

Introduction to Twitter Data in R

Aditya Ranganath

12/6/2021

R Markdown

```
library(rtweet)
library(httputil)
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.3      v purrr  0.3.4
## v tibble  3.0.6      v dplyr  1.0.4
## v tidyr   1.1.2      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x purrr::flatten() masks rtweet::flatten()
## x dplyr::lag()     masks stats::lag()
```

```
library(tidytext)
library(wordcloud2)
library(qdapRegex)
```

```
##
## Attaching package: 'qdapRegex'
```

```
## The following object is masked from 'package:dplyr':
##
##     explain
```

```
## The following object is masked from 'package:ggplot2':
##
##     %+%
```

```
library(tm)
```

```
## Loading required package: NLP
```

```
##  
## Attaching package: 'NLP'
```

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

```
student_debt_tweets<-search_tweets(q="#CancelStudentDebt",  
                                   n=1000,  
                                   include_rts=FALSE,  
                                   `~filter`="replies",  
                                   lang="en")
```

```
twitter_wordcloud<-function(twitterhandle, tweet_number){  
  tweet_timeline<-get_timeline(twitterhandle, n=tweet_number)  
  tweet_timeline_text<-str_c(tweet_timeline$text, collapse="")  
  tweet_timeline_text<-str_remove_all(tweet_timeline_text, pattern='[:emoji:]')  
  tweet_timeline_text<-tweet_timeline_text %>%  
    str_remove("\\n") %>%                # remove linebreaks  
    rm_twitter_url() %>%                 # Remove URLs  
    rm_url() %>%  
    str_remove_all("#\\S+") %>%          # Remove any hashtags  
    str_remove_all("@\\S+") %>%         # Remove any @ mentions  
    removeWords(stopwords("english")) %>% # Remove common words (a, the, it  
    removeNumbers() %>%  
    stripWhitespace() %>%  
    removeWords(c("amp")) %>%  
    removePunctuation()  
  
  textCorpus <-  
    Corpus(VectorSource(tweet_timeline_text)) %>%  
    TermDocumentMatrix() %>%  
    as.matrix()  
  
  textCorpus <- sort(rowSums(textCorpus), decreasing=TRUE)  
  textCorpus <- data.frame(word = names(textCorpus), freq=textCorpus, row.names = NULL)  
  
  wordcloud <- wordcloud2(data = textCorpus, minRotation = 0, maxRotation = 0, ellipticity = 0.2)  
  return(wordcloud)  
}
```

```
lebron_wordcloud<-twitter_wordcloud("KingJames", 400)  
elon_wordcloud<-twitter_wordcloud("elonmusk", 400)
```

```
elon_wordcloud
```

just always
will can
launch

next

dddddddsfadfa dafafadfafa

```
lebron_wordcloud<-twitter_wordcloud("KingJames", 400)  
lebron_wordcloud
```

love
now
the man

 <https://github.com/Lchiffon/wordcloud2/issues/65>