Drone Sentiment Analysis

Opinions about Drones, sentiment analysis

```r
# Basic
library(tidyverse)
library(scales)
library(knitr)

# For cleaning / filtering text
library(stringr)
library(wordcloud)
library(tm)
library(janeaustenr)
library(textstem)
library(syuzhet)
library(sentimentr)

# For rendering
library(rmarkdown)
library(stringr)
library(gridExtra)
library(grid)
```

```r
data1 <- read_csv("data/data1.csv")

# The number of total responses
pre_length = nrow(data1)

# Filtering and cleaning responses
response <- data1 |>
  select(Q6, Q8) |>
```

```r
  rename(
    reason = Q6, # text
    opinion = Q8 # number
  )|>
  mutate(reason = gsub("(\n|<br />)"," ",reason)) |>
  mutate(reason = gsub("'","",reason)) |>
  mutate(reason = gsub("'","",reason)) |>
  mutate(reason = gsub("."," ",reason, fixed = TRUE)) |> #prevent sentence splitting
  mutate(reason = gsub("?","",reason, fixed = TRUE)) |> #prevent sentence splitting
  mutate(reason = gsub("!","",reason, fixed = TRUE)) |> #prevent sentence splitting
  drop_na()



# Removing first 2 rows
response = response[-c(1, 2),]

# Reformatting opinion column
response$opinion = as.numeric(response$opinion)

# Lemmatizing words - reducing them to base form
lemmatize <- function(sentence) {
  return(paste(lemmatize_words(strsplit(sentence, " ")[[1]]),
        collapse=" "))
}
response[c("reason")] <- apply(response[c("reason")], 1,lemmatize)


# Adding sentence-level sentiments to each response
#response |>
#  mutate(sentiment = sentiment(reason)$sentiment)

# Function to turn a column into a corpus
create_corpus <- function(column) {

  # Creating Corpus for ALL RAW REASONS
  corpus <- VCorpus(
    VectorSource(
      as.vector(column))
    )
```

```r
  corpus <- corpus |>
    tm_map(removeNumbers) |>
    tm_map(removePunctuation) |>
    tm_map(stripWhitespace)

  corpus <- tm_map(corpus, content_transformer(tolower))
  corpus <- tm_map(corpus, removeWords, stopwords("english"))

  return (corpus)
}

# Corpus for all REASONS
q6_corpus = create_corpus(response$reason)


# Creating Corpus for OPINIONS & REASONS
reason_list <- split(response, response$opinion)
opinion_corpus_list <- list()


# Corpuses for the REASONS, split per OPINION (1-5)
opinion_1_corpus = create_corpus(reason_list[[1]]$reason)
opinion_2_corpus = create_corpus(reason_list[[2]]$reason)
opinion_3_corpus = create_corpus(reason_list[[3]]$reason)
opinion_4_corpus = create_corpus(reason_list[[4]]$reason)
opinion_5_corpus = create_corpus(reason_list[[5]]$reason)

# Corpuses for the REASONS, split by GENERAL OPINION
# where scores of 1 or 2 = oppose & 4 or 5 = support


# Creating corpuses for general support / opposition
opinion_support_corpus = create_corpus(
  rbind(reason_list[[1]], reason_list[[2]])$reason
)

opinion_oppose_corpus = create_corpus(
  rbind(reason_list[[4]], reason_list[[5]])$reason
)

# Calculating the number of real responses
```

```
post_length = nrow(response)
yield = round((((post_length / pre_length) * 100), 3)

#response
```

The data has 2146 responses. The percentage of non-blank responses is 93.476 %

```
# Returns a word-frequency matrix from a corpus
get_wfm <- function(corpus) {
  dtm <- TermDocumentMatrix(corpus)
  matrix <- as.matrix(dtm)
  words <- sort(rowSums(matrix),decreasing=TRUE)
  wfm <- data.frame(word = names(words),freq=words)

  return (wfm)
}


# Draws the word-frequency graph between 2 corpuses
draw_wfm_diff_graph <- function(corpus_1, corpus_2, n_width, p_width) {

  # Creating word-frequency matrix
  opinion_1_wfm = get_wfm(corpus_1)
  opinion_5_wfm = get_wfm(corpus_2)

  # Adding negative sign to opposing views
  opinion_1_wfm <- opinion_1_wfm |>
    mutate(freq = -freq)

  # Calculating the relative frequencies
  opinion_1_wfm <- opinion_1_wfm |>
    mutate(ratio = freq / (nrow(opinion_1_wfm)))

  opinion_5_wfm <- opinion_5_wfm |>
    mutate(ratio = freq / (nrow(opinion_5_wfm)))

  # Finding the difference in word frequencies
  opinion_diff <- rbind(opinion_1_wfm, opinion_5_wfm)
  opinion_diff <- opinion_diff |>
    group_by(word) |>
    summarize(diff_freq = sum(ratio)) |>
```

```r
    arrange(desc(diff_freq))

  # Joining the most significant words
  largest_diff <- rbind(head(opinion_diff, p_width), tail(opinion_diff, n_width))


  # Drawing the grap
  largest_diff |>
    ggplot(
      aes(reorder(word, -diff_freq, sum),
          diff_freq, fill = diff_freq < 0)) +

    geom_bar(stat="identity") +

    coord_flip() +
    ylab("Difference in word frequency") +
    xlab("Word") +
    ggtitle("Word frequency of drone export supporters / opposers") +

    scale_fill_manual(
      values=c("#77dd76", "#f69185"),
      name="Drone export\nopinion",
                      breaks=c("FALSE", "TRUE"),
                      labels=c("Support", "Oppose")) +


    ylim(-0.1,0.09) +
    scale_y_continuous(labels = scales::percent_format(accuracy = 2))
}

draw_wfm_diff_graph(opinion_support_corpus, opinion_oppose_corpus, 15, 15)
```
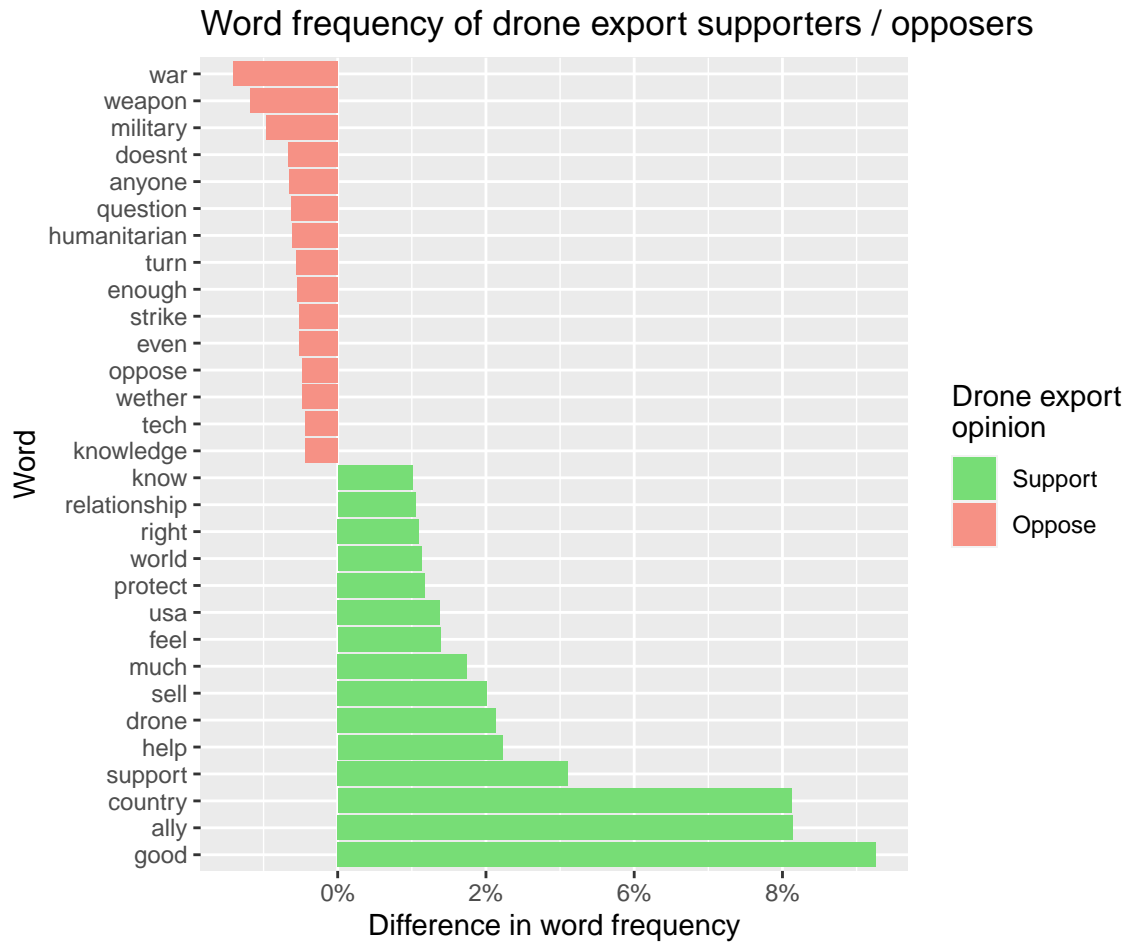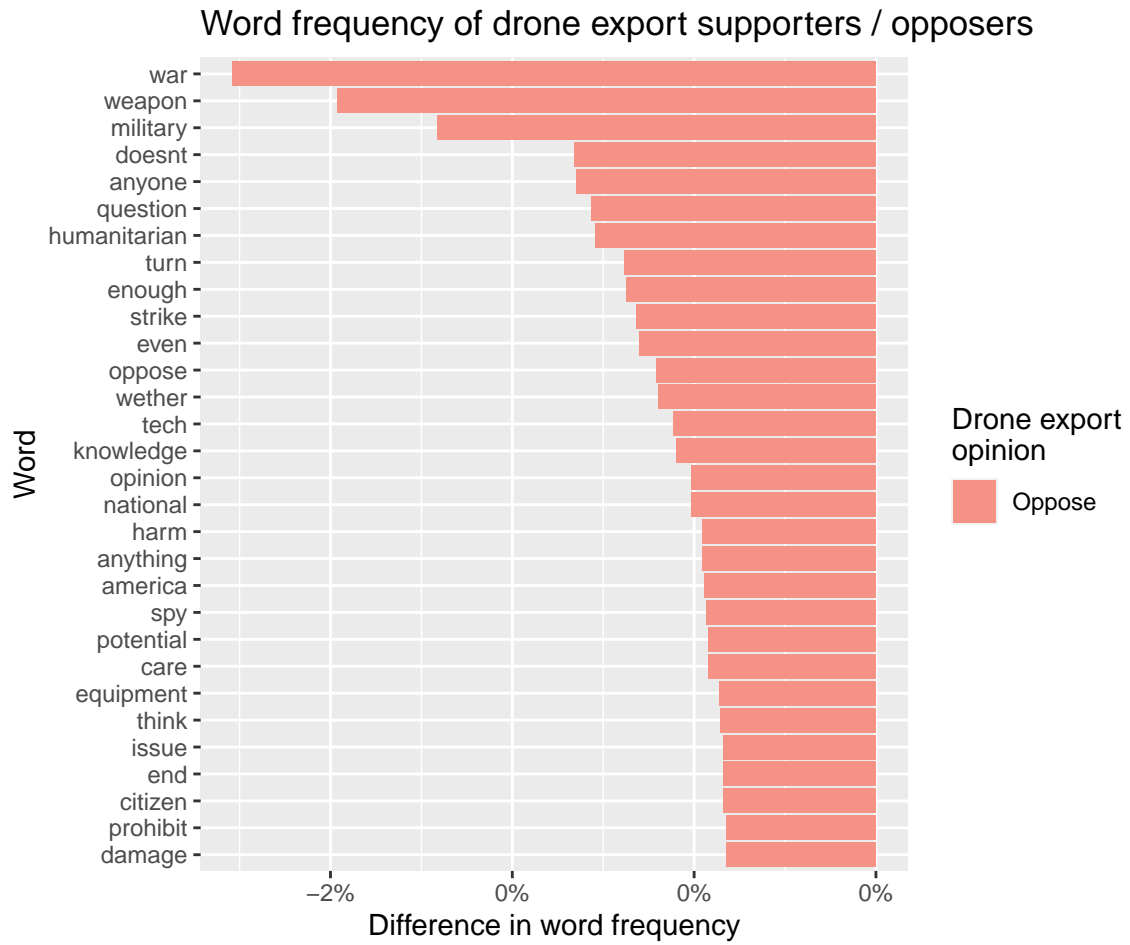
Word frequency of drone export supporters / opposers

The difference in word frequency represents the percentage of the difference between the words used. For example, "good" is featured 8% more frequently in supportive votes than in opposing ones. On the other hand, "war" and "weapon" are featured more frequently in opposing votes.
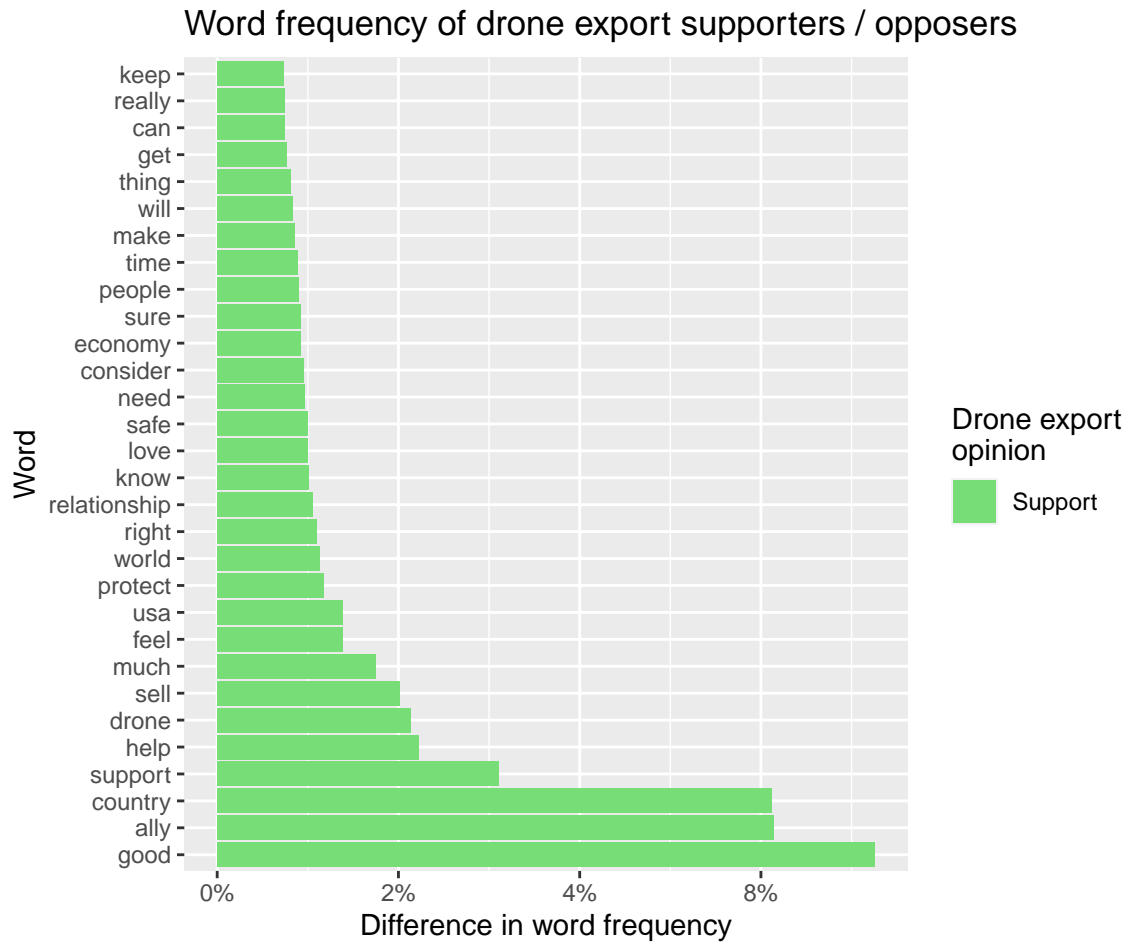
The reason the word frequencies for supporters peak much more than for people opposing drone exports is because opposing viewpoints tended to use a greater variety of language, whereas drone export supporters mostly used the same words like "good", "country" and "ally".

```
draw_wfm_diff_graph(opinion_support_corpus, opinion_oppose_corpus, 30, 0)
```

Word frequency of drone export supporters / opposers

A closer analysis at the comments of the people opposing drone exports shows many of the factors people are concerned about.

```
draw_wfm_diff_graph(opinion_support_corpus, opinion_oppose_corpus, 0, 30)
```

Word frequency of drone export supporters / opposers

Similarly, looking at the comments by the people supporting drone exports reveals the reasoning behind their opinions.

```
# Draws a Word Cloud based from a corpus
generate_wordcloud <- function(corpus) {
  wfm <- get_wfm(corpus)
  wfm <- wfm[c(-1, -2),] # removing the words "drone" and "country"
  set.seed(1234)
  wordcloud(words = wfm$word,
            freq = wfm$freq,
            min.freq = 1,
            max.words=200,
            random.order=FALSE,
            rot.per=0.35,
```

```
            colors=brewer.pal(8, "Dark2"))
}

# Drawing worldcloud for STRONG OPPOSITION (1)
generate_wordcloud(opinion_1_corpus)
```


A word cloud of words used by people who voted (1), "Strongly Opposing" drone exports.

```
# Drawing worldcloud for STRONG SUPPORT (5)
generate_wordcloud(opinion_5_corpus)
```

A word cloud of words used by people who voted (5), "Strongly Supporting" drone exports.

10 Most Positive responses

```
score_responses <- response |>
  mutate(score = get_sentiment(reason)) |>
  arrange(score)


head(score_responses, 10) |>
  select(reason, score)
```

```
# A tibble: 10 x 2
```

```
      reason                                                      score
      <chr>                                                       <dbl>
 1 I consider the privacy issue involve Also that of insurance And the im~ -2.65
 2 None at the moment bitch ass whore                             -2.5
 3 good the us have problem with some of this country to sell them drone ~ -2.4
 4 this country be our friend and we should support them Ukraine be unjus~ -2.4
 5 I think as long as the item be not be use for threat like terrorist at~ -2.25
 6 I consider the fact drone cant be see as good as a tank or helicopter ~ -2.15
 7 I do support any of it We be definitely fight a war that we do fully u~ -2.15
 8 The surveillance state will soon turn into fascism Authoritarianism be~ -2.15
 9 It be really vague on where Turkey be sell the drone and other weapon ~ -2.15
10 The history of the conflict the government have be fight            -2
```

10 Most Negative responses

```
  tail(score_responses, 10) |>
    select(reason, score)
```

```
# A tibble: 10 x 2
   reason                                                      score
   <chr>                                                       <dbl>
 1 "I consider whether or not the country be a ally of ours or not  In ad~  3.1
 2 "I believe that, other country such as the one listed, often face thre~  3.15
 3 "I consider whether I think the recipient would use it in a malicious ~  3.2
 4 "That how versatile and useful than can be for our ally in aid in what~  3.3
 5 "Humanitarian aid, current ally with other country I support defensive~  3.3
 6 "hello good even babe good morning babe good night love you merry merr~  3.6
 7 "How it will assist ally in their defense mechanism and protect essent~  3.7
 8 "be a large-scale community map  Its objective be to improve urban  it~  3.95
 9 "The only owner I know that have any interest be a real estate attorne~  4.6
10 "To support our ally because they provide safety and easy to navigate ~  5.25
```

This shows the most positive and negative arguments for drone exports to foreign exports.

```
  # Adds line breaks to text
  prnt.test <- function(x){
     cat(x, sep="\n\n")
  }

  # Adds comma flags to text
  break_words <- function(startstring) {
```

```r
    words = strsplit(startstring, ' ')[[1L]]
    splits = cut(seq_along(words), breaks = seq(0L, length(words) + 12L, by = 12L))
    paste(lapply(split(words, splits), paste, collapse = ' '), collapse = '\n')
  }

  # Prints everything with line breaks every 12 words.
  prnt_all_multi <- function(neg_lst) {
    for (x in 1:nrow(neg_lst)) {
      print(sprintf("[Response %s]", x))
      writeLines(break_words(neg_lst$reason[[x]]))
      writeLines("\n")
    }
  }
```

The full text for the 10 most **positive** responses is below.

```r
  prnt_all_multi(neg_lst = tail(score_responses, 10) |>
                   select(reason))
```

```
[1] "[Response 1]"
I consider whether or not the country be a ally of ours
or not  In addition, certain country that may not always be
friendly with us, I consider that as good


[1] "[Response 2]"
I believe that, other country such as the one listed, often face
threat and should have a right to safety and access to equipment
that can alert them of dangerous activity I also consider the fact
that we as the US have acquire help system from many different
region and it wouldnt be very considerate of us to not share
our useful asset


[1] "[Response 3]"
I consider whether I think the recipient would use it in a
malicious manor or not, I also try to think if there can
be any possible benefit assist humanitarian be a good benefit


[1] "[Response 4]"
That how versatile and useful than can be for our ally in
```

aid in what they can use them for

[1] "[Response 5]"
Humanitarian aid, current ally with other country I support defensive sale to
ally along with humanitarian aide But I do want to help support
and sponsor war with our technology Im split on the Ukraine support
But we need to assist with aide for that in need

[1] "[Response 6]"
hello good even babe good morning babe good night love you merry
merry dear dear amen

[1] "[Response 7]"
How it will assist ally in their defense mechanism and protect essential
asset

[1] "[Response 8]"
be a large-scale community map Its objective be to improve urban
 itresilience to flood by provide accurate spatial information to support

[1] "[Response 9]"
The only owner I know that have any interest be a real
estate attorney and he be the one   I know that be
a good person to be around and that be the only thing
that can be fix and he be very nice and trustworthy and
he be a very good   I be very grateful for the
opportunity he have offer me I be so grateful   to him
and my family for all of the work he put in to
help us and for his work

[1] "[Response 10]"
To support our ally because they provide safety and easy to navigate
, precision good than aircraft

Similarly, the 10 most **negative** responses are here:

```
prnt_all_multi(neg_lst = head(score_responses, 10) |>
                 select(reason))
```

[1] "[Response 1]"
I consider the privacy issue involve Also that of insurance And the
implication of damage cause by fail drone And furthermore the copyright infringement
of drone fly over outdoor concert or other public forum

[1] "[Response 2]"
None at the moment bitch ass whore

[1] "[Response 3]"
good the us have problem with some of this country to sell
them drone to strike in the near future I would have to
oppose because thats putting a threat on be country

[1] "[Response 4]"
this country be our friend and we should support them Ukraine be
unjustly attack and invade by Russia and we should do everything to
support them in theyre struggle against the Russian invasion

[1] "[Response 5]"
I think as long as the item be not be use for
threat like terrorist attack or anything else that would bring harm, as
long as they be sell this device and not give them away
for free like Biden have be do to the Ukraine

[1] "[Response 6]"
I consider the fact drone cant be see as good as a
tank or helicopter and can pose serious harm if use against us
Why arm the enemy of America with deadly equipment to use against
us or spy on us

[1] "[Response 7]"
I do support any of it We be definitely fight a war
that we do fully understand the depth of destruction it will reap

on us and generation and generation to come And for what A
few certain people to have the much power and unlimited fund while
they watch their own people starve and suffer live in high violent
crime area

[1] "[Response 8]"
The surveillance state will soon turn into fascism Authoritarianism be rise everywhere
The next war be come but we can delay it if Ukraine
win

[1] "[Response 9]"
It be really vague on where Turkey be sell the drone and
other weapon to It be a really naive assumption to make that
once you sell something to someone it wont wind up in another
person hand

[1] "[Response 10]"
The history of the conflict the government have be fight

Note that these responses have grammatical issues because all the words have been reduced
to their base form, and all their punctuation has been removed.