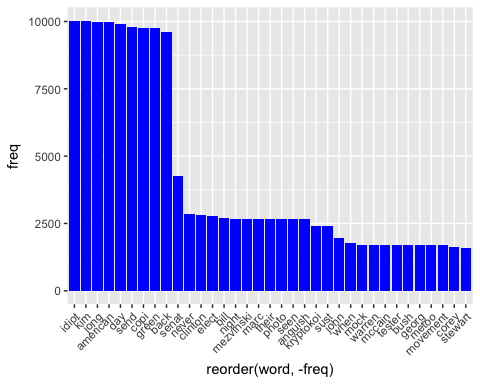
# WORDCLOUD of most common used words (Top 500)



# List of The Top 15 Frequently used Words

## word freq  
## idiot idiot 10026  
## kim kim 10008  
## jong jong 9984  
## american american 9982  
## day day 9905  
## send send 9797  
## copi copi 9762  
## green green 9755  
## back back 9623  
## senat senat 4251  
## never never 2858  
## clinton clinton 2821  
## elect elect 2759  
## bill bill 2703  
## night night 2679

# 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 clinton anguish bill elect   
## 0.96 0.67 0.65 0.65 0.65   
## marc mezvinski night photo seen   
## 0.65 0.65 0.65 0.65 0.65   
## their kryptokoi never sust border   
## 0.65 0.62 0.62 0.61 0.22   
## ice sen abolish politifact want   
## 0.22 0.20 0.18 0.18 0.18   
## after rule sustain buddoggin advanc   
## 0.18 0.18 0.18 0.18 0.18   
## brianefallon jam monday tabl call   
## 0.18 0.18 0.18 0.18 0.17   
## elton said cant issu give   
## 0.17 0.17 0.17 0.17 0.16   
## pompeo trustrestor take tri arrest   
## 0.16 0.16 0.16 0.16 0.15   
## son forget pant republican rid   
## 0.15 0.15 0.14 0.14 0.14   
## loser amalmonstah thehil went now   
## 0.14 0.14 0.13 0.13 0.13   
## immigr deport yooooo patrol point   
## 0.13 0.13 0.13 0.12 0.12   
## mic claim doe get rate   
## 0.12 0.11 0.11 0.11 0.11   
## hillari attende samswey where barrett   
## 0.11 0.11 0.10 0.10 0.10   
## harmonica chelsea matthewstol outreach strategi   
## 0.10 0.10 0.10 0.10 0.10

## $kain  
## tim clinton anguish bill elect   
## 0.96 0.67 0.65 0.65 0.65   
## marc mezvinski night photo seen   
## 0.65 0.65 0.65 0.65 0.65   
## their kryptokoi never sust border   
## 0.65 0.62 0.62 0.61 0.22   
## ice sen abolish politifact want   
## 0.22 0.20 0.19 0.18 0.18   
## after sustain buddoggin advanc brianefallon   
## 0.18 0.18 0.18 0.18 0.18   
## jam monday tabl call elton   
## 0.18 0.18 0.18 0.17 0.17   
## said cant issu give pompeo   
## 0.17 0.17 0.17 0.16 0.16   
## trustrestor rule take tri arrest   
## 0.16 0.16 0.16 0.16 0.15   
## son forget pant republican rid   
## 0.15 0.15 0.14 0.14 0.14   
## loser amalmonstah thehil now immigr   
## 0.14 0.14 0.13 0.13 0.13   
## deport yooooo joke went patrol   
## 0.13 0.13 0.12 0.12 0.12   
## point mic claim doe get   
## 0.12 0.12 0.11 0.11 0.11   
## rate attende samswey where harmonica   
## 0.11 0.11 0.11 0.10 0.10   
## chelsea outreach strategi   
## 0.10 0.10 0.10

## $timkain  
## back copi send day green idiot american jong   
## 0.61 0.59 0.59 0.59 0.59 0.56 0.56 0.55   
## kim and mock tester movement warren bush georg   
## 0.55 0.48 0.20 0.20 0.20 0.20 0.20 0.20   
## mccain metoo when senat john stupid war piec   
## 0.20 0.20 0.18 0.18 0.13 0.11 0.11 0.11   
## his trade add certain aluminum industri steel tariff   
## 0.11 0.11 0.11 0.11 0.11 0.11 0.11 0.11

## $corey  
## stewart cssc protrump restaur hen candid   
## 0.98 0.80 0.80 0.80 0.79 0.78   
## run hold red for ralli racist   
## 0.78 0.78 0.75 0.74 0.61 0.41   
## promis vicious activist proconfeder rather trumplov   
## 0.41 0.41 0.41 0.41 0.41 0.41   
## campaign politifact support republican want sen   
## 0.34 0.29 0.27 0.26 0.26 0.25   
## border ice pant rid claim rate   
## 0.24 0.24 0.22 0.22 0.20 0.18   
## patrol abolish doe get away broadcast   
## 0.18 0.17 0.17 0.17 0.14 0.14   
## facebook dure pat said fire mincemeat   
## 0.14 0.13 0.13 0.12 0.12 0.12   
## stew video despit watchingfox   
## 0.11 0.11 0.11 0.10

## $stewart  
## corey cssc protrump restaur candid run   
## 0.98 0.81 0.81 0.81 0.80 0.80   
## hen hold for red ralli racist   
## 0.80 0.79 0.76 0.76 0.62 0.42   
## promis vicious activist proconfeder trumplov rather   
## 0.42 0.42 0.42 0.42 0.42 0.41   
## campaign support politifact pant republican rid   
## 0.34 0.28 0.24 0.23 0.22 0.22   
## want sen border ice claim rate   
## 0.22 0.21 0.20 0.20 0.19 0.18   
## abolish doe get dure broadcast facebook   
## 0.16 0.16 0.16 0.14 0.14 0.14   
## pat video fire patrol   
## 0.14 0.12 0.12 0.11

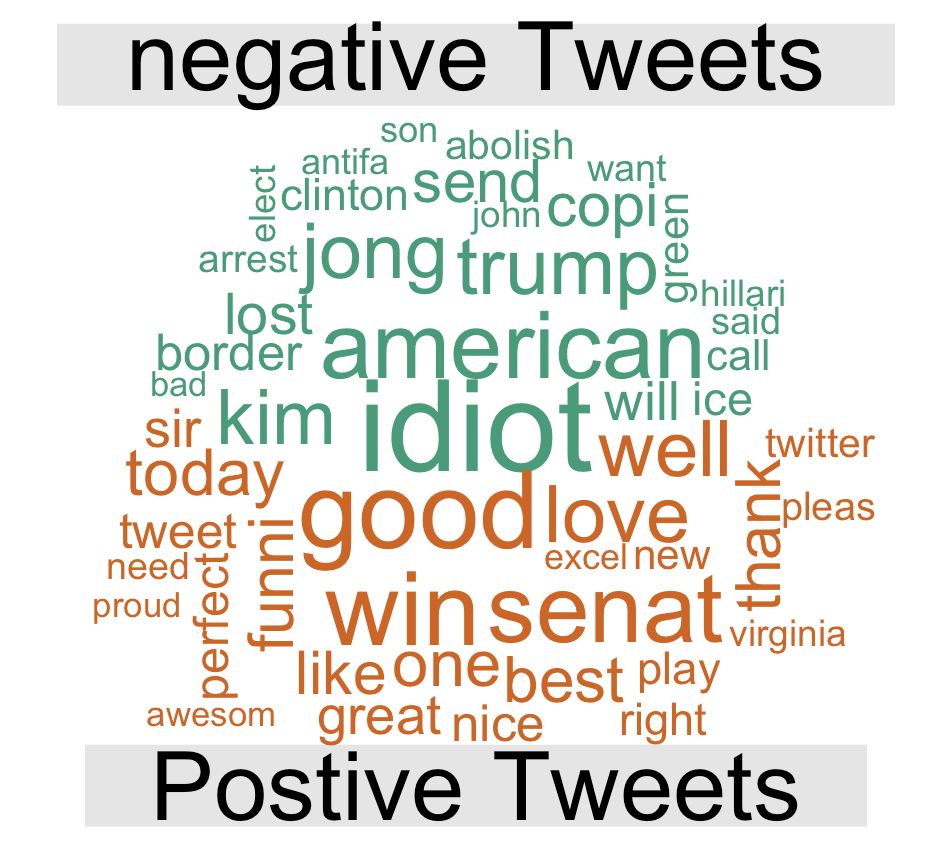
## $coreystewartva  
## experienc economi someth there   
## 0.74 0.73 0.69 0.69   
## thank peopl vabeachdem scotttaylorva   
## 0.57 0.45 0.44 0.43   
## work campaigni imag pictur   
## 0.39 0.36 0.36 0.34   
## highlight job word clear   
## 0.33 0.31 0.31 0.24   
## worth elaineluriava never fightingveteran   
## 0.23 0.18 0.14 0.14   
## vagop mincemeat islam tie   
## 0.13 0.12 0.12 0.12   
## abolishic proamnesti watchingfox pacifist   
## 0.11 0.11 0.11 0.11   
## wimpi everywheregari dis   
## 0.11 0.11 0.11

**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 is not stating if these words themselves are positive or negative.**

