

# data\_analysis

Lindsey Greenhill

10/11/2021

*# turning corpus into tokens. Removing unimportant words and selecting for ngrams  
# = 2. Code taken from gov51 final project*

```
toks <- tokens(text_corpus,
               remove_punct = TRUE,
               remove_symbols = TRUE,
               remove_numbers = TRUE,
               remove_url = TRUE) %>%
tokens_tolower() %>%
tokens_remove(pattern=stopwords("en")) %>%
tokens_remove(pattern = c("u201c", "u00b0", "u2014", "wopi", "avi", "ooo",
                           "000", "ito", "ynl", "f'avl", "foxnews.com",
                           "ufb02pi", "ufb021", "ufbo2l", "ufb02L", "rpm")) %>%
tokens_select(min_nchar = 3) %>%
tokens_ngrams(n = 2)
```

```
wordcloud_dfm <- dfm(toks, groups = "channel")
textplot_wordcloud(wordcloud_dfm, comparison = T,
                   min_count = 200)
```

```
## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## president_trump could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## house_reporter could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## newspaper_shooting could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## subpoena_full could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## maryland_newspaper could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## pres_trump's could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## political_reporter could not be fit on page. It will not be plotted.
```

```

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## house_counsel could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## national_political could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## tariff_threat could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## record_new could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## judiciary_committee could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## epa_chief could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## deputy_director could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## phone_call could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## ntrump_claims could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## nsource_trump could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## memo_alleging could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## los_angeles could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## court_justice could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## ntrump_threatens could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## envoy_ukraine could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## trump's_ukraine could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## nhouse_dems could not be fit on page. It will not be plotted.

```

```

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## saudi_crown could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## capital_gazette could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## ntrump_calls could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## source_trump could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## biden_campaign could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## federal_prosecutor could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## trump_tower could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## elizabeth_warren could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## pressure_ukraine could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## russia_probe could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## house_speaker could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## ntrump_admin could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## poll_biden could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## ukraine_docs could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## reopen_schools could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## ukraine_documents could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## attorney_general could not be fit on page. It will not be plotted.

```

```

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## stone's_prison could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## trump_commuters could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## commutes_roger could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## roger_stone's could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## inspector_general could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## state_department could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## election_meddling could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## donald_trump could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## mask_mandate could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## shooting_suspect could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## threat_close could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## subpoena_mueller could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## poll_shows could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## release_gop could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## russia_investigation could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## coronavirus_cases could not be fit on page. It will not be plotted.

## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :
## bernie_sanders could not be fit on page. It will not be plotted.

```

