Star Wars (4-6) Text Analysis of Movie Scripts

Contents

Dependencies Reading in data

```
library(tidyverse)
library(tidytext)
```

1   Tokenize and remove stop words

```
# sw_tokens <- sw_trilogy %>%
#   unnest_tokens(
#     word,
#     dialogue
#   ) %>%
#   anti_join(stop_words)
#
# sw_tokens
```

```
# lexicons <- c("bing", "afinn", "loughran", "nrc")
#
# df <- tibble(sentiments = lexicons %>% map(~left_join(sw_tokens, get_sentiments(.), by = "word")))
#
# for (lexicon in 1:nrow(df)){
#   print(sum(is.na(df[[1]][[lexicon]][5])))
# }
```