

```

#Downloading required packages
library("rtweet")
library(sf)
library(raster)
library(dplyr)
library(rtweet)
library(httput)
library(magrittr)
library(tidytext)
devtools::install_github("mkearney/rtweet")
library(httr)
library(knitr)
library(knitr)
library(ggplot2)
library(tidytext)
library(stringr)
library(textdata)
library(tibble)
library(widyr)
library(ggrepel)
library(gridExtra)
library(kableExtra)
library(formattable)
library(circlize)
library(memery)
library(magick)
library(yarr)
library(radarchart)
library(igraph)
library(ggraph)
library(tidyr)
library(ggmap)
library(sf)
library(raster)
library(dplyr)
library(spData)
library(spDataLarge)
library(tmap) # for static and interactive maps
library(leaflet) # for interactive maps
library(mapview) # for interactive maps
library(ggplot2) # tidyverse data visualization package
library(shiny)
install.packages("formattable")
library(formattable)
library(data.table)

```

2

```

#create the token environment
my_tokens <- create_token (app = "UChicago_Lab",
                           consumer_key = "ldvwVWOFLJjVj8LCEaZUH1L0J",
                           consumer_secret =
"vtFk0lmNBrDpcixuSQjvZPRMbF7mnQ9Qno4ChQk3M6pStW8jw5")

```

```

#March

Reade_March <- search_fullarchive(q="Tara Reade",
  n=10000,
  fromDate = 202003010000,
  toDate = 202003312359,
  env_name="development1")

Reade_March2 <- search_fullarchive(q="TaraReade",
  n=10000,
  fromDate = 202003010000,
  toDate = 202003312359,
  env_name="development1")

Reade_March_complete <- rbind(Reade_March, Reade_March2)

Reade_March_complete

#Creating another app
my_token1 <- create_token (app = "Reade_Tweets",
  consumer_key = "OpvGvkxRDcJg3c8nJsM9HgWGR",
  consumer_secret =
"39F0T6wtQUGiBYZQ7UEmsnkbwRBXXVPcjXOWmQaa4l6ilZXvdY")

Reade_Twitter <- search_tweets("Tara Reade", n = 20,000, geocode =
lookup_coords("usa"), type = "recent", retryonratelimit = TRUE)

Reade_Twitter <- lat_lng(Reade_Twitter)
par(mar = c(0, 0, 0, 0))
maps::map("state", lwd = .25)
with(Reade_Twitter, points(lng, lat, pch = 20, cex = .75, col = rgb(0, .3,
.7, .75)))

March_geolocated <- lat_lng(Reade_March_complete)
par(mar = c(0, 0, 0, 0))
maps::map("state", lwd = .25)
with(March_geolocated, points(lng, lat, pch = 20, cex = .75, col = rgb(0,
.3, .7, .75)))

ReadeTwitter <- search_tweets("Tara Reade", n = 50,000, geocode =
lookup_coords("usa"), type = "recent", retryonratelimit = TRUE)
ReadeTwitter2 <- search_tweets("TaraReade", n = 50,000, geocode =
lookup_coords("usa"), type = "recent", retryonratelimit = TRUE)
ReadeTwitter2

big_Reade <-

mega_Reade_no_duplicates <- mega_Reade[!duplicated(mega_Reade$status_id),]

objects(mega_Reade_no_duplicates)
mega_Reade_no_duplicates
par(mar = c(0, 0, 0, 0))
maps::map("state", lwd = .25)

```

```

with(mega_Reade_no_duplicates, points(lng, lat, pch = 20, cex = .75, col =
rgb(0, .3, .7, .75)))
past_week_Reade_geo <- mega_Reade_no_duplicates
March_no_geo <- March_geolocated

#####STARTING TUTORIALS HERE
# DATASETS: past_week_Reade_geo (5000ish), March_no_geo (5600)
# I'm going to do this first with the past week, because that's what I have
geolocated

#Big question: how do people feel about Tara Reade?

##### ALL MY ATTEMPTS TO TOKENIZE BY WORD WHICH DID NOT WORK

get_sentiments("afinn")
get_sentiments("bing")
get_sentiments("nrc")
1

Reade_tb <- tibble(text = seq_along(past_week_Reade_geo))

Reade_by_word <- past_week_Reade_geo %>%
  unnest_tokens(word, text)

past_week_Reade_geo %>% unnest_tokens(word, Reade_df$text)

Reade_df <- tibble(past_week_Reade_geo)
Reade_df1 <- mutate(Reade_df, text = text)

unnest <- Reade_df1 %>%
  unnest_tokens(text,
                output = STRING,
                token="words")
View(unnest)
#that worked ish

unnest1 <- unnest_tokens(Reade_df1, output=string, token="words",
format=c("text"), to_lower=TRUE,drop=FALSE)

unnest2 <- unnest_tokens(
  tbl = Reade_df1$text,
  output=string,
  token = "words",
  to_lower = TRUE,
  drop = FALSE,
  collapse = NULL,
)

View(unnest1)
View(past_week_Reade_geo)
head(unnest1)

##### trying something else!

```

```

reade_tidy <- past_week_Reade_geo %>%
  unnest_tokens(word, text)
View(reade_tidy)

glimpse(reade_tidy)

my_colors <- c("#E69F00", "#56B4E9", "#009E73", "#CC79A7", "#D55E00",
"#D65E00")

my_theme <- function(aticks = element_blank(),
                     pgminor = element_blank(),
                     lt = element_blank(),
                     lp = "none")
  theme(plot.title = element_text(hjust = 0.5),
        axis.ticks = aticks,
        panel.grid.minor = pgminor,
        legend.title = lt,
        legend.position = lp)
my_kable_styling <- function(dat, caption) {
  kable(dat, "html", escape = FALSE, caption = caption) %>%
    kable_styling(bootstrap_options = c("striped", "condensed",
"bordered"),
                  full_width = FALSE)
}

### getting into the SA part

reade_tidy %>%
  right_join(get_sentiments("nrc")) %>%
  filter(!is.na(sentiment)) %>%
  count(sentiment, sort = TRUE)

reade_bing <- reade_tidy %>%
  inner_join(get_sentiments("bing"))

View(reade_bing)

reade_bing_by_sentiment <- reade_bing %>%
  group_by(sentiment) %>%
  count(word, sort = TRUE) %>%
  arrange(desc(n)) %>%
  slice(seq_len(8)) %>%
  ungroup()

reade_bing_by_sentiment %>%
  ggplot(aes(word, 1, label = word, fill = sentiment )) +
  geom_point(color = "transparent") +
  geom_label_repel(force = 1, nudge_y = .5,
                  direction = "y",
                  box.padding = 0.04,
                  segment.color = "transparent",
                  size = 3) +
  facet_grid(~sentiment) +
  theme_lyrics()() +

```

```

theme(axis.text.y = element_blank(), axis.text.x = element_blank(),
      axis.title.x = element_text(size = 6),
      panel.grid = element_blank(), panel.background = element_blank(),
      panel.border = element_rect("lightgray", fill = NA),
      strip.text.x = element_text(size = 9)) +
xlab(NULL) + ylab(NULL) +
ggtitle("Reade Tweets by Sentiment, end of May") +
coord_flip()

reade_nrc <- reade_tidy %>%
  inner_join(get_sentiments("nrc"))

reade_nrc_sub <- reade_tidy %>%
  inner_join(get_sentiments("nrc")) %>%
  filter(!sentiment %in% c("positive", "negative"))

reade_nrc_by_sentiment <- reade_nrc_sub %>%
  group_by(sentiment) %>%
  count(word, sort = TRUE) %>%
  arrange(desc(n)) %>%
  slice(seq_len(8)) %>%
  ungroup()

reade_nrc_by_sentiment %>%
  ggplot(aes(word, 1, label = word, fill = sentiment )) +
  geom_point(color = "transparent") +
  geom_label_repel(force = 1, nudge_y = .5,
                  direction = "y",
                  box.padding = 0.04,
                  segment.color = "transparent",
                  size = 3) +
  facet_grid(~sentiment) +
  theme_lyrics()() +
  theme(axis.text.y = element_blank(), axis.text.x = element_blank(),
        axis.title.x = element_text(size = 6),
        panel.grid = element_blank(), panel.background = element_blank(),
        panel.border = element_rect("lightgray", fill = NA),
        strip.text.x = element_text(size = 9)) +
  xlab(NULL) + ylab(NULL) +
  ggtitle("Reade Tweets by sentiment, end of May") +
  coord_flip()

View(reade_nrc_by_sentiment)

#SPIDER CHART

year_sentiment_nrc <- reade_nrc_sub %>%
  group_by(sentiment) %>%
  count(sentiment) %>%
  select(sentiment, "Sentiment Count" = n)

reade_sentiment_nrc <- year_sentiment_nrc

#words by sentiment

```

```

reade_sentiment_nrc

spider_reade <- reade_sentiment_nrc %>%
  radarchart::chartJSRadar(showToolTipLabel = TRUE,
    main = "end-May Sentiment Counts")

spider_reade

## DOING THAT WHOLE THING AGAIN W THE MARCH

reade_March_tidy <- March_no_geo %>%
  unnest_tokens(word, text)

glimpse(reade_March_tidy)

reade_March_tidy %>%
  right_join(get_sentiments("nrc")) %>%
  filter(!is.na(sentiment)) %>%
  count(sentiment, sort = TRUE)

reade_bing_March <- reade_March_tidy %>%
  inner_join(get_sentiments("bing"))

View(reade_bing)

reade_bing_by_sentiment_March <- reade_bing_March %>%
  group_by(sentiment) %>%
  count(word, sort = TRUE) %>%
  arrange(desc(n)) %>%
  slice(seq_len(8)) %>% #consider top_n() from dplyr also
  ungroup()

reade_bing_by_sentiment %>%
  #Set `y = 1` to just plot one variable and use word as the label
  ggplot(aes(word, 1, label = word, fill = sentiment )) +
  #You want the words, not the points
  geom_point(color = "transparent") +
  #Make sure the labels don't overlap
  geom_label_repel(force = 1, nudge_y = .5,
    direction = "y",
    box.padding = 0.04,
    segment.color = "transparent",
    size = 3) +
  facet_grid(~sentiment) +
  theme_lyrics()() +
  theme(axis.text.y = element_blank(), axis.text.x = element_blank(),
    axis.title.x = element_text(size = 6),
    panel.grid = element_blank(), panel.background = element_blank(),
    panel.border = element_rect("lightgray", fill = NA),
    strip.text.x = element_text(size = 9)) +
  xlab(NULL) + ylab(NULL) +
  ggtitle("Reade Tweets by Sentiment, March") +
  coord_flip()

```

```

reade_nrc_March <- reade_March_tidy %>%
  inner_join(get_sentiments("nrc"))

reade_nrc_sub_March <- reade_March_tidy %>%
  inner_join(get_sentiments("nrc")) %>%
  filter(!sentiment %in% c("positive", "negative"))

reade_nrc_by_sentiment_March <- reade_nrc_sub_March %>%
  group_by(sentiment) %>%
  count(word, sort = TRUE) %>%
  arrange(desc(n)) %>%
  slice(seq_len(8)) %>% #consider top_n() from dplyr also
  ungroup()

reade_nrc_by_sentiment_March %>%
  #Set `y = 1` to just plot one variable and use word as the label
  ggplot(aes(word, 1, label = word, fill = sentiment )) +
  #You want the words, not the points
  geom_point(color = "transparent") +
  #Make sure the labels don't overlap
  geom_label_repel(force = 1, nudge_y = .5,
    direction = "y",
    box.padding = 0.04,
    segment.color = "transparent",
    size = 3) +
  facet_grid(~sentiment) +
  theme_lyrics()() +
  theme(axis.text.y = element_blank(), axis.text.x = element_blank(),
    axis.title.x = element_text(size = 6),
    panel.grid = element_blank(), panel.background = element_blank(),
    panel.border = element_rect("lightgray", fill = NA),
    strip.text.x = element_text(size = 9)) +
  xlab(NULL) + ylab(NULL) +
  ggtitle("Reade Tweets by sentiment, March") +
  coord_flip()

View(reade_nrc_by_sentiment)

#SPIDER CHART - MARCH

year_sentiment_nrc_March <- reade_nrc_sub_March %>%
  group_by(sentiment) %>%
  count(sentiment) %>%
  select(sentiment, "Sentiment Count" = n)

reade_sentiment_nrc_March <- year_sentiment_nrc_March

#words by sentiment

spider_reade_March <- reade_sentiment_nrc_March %>%
  radarchart::chartJSRadar(showToolTipLabel = TRUE,
    main = "March Sentiment Counts")

spider_reade_March

```

```
###THINKING ABOUT HOW TO MAP THE COLORS
```

```
reade_nrc_sub_March <- lat_lng(reade_nrc_sub_March)
par(mar = c(0, 0, 0, 0))
maps::map("state", lwd = .25)
with(Reade_Twitter, points(lng, lat, pch = 20, cex = .75,
theme(theme_lyrics())))
```

```
View(reade_nrc_sub_March)
glimpse(reade_nrc_sub_March)
```

```
trust_reade_March <- subset(reade_nrc_sub_March, sentiment=="trust")
View(trust_reade_March)
glimpse(trust_reade_March)
```

```
trust_geo <- lat_lng(trust_reade_March)
```

```
anger_reade_March <- subset(reade_nrc_sub_March, sentiment=="anger")
glimpse(anger_reade_March)
```

```
anger_geo <- lat_lng(anger_reade_March)
```

```
anticipation_reade_March <- subset(reade_nrc_sub_March,
sentiment=="anticipation")
anticipation_geo <- lat_lng(anticipation_reade_March)
glimpse(anticipation_geo)
```

```
fear_reade_March <- subset(reade_nrc_sub_March, sentiment=="fear")
fear_geo <- lat_lng(fear_reade_March)
glimpse(fear_geo)
```

```
joy_reade_March <- subset(reade_nrc_sub_March, sentiment=="joy")
joy_geo <- lat_lng(joy_reade_March)
glimpse(joy_geo)
```

```
disgust_reade_March <- subset(reade_nrc_sub_March, sentiment=="disgust")
disgust_geo <- lat_lng(disgust_reade_March)
glimpse(disgust_geo)
```

```
surprise_reade_March <- subset(reade_nrc_sub_March, sentiment=="surprise")
surprise_geo <- lat_lng(surprise_reade_March)
glimpse(surprise_geo)
```

```
sadness_reade_March <- subset(reade_nrc_sub_March, sentiment=="sadness")
sadness_geo <- lat_lng(sadness_reade_March)
glimpse(sadness_geo)
```

```
par(mar = c(0, 0, 0, 0))
maps::map("state", lwd = .25)
with(trust_geo, points(lng, lat, pch = 20, cex = .2, col="magenta2"))
with(surprise_geo, points(lng, lat, pch = 20, cex = .5, col = "purple1"))
```



```

with(anger_geo, points(lng, lat, pch = 20, cex = .6, col = "tomato2"))
with(anticipation_geo, points(lng, lat, pch = .8, cex = .7, col =
"darkgoldenrod1"))
with(fear_geo, points(lng, lat, pch = 20, cex = 1, col = "mediumseagreen"))
with(joy_geo, points(lng, lat, pch = 20, cex = 1.2, col = "lightseagreen"))
with(disgust_geo, points(lng, lat, pch = 20, cex = 1.5, col =
"yellowgreen"))
with(sadness_geo, points(lng, lat, pch = 20, cex = 2, col =
"deepskyblue2"))

```

FOR THE LAST WEEK OF MAY

```

trust_reade_May <- subset(reade_nrc_sub, sentiment=="trust")
View(trust_reade_May)
glimpse(trust_reade_May)
trust <- lat_lng(trust_reade_May)

anger_reade_May <- subset(reade_nrc_sub, sentiment=="anger")
glimpse(anger_reade_May)

anger <- lat_lng(anger_reade_May)

anticipation_reade_May <- subset(reade_nrc_sub, sentiment=="anticipation")
anticipation <- lat_lng(anticipation_reade_May)
glimpse(anticipation)

fear_reade_May <- subset(reade_nrc_sub, sentiment=="fear")
fear <- lat_lng(fear_reade_May)
glimpse(fear)

joy_reade_May <- subset(reade_nrc_sub, sentiment=="joy")
joy <- lat_lng(joy_reade_May)
glimpse(joy)

disgust_reade_May <- subset(reade_nrc_sub, sentiment=="disgust")
disgust <- lat_lng(disgust_reade_May)
glimpse(disgust_geo)

surprise_reade_May <- subset(reade_nrc_sub, sentiment=="surprise")
surprise <- lat_lng(surprise_reade_May)
glimpse(surprise)

sadness_reade_May <- subset(reade_nrc_sub, sentiment=="sadness")
sadness <- lat_lng(sadness_reade_May)
glimpse(sadness)

par(mar = c(0, 0, 0, 0))
maps::map("state", lwd = .25)
with(trust, points(lng, lat, pch = 20, cex = .5, col="lightskyblue"))
with(surprise, points(lng, lat, pch = 20, cex = .5, col = "purple3"))
with(anger, points(lng, lat, pch = 20, cex = .5, col = "darkorange1"))
with(anticipation, points(lng, lat, pch = 20, cex = .5, col = "orchid"))
with(fear, points(lng, lat, pch = 20, cex = .5, col = "firebrick4"))
with(joy, points(lng, lat, pch = 20, cex = .5, col = "dodgerblue2"))

```

```
with(disgust, points(lng, lat, pch = 20, cex = .5, col = "firebrick2"))
with(sadness, points(lng, lat, pch = 20, cex = .5, col = "gold"))
```

```
theme_lyrics
```

```
View(reade_bing_by_sentiment)
```

```
reade_latlong <- lat_lng(reade_nrc_sub)
1
```

```
#unnested: reade_tidy
```

```
lexicon <- tibble(word =c(
                        'ibelievetara',
                        'believewomen',
                        'sexualpredator',
                        'ibelievebiden',
                        'bluenomatterwho',
                        'tarareadeisafraud'
                      ),
                  scores=c(1, 1, 1, -1, -1, -1))
```

```
beliefMarch <- merge(tidy_March, lexicon, by="word")
```

```
View(lexicon)
```

```
beliefdoubtlexicon <- merge(reade_tidy, lexicon, by="word")
View(beliefdoubtlexicon)
glimpse(beliefdoubtlexicon)
View(d_march)
```

```
View(beliefMarch)
```

```
View(beliefdoubtlexicon)
```

```
View(b_may)
```

```
View(d_may)
```

```
table <- matrix(c(888,0,29,5), ncol=2)
colnames(table) <- c('March', 'May')
rownames(table) <- c('Belief', 'Doubt')
table.table <- as.table(table)
formattable(table)
format_table(table)
```

```
formattable(d_may)
formattable(table)
```

```

table_tibble <- data_frame(table)

beliefdoubt <- lat_lng(beliefdoubtlexicon)

belief <- subset(beliefdoubtlexicon, scores=="1")
View(belief)

doubt <- subset(beliefdoubtlexicon, scores=="-1")
View(doubt)

par(mar = c(0, 0, 0, 0))
maps::map("state", lwd = .25)
with(belief, points(lng, lat, pch = 20, cex = .5, col="green"))
with(doubt, points(lng, lat, pch = 20, cex = .5, col = "purple"))

#okay let's try it with the March data

  tidy_March <- March_no_geo %>%
    unnest_tokens(word, text)
View(tidy_March)

##my attempt to bring together the two

###so we're going to take the beliefdoubt set and merge it with the tidy
one

tidy_w_scores <- rbind(tidy_March, beliefdoubt, by = "status_id")

View(tidy_w_scores)
tidy_nested <- tidy_w_scores %>%
  group_by(status_id)
View(tidy_w_scores)

View(beliefdoubt)

View(tidy_March)

score <- cbind(t())

X <- data.frame(name=LETTERS[1:3], value=1:3, stringsAsFactors=FALSE)

grouped <- beliefdoubt %>% group_by(scores) %>%
  mutate(words_in_lexicon = n_distinct(word)) %>%
  ungroup()

glimpse(beliefdoubt)

table
View(table)

install.packages("gt")
library(gt)

```

```
table %>% gt()
```