Climate

Krishangi

6/16/2020

Word Cloud: Climate and COVID-19.

```
## install devtools package if it's not already
if (!requireNamespace("devtools", quietly = TRUE)) {
  install.packages("devtools") }
```

```
library(rtweet)
```

```
## Warning: package 'rtweet' was built under R version 3.6.3
```

```
create_token(
  app = "TwitterData_Analysis2020",
  consumer_key = 'Vto757F3nGIAcj1BLCmHEsRvb',
  consumer_secret =  'jbGQQXiSXHlpD6XDJwyIfzu2O0jijODLFhqydpGElRztM5qP1p',
  access_token = '1272250817064710147-0Hy0yNcj9nqiAsp6U23Rl3MBOlsdvX',
  access_secret = 'LYPYAUvDvhb1idOr1CdAOYOSY3AByNAlwEjqZdGlecUVa')
```

```
## <Token>
## <oauth_endpoint>
##  request:   https://api.twitter.com/oauth/request_token
##  authorize: https://api.twitter.com/oauth/authenticate
##  access:    https://api.twitter.com/oauth/access_token
## <oauth_app> TwitterData_Analysis2020
##   key:    Vto757F3nGIAcj1BLCmHEsRvb
##   secret: <hidden>
## <credentials> oauth_token, oauth_token_secret
## ---
```

```
library(wordcloud)
```

```
## Warning: package 'wordcloud' was built under R version 3.6.3
```

```
## Loading required package: RColorBrewer
```

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 3.6.3
```

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

## v ggplot2 3.3.0     v purrr   0.3.3
## v tibble  3.0.0     v dplyr   0.8.5
## v tidyr   1.0.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0

## Warning: package 'ggplot2' was built under R version 3.6.3

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

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

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

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

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

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

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

```
library(tidyr)
library(tidytext)
```

```
## Warning: package 'tidytext' was built under R version 3.6.3
```

```
library(dplyr)
library(ggplot2)
library(RColorBrewer)
library(tm)
```

```
## Warning: package 'tm' was built under R version 3.6.3

## Loading required package: NLP

##
## Attaching package: 'NLP'

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

```
stream_tweets(
q = "climate, global warming, environment, carbondioxide, nitrogen, dioxide, greenhouse gases, anomaly,
timeout = 60,
parse = FALSE,
file_name = "tweets_climate.json")
```

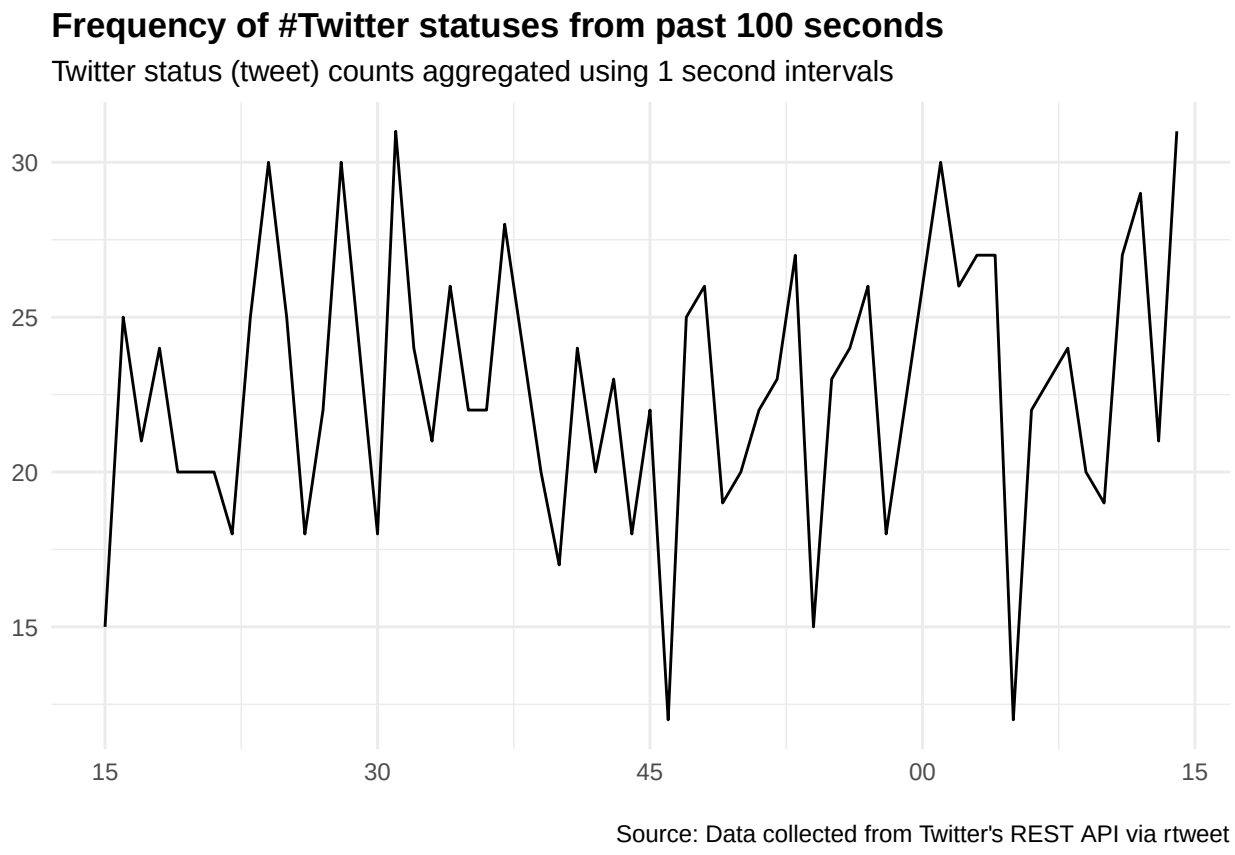
```
## Streaming tweets for 60 seconds...

## Finished streaming tweets!

## streaming data saved as tweets_climate.json
```

```
tweets_climate <- parse_stream("tweets_climate.json")
```

```
tweets_climate %>%
  ts_plot("1 second") +
  ggplot2::theme_minimal() +
  ggplot2::theme(plot.title = ggplot2::element_text(face = "bold")) +
  ggplot2::labs(
    x = NULL, y = NULL,
    title = "Frequency of #Twitter statuses from past 100 seconds",
    subtitle = "Twitter status (tweet) counts aggregated using 1 second intervals",
    caption = "\nSource: Data collected from Twitter's REST API via rtweet"
  )
```

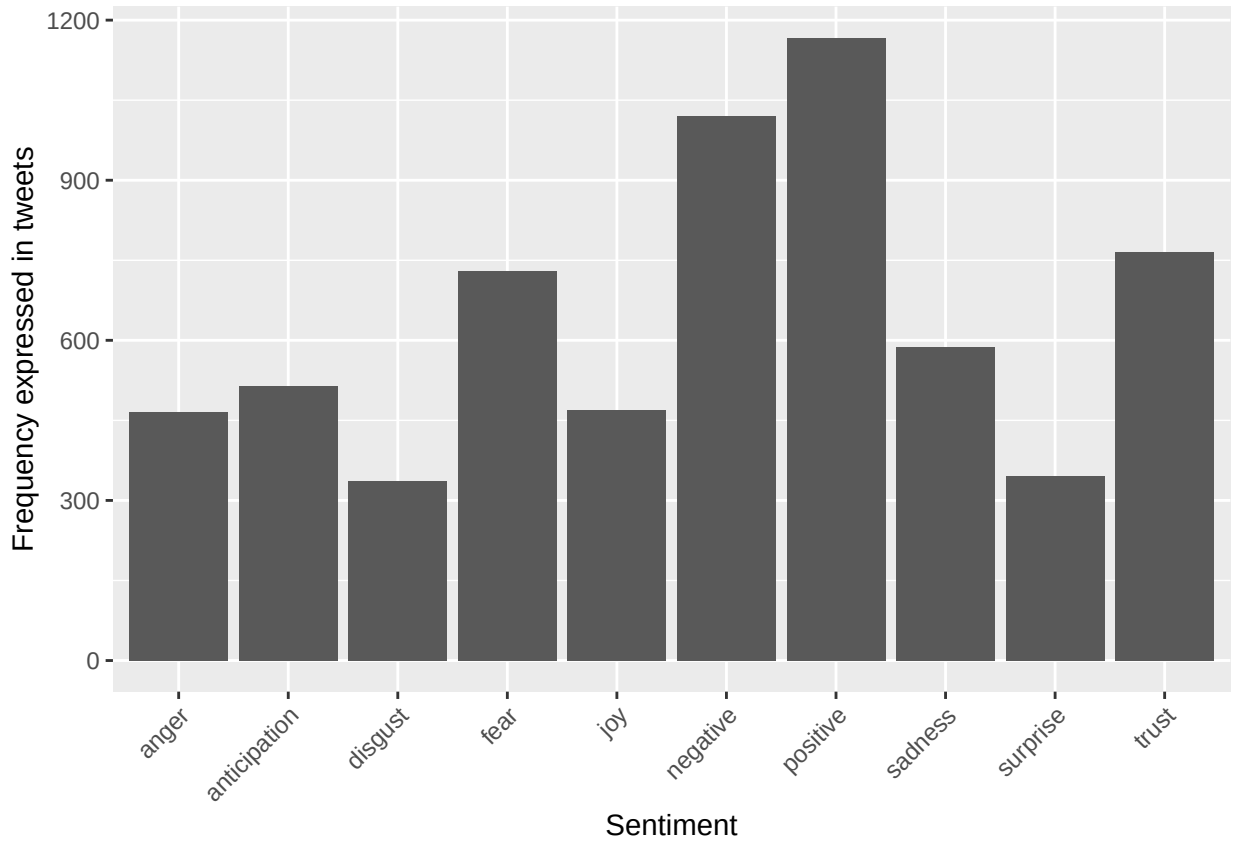


```
reg <- "([^A-Za-z\\d#@']|'(?![A-Za-z\\d#@]))"
climate_vocab <-
  tweets_climate %>%
  select(user_id,source,created_at,text) %>%
  filter(!str_detect(text, '^"')) %>%
  mutate(text = str_replace_all(text, "https://t.co/[A-Za-z\\d]+|&amp;", "")) %>%
  unnest_tokens(word, text, token = "regex", pattern = reg) %>%
  filter(!word %in% stop_words$word,
         str_detect(word, "[a-z]"))
```

```
corpus <- VCorpus(VectorSource(climate_vocab$word))
dtm <- DocumentTermMatrix(corpus)
freq <- colSums(as.matrix(dtm))
freq <- data.frame(names(freq), count = as.numeric(freq))
freq <- freq[order(-freq$count),]
#head(freq)
```

```
climatesent <- inner_join(
  get_sentiments("nrc"),
  climate_vocab,
  by="word") %>%
  count(sentiment)
```

```
ggplot(climatesent, aes(sentiment, n)) +
  geom_col() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  xlab("Sentiment") +
  ylab("Frequency expressed in tweets")
```




```
head(get_sentiments("nrc"))
```

```
## # A tibble: 6 x 2
##   word      sentiment
##   <chr>     <chr>
## 1 abacus    trust
## 2 abandon   fear
## 3 abandon   negative
## 4 abandon   sadness
## 5 abandoned anger
## 6 abandoned fear
```

```
#pal2 <- brewer.pal(8,"Dark2")
#plot_wc <- wordcloud(climate_vocab$word,scale=c(8,.2),min.freq=2,
#max.words=Inf, random.order=FALSE, rot.per=.15, colors=pal2)
```

```
require(devtools)
```

```
## Loading required package: devtools
```

```
## Loading required package: usethis
```

```
## Warning: package 'usethis' was built under R version 3.6.3
```

```
install_github("lchiffon/wordcloud2")
```

```
## Skipping install of 'wordcloud2' from a github remote, the SHA1 (8a12a3b6) has not changed since last
##   Use 'force = TRUE' to force installation
```

```
library(wordcloud2)
df <- as.data.frame(freq)
wcloud2 <- wordcloud2(data = df, size = 2.6)
wcloud2
```