Drone Sentiment Analysis

Alexey Elkin

6/15/23

What do people think about drone exports?

```
# 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)
```

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

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

# Filtering responses
```

```r
response <- data1 |>
  select(Q6, Q8) |>
  rename(
    reason = Q6, # text
    opinion = Q8 # number
  )

# Adding row indices
response$index <- 1:nrow(response)

# Response without filtering
pure_response <- response |>
  drop_na()

# Cleaning responses
response <- response |>

  # Removing punctuation
  mutate(reason = gsub("(\n|<br />)"," ",reason)) |>
  mutate(reason = gsub("'","",reason)) |>
  mutate(reason = gsub("’","",reason)) |>

  # Prevent sentence splitting
  mutate(reason = gsub(".","",reason, fixed = TRUE)) |>
  mutate(reason = gsub("?","",reason, fixed = TRUE)) |>
  mutate(reason = gsub("!","",reason, fixed = TRUE)) |>
  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)
```

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

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

  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)
```

First, we clean and format the survey data to make it easier to analyse. All punctuation and special characters are removed, and all the words are *lemmatized* - reduced to their base form. For example, "walked", "walking" and "walks" will be reduced to "walk". This ensures that the words frequency tables and sentiment analysis doesn't have doubled words.

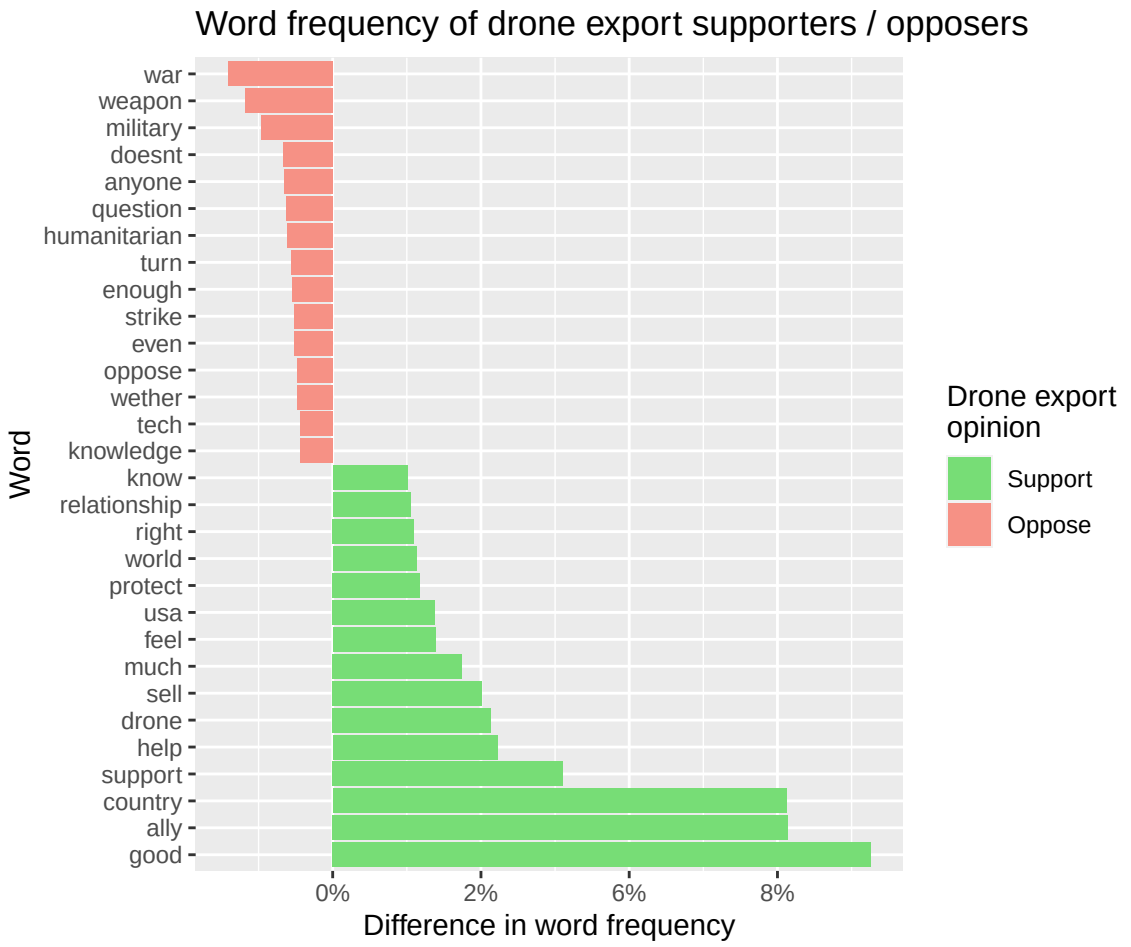
Then, the data is split along the person's answer to Question 8: *"To what extent do you agree that U.S. officials have a moral obligation to sell U.S.-manufactured drones to allies and partners?"*. The responses are split into 5 groups based on their response to on a 1-5 scale, where users that selected [1] Strongly Oppose drone exports and those who selected [5] Strongly Support them.

Overall, 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
```

```r
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)) |>
  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)
```



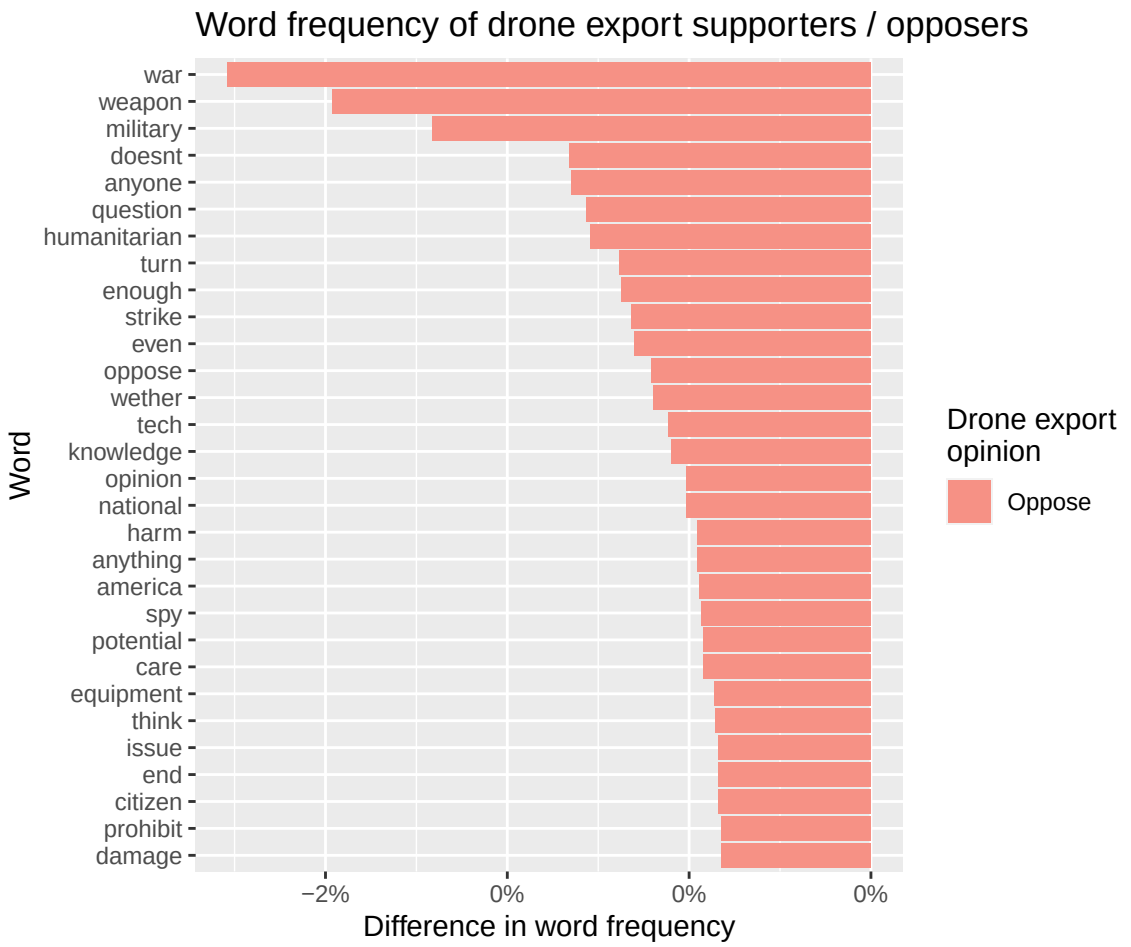
Then, we create a word frequency graph that covers all of the responses to Question 6: *"What factors did you consider while evaluating your support for the export of U.S.-manufactured drones?"*. Then, the respondents are split into 2 groups, those who support drone exports [Question 8: 4 or 5], and those who oppose them [Question 8: 1 or 2]. We calculate the ratio of each word to the total response length in their category. For example, "good" might make up 10% of words in the positive responses. After doing this for both categories, we find the difference between them, and plot a graph of a results.

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.

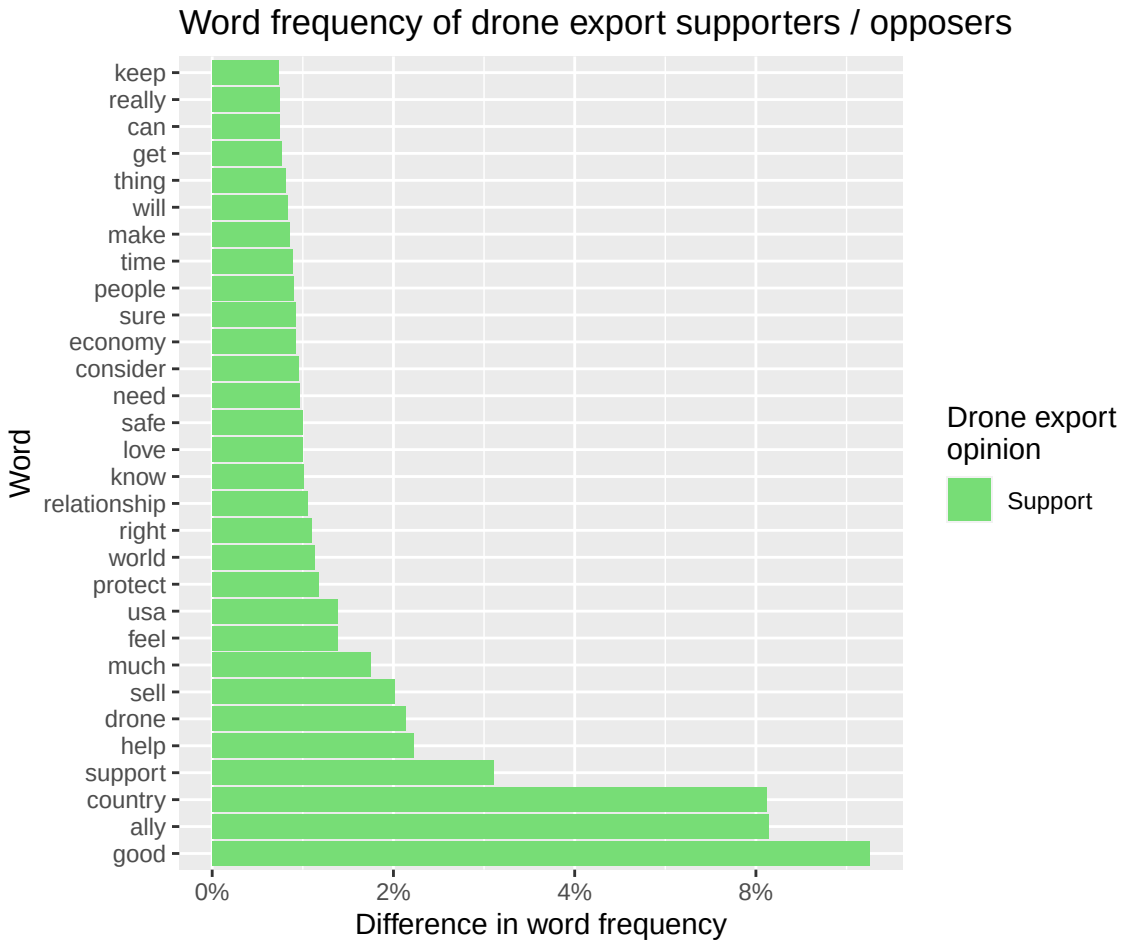
The reason the word frequencies for supporters peak much more than for people opposing drone exports is likely because opposing viewpoints tended to use a greater variety of language, whereas drone export supporters mostly used the same words like "good", "country" and "ally".

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



A closer analysis at the comments of the people opposing drone exports shows many of the factors people are concerned about. For example, people are concerned about the "humanitarian" aspects of drone exports, as well as mentioning "strikes" and "spy".

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

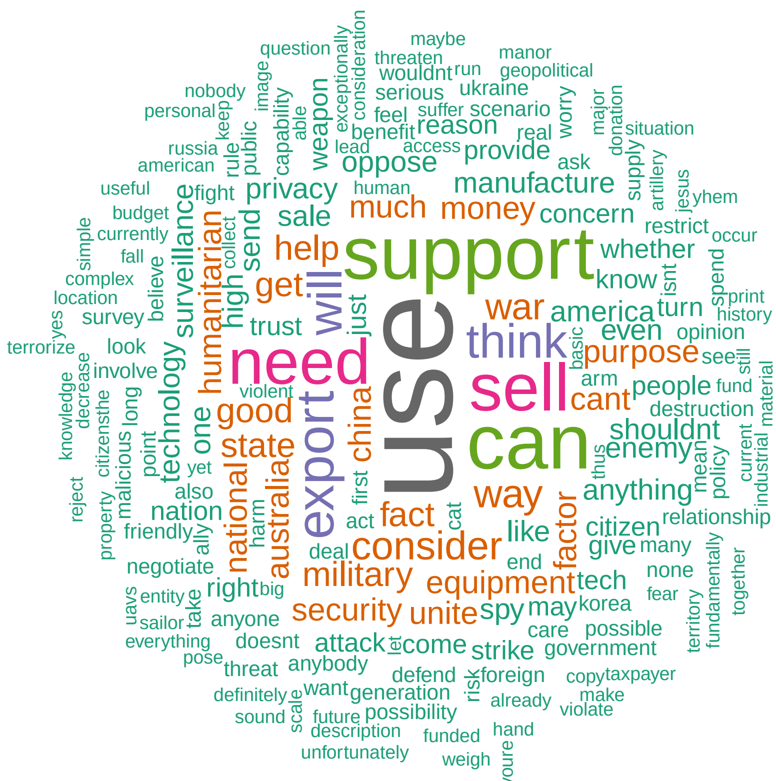


Similarly, looking at the comments by the people supporting drone exports reveals the reasoning behind their opinions. People mention "relationships", "economy", "sell", "help", and "support" very frequently.

```
# 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 to [1] Strongly Oppose drone exports.

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

A word cloud of words used by people who voted to [5] Strongly Support drone exports.

Sentiment Analysis of Responses

To Question 6: *"What factors did you consider while evaluating your support for the export of U.S.-manufactured drones?"*

10 Most positive responses

```r
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

```r
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
```

We performed a sentiment analysis of the responses to Question 6, and assigned "scores" to reach responses that measure how positive or negative they are. Then, we selected the 10 most extreme responses in each category, showing the most positive and negative arguments for drone exports to foreign countries.

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

# Adds comma flags to text
break_words <- function(startstring) {
  words = strsplit(startstring, ' ')[[1L]]
  splits = cut(seq_along(words), breaks = seq(0L, length(words) + 10L, by = 10L))
  paste(lapply(split(words, splits), paste, collapse = ' '), collapse = '\n')
}

# Prints everything with line breaks every 10 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")
  }
}
```

```
# Returns the 'pure' version of the text.
# This is the version without the punctuation filtering
# And without the words being reduced to their base form.

get_pure <- function(x) {

  df <- merge(x, pure_response, by="index", all=TRUE) |>
    drop_na(reason.x, score) |>
    select(reason.y)

  return (df)
}

hd <- head(score_responses, 100) |>
  select(reason, score, index)
```

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

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

```
[1] "[Response 1]"
I considered whether or not the country was an ally
of ours or not.  In addition, certain countries that
may not always be friendly with us, I considered that
as well.


[1] "[Response 2]"
I believe that, other countries such as the ones listed,
often face threats and should have a right to safety
and access to equipment that can alert them of dangerous
activity. I also considered the fact that we as the
U.S have acquired help systems from many different regions and
it wouldn't be very considerate of us to not share
our useful assets.


[1] "[Response 3]"
Hi good evening babe good morning babe good night love
you merry merry dear dear amen


[1] "[Response 4]"
Humanitarian aid, current Allie’s with other countries. I support defensive
sales to Allie’s along with humanitarian aide. But I don’t
want to help support and sponsor war with our technology.
I’m split on the Ukraine support. But we need to
assist with aide for those in need.


[1] "[Response 5]"
is a large-scale community mapping  Its objective is to
improve urban  itresilience to flooding by providing accurate spatial
```

```
information to support


[1] "[Response 6]"
The only owner I know that has any interest is
a real estate attorney and he is the one
I know that is a good person to be around
and that is the only thing that can be fixed
and he is very nice and trustworthy and he is
a very good  I am very grateful for the
opportunity he has offered me I am so grateful
to him and my family for all of the work
he put in to help us and for his work



[1] "[Response 7]"
How it will assist allies in their defense mechanisms and
protecting essential assets.


[1] "[Response 8]"
That how versatile and useful than can be for our
Allie's in aiding in what they can use them for.




[1] "[Response 9]"
I consider whether I think the recipient would use it
in a malicious manor or not, I also try to
think if there could be any possible benefits. Assisting humanitarians
is a good benefit.


[1] "[Response 10]"
To support our allies because they provide safety and easier
to navigate , precision better than aircrafts
```

Similarly, the full text for the10 most **negative** responses are here:

```r
prnt_all_multi(
  get_pure(
    head(score_responses, 10) |>
      select(reason, score, index)
    )
  )
```

```
[1] "[Response 1]"
I considered the privacy issues involved. Also those of insurance.
And the implications of damaged caused by failing drones. And
furthermore the copyright infringements of drones flying over outdoor concerts
or other public forums.


[1] "[Response 2]"
I think as long as the items are not being
used for threats like terrorist attacks or anything else that
would bring harm, as long as they are selling these
devices and not giving them away for free like Biden
has been doing to the Ukraine


[1] "[Response 3]"
Well the us has problem with some of these countries
to sell them drones to strike in the near future
I would have to oppose because that's putting a threat
on are country


[1] "[Response 4]"
I considered the fact drones can't be seen as well
as a tank or helicopter and can pose serious harm
if used against us. Why arm the enemies of America
with deadly equipment to use against us or spy on
us.


[1] "[Response 5]"
I don’t support any of it. We are definitely fighting
a war that we don’t fully understand the depth of
destruction it will reap on us and generations and generations
to come. And for what? A few certain people to
```

have the most power and unlimited funds while they watch
their own people starve and suffer living in high violent
crime areas.

[1] "[Response 6]"
The surveillance state will soon turn into fascism. Authoritarianism is
rising everywhere. The next war is coming but we can
delay it if Ukraine wins.

[1] "[Response 7]"
The history of the conflict the government has been fighting.

[1] "[Response 8]"
It is really vague on where Turkey is selling the
drones and other weapons to. It is a really naive
assumption to make that once you sell something to someone
it won't wind up in another person's hands.

[1] "[Response 9]"
These countries are our friends and we should support them
Ukraine
was unjustly attacked and invaded by Russia and we should
do everything to support them in they're struggle against the
Russian invasion!

[1] "[Response 10]"
None at the moment bitch ass whore