



Preprocessing text

Mees van Stiphout

11 july 2022

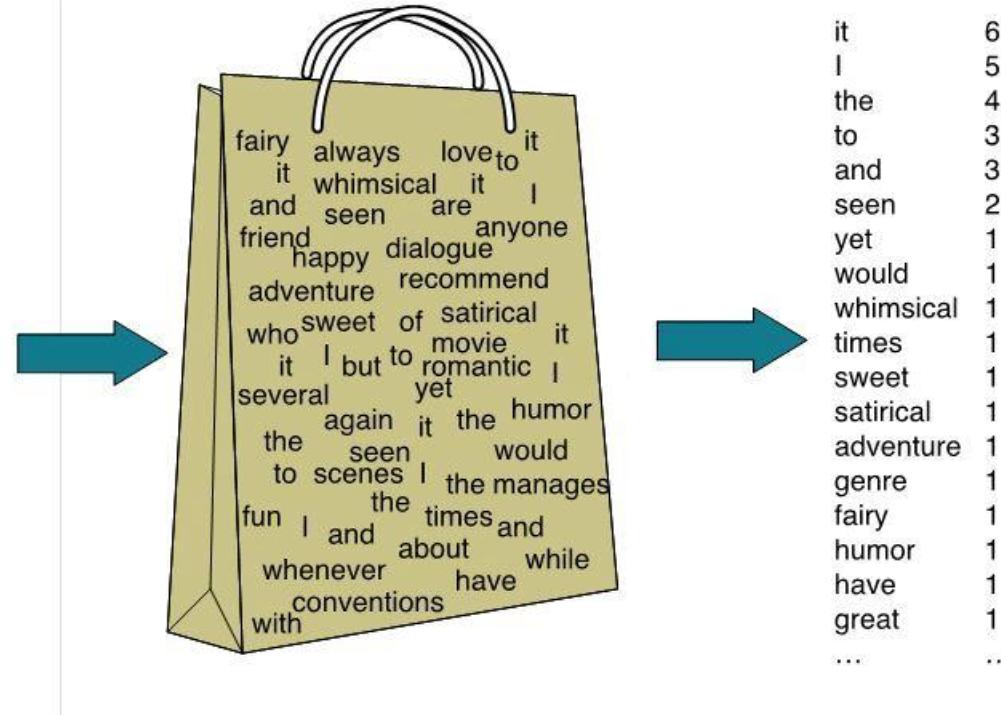
This afternoon

**Insight: Cleaning of data and text
is a vitally important skill set.**

Bag of Words

The Bag of Words Representation

I love this movie! It's sweet, but with satirical humor. The dialogue is great and the adventure scenes are fun... It manages to be whimsical and romantic while laughing at the conventions of the fairy tale genre. I would recommend it to just about anyone. I've seen it several times, and I'm always happy to see it again whenever I have a friend who hasn't seen it yet!



POStagging

Parts of Speech



Subjects

- Purposes of text cleaning
- Getting rid of noise and stop words
- Tokenizing and Ngrams
- Stemming and lemmatizing
- Language guesser
- POS tagging

Install or call relevant libraries/packages

```
library (cld2)          # for language detection
library (corpus)        # for text analysis with support for international text
library (dplyr)         # for data manipulation
library (entity)         # for easy NER
library (ggraph)         # for graphs
library (gridExtra)      # for working with grids to obtain nice layouts
library (hunspell)       # for high-performance stemming
library (igraph)         # for easy graphs
library (janitor)        # for a pretty table
library (lattice)         # for easy charts
library (knitr)          # for manipulating the output of this markdown document
library (plyr)           # for data manipulation
library (NLP)            # for natural language processing tools
library (openNLP)         # for named entity recognition
library (pander)          # for nice slide output
library (qdap)           # for using parsing tools to prepare transcript data
```

Point to the right working directory:

```
setwd("C:/Noodfolder")
```

Read CSV file:

- Articles on John Maynard Keynes taken from The Times newspaper
- `read.csv ("Keynes.csv")`

```
Keynes <- read.csv("Keynes.csv", stringsAsFactors = FALSE)
```

Dataframe Times articles

X	edition	date.pub	author	category	content	title
10	10	NA	February 1, 1937	['Arts and Entertainment']	The Arts Theatre of Cambridge, now in its second year tinde...	The Arts Theatre, Cambridge
70	70	NA	January 10, 1981	['Arts and Entertainment']	BBC 1 9.00 am Heads and Tails: Light- hearted view of the a...	Sunday's programmes
32	32	NA	August 27, 1969	['Business and Finance']	At the height of the Depression a meeting was engineered ...	A new maths lesson for the Treasury
62	62	NA	August 3, 1977	['Business and Finance']	Although he is not prepared to grumble about it, Quinton H...	Business Diary: Banging Brum's drum Â· CBI fire power
60	60	NA	December 24, 1976	['Business and Finance']	Gather round children, and Uncle Francis will tell you a love...	Once upon a time in a land on the never-never...
30	30	NA	December 9, 1968	['Business and Finance']	The gold crisis of March this year, which precipitated the tw...	The safety valve in a floating exchange rate
4	4	NA	February 1, 1923	['Business and Finance']	The N1lxMY-MIRD AxNUS.iL GmreRAL MEEWtfG of the Nati...	National Mutual Life Assurance Society.
5	5	NA	January 26, 1928	['Business and Finance']	The Wzsa -EaGHmT A-NtrAL GEcEBaL Mr.jirxo of the Nation...	Company Meetings.
45	45	NA	January 4, 1973	['Business and Finance']	Anthony Thomas discusses the immediate U S energy dilem...	Black outlook for American oil
31	31	NA	May 7, 1969	['Business and Finance']	The. One Hundred and Thirty- Ninth Annual General Meetin...	Company Meeting
28	28	NA	October 20, 1967	['Business and Finance']	A desperate plea for the coal mines The somewhat eccentric...	Business Diary
95	99	NA	October 9, 1985	['Business and Finance']	One of the most important themes to emerge from the ann...	Financial Services
1	1	NA	September 12, 1913	['Business and Finance']	The Royal Commission on Indian Finance and Currency yest...	Indian Currency Commission.
14	14	NA	September 19, 1941	['Business and Finance']	An announcement that Mr. John Maynard Keynes is to be pr...	Bank Of England Directorate
40	40	NA	September 27, 1971	['Business and Finance']	Bretton Woods and after What would the Founding Father...	Keynes and White on the dollar crisis
2	2	NA	September 29, 1919	['Business and Finance']	Mr. Pierce Lacy, of the firm of Cutler and Lacy, of Birmingham...	City News in Brief.
33	33	NA	September 9, 1970	['Business and Finance']	WILLIAM WALLACE is likely to be feeling a little un- easy ab...	BUSINESS diary
13	13	NA	September 19, 1941	['Classified Advertising']	3 lines 9s. (nininmu,n) RUTLAND GATE HOTEL, RUTLAND G...	London Hotels
74	74	NA	June 9, 1981	['Deaths']	DEATHS ASHTON.-On June 6th, 1981, peacefully at St Paul's ...	Deaths
43	43	NA	March 21, 1972	['Display Advertising']	In association with the Royal Economic Society Macmillan ar...	Macmillan

EXAMPLE TEXT FIELD

- Title: Mr. Roosevelt's Experiments
- Publication date: February 27, 1940

TEXT <- "HOW TO PAY FOR THE WAR. By JOHN MAYNARD KEYNES. Macmillan. Is.
In three articles published in The Times last November Mr. J. M. Keynes put forward what was expressly a first draft of proposals for compulsory savings in wartime. His plan now reappears in the form of a pamphlet in which Mr. Keynes elaborates his earlier argument and varies it to meet the comment and suggestion which it has so widely provoked. The pamphlet includes four appendices on the national income, the extent of our resources abroad, the cost of family allowances and the formula for the aggregate of deferred pay and direct taxes. The pamphlet is the subject of a leading article in today's issue"

EXAMPLE TEXT FIELD Tokenized

```
TEXTtokens <- str_split(TEXT, " ")  
TEXTtokens <- unlist(TEXTtokens)  
print (TEXTtokens [1:50])  
  
## [1] "HOW"          "TO"           "PAY"          "FOR"  
## [5] "THE"          "3WAR."        "By"           "JOHN\nMAYNARD"  
## [9] "KEYNES."       "Macmillan."    "Is.\n\nIn"      "three"  
## [13] "articles"     "published"     "i6n"          "Tlhe"  
## [17] "Times"         "last"          "November"    "Mr."  
## [21] "J."            "M.\nKeynes"   "put"          "forward"  
## [25] "what"          "was"           "expressly"    "a"  
## [29] "first"         "draft"         "of"           "proposals"  
## [33] "f6or\ncompulsory" "savings"      "in"           ""  
## [37] "wartime."      "3388"         "His"          "plan"  
## [41] ""              "now"          "reappears"    "in"  
## [45] "the"           "form\nof"      ""             "a"  
## [49] "pamphlet"      "in"
```

Wordcloud of tokens

```
wordcloud(TEXTtokens, scale = c(4,.5), min.freq =1, colors = brewer.pal(1, "Dark2"))
```



Clean text: all words in lower case

```
TEXT2 <- tolower(TEXT)
cat(TEXT2)

## how to pay for the 3war. by john
## maynard keynes. macmillan. is.
##
## in three articles published i6n tlhe times last november mr. j. m.
## keynes put forward what was expressly a first draft of proposals f6or
## compulsory savings in wartime. 3388 his plan now reappears in the form
## of a pamphlet in which mr. keynes elaborates his earlier argument and
## varies it to meet the comment and suggestion which it has so widely
## provoked.the pamphlet includes four appendices on the national income5the
## extent of our resources abroad the cost of family allowances and
## the formula
## for the aggregate of deferred pay and direct taxes
## the
## pamphlet
## is the subject of a leading article in 344 todays issue
```

Remove numbers

```
TEXT2 <- tm::removeNumbers(TEXT2)
cat(TEXT2)

## how to pay for the war. by john
## maynard keynes. macmillan. is.
##
## in three articles published in tlhe times last november mr. j. m.
## keynes put forward what was expressly a first draft of proposals for
## compulsory savings in wartime. his plan now reappears in the form
## of a pamphlet in which mr. keynes elaborates his earlier argument and
## varies it to meet the comment and suggestion which it has so widely
## provoked.the pamphlet includes four appendices on the national incomethe
## extent of our resources abroad the cost of family allowances and
## the formula
## for the aggregate of deferred pay and direct taxes
## the
## pamphlet
## is the subject of a leading article in todays issue
```

Remove double spaces

```
TEXT2<- str_replace_all(TEXT2, "  ", " ")
cat(TEXT2)

## how to pay for the war. by john
## maynard keynes. macmillan. is.
##
## in three articles published in tlhe times last november mr. j. m.
## keynes put forward what was expressly a first draft of proposals for
## compulsory savings in wartime. his plan now reappears in the form
## of a pamphlet in which mr. keynes elaborates his earlier argument and
## varies it to meet the comment and suggestion which it has so widely
## provoked.the pamphlet includes four appendices on the national incomethe
## extent of our resources abroad the cost of family allowances and
## the formula
## for the aggregate of deferred pay and direct taxes
## the
## pamphlet
## is the subject of a leading article in todays issue
```

Remove punctuation

```
TEXT2<-tm::removePunctuation(TEXT2)
cat(TEXT2)

## how to pay for the war by john
## maynard keynes macmillan is
##
## in three articles published in tlhe times last november mr j m
## keynes put forward what was expressly a first draft of proposals for
## compulsory savings in wartime his plan now reappears in the form
## of a pamphlet in which mr keynes elaborates his earlier argument and
## varies it to meet the comment and suggestion which it has so widely
## provokedthe pamphlet includes four appendices on the national incomethe
## extent of our resources abroad the cost of family allowances and
## the formula
## for the aggregate of deferred pay and direct taxes
## the
## pamphlet
## is the subject of a leading article in todays issue
```

Is the number of tokens reduced?

```
testtokens <- text_ntoken(TEXT)
print(testtokens)
```

```
## [1] 126
```

```
testtokens2 <- text_ntoken(TEXT2)
print(testtokens2)
```

```
## [1] 115
```

CLEAN THE TEXTS IN KEYNES DATFRAME

```
#All words in lower case
cleancontent <- tolower(Keynes$content)

## Remove numbers

cleancontent <- tm:::removeNumbers(cleancontent)

## Remove double spaces

cleancontent<- str_replace_all(cleancontent, " ", " ")

## Remove punctuation

cleancontent<-tm:::removePunctuation(cleancontent)

## Store results in new column in the dataframe

Keynes$clean_content <- cleancontent
```

Reduction in the number of tokens?

```
testtokens <- text_ntoken(Keynes$content)
print (sum(testtokens))

## [1] 141015

testtokens2 <- text_ntoken(Keynes$clean_content)
print (sum(testtokens2))

## [1] 116667
```

stopwords

Store all too common words in a variable

```
#Stopwordslist
```

```
Joseesstopwords <- c("the", "a", "an", "and", "i","in", "if", "he","she",
  "him", "should", "or", "we", "no", "over", "only",
  "could", "to", "of", "it", "is", "that", "this",
  "was", "/'s", "with","/'t", "as", "'t", "on", "are",
  "one", "so", "be","me", "are", "at", "s", "has",
  "by", "/", "for", "not", "from", "have", "which",
  "but", "his","than", "their", "were", "any", "some",
  "can", "what", "will", "would", "been", "per",
  "more", "they", "there", "tibble", "its", "had",
  "mrs", "mr", "when", "all", "our", "cent", "her",
  "who")
```

Remove these words:

```
Keynestext <- removeWords(Keynes$clean_content, Joseesstopwords)
#And store clean text in a new field
Keynes$clean_content <- Keynestext
```

Clean content in dataframe

content	clean_content
1 The Royal Commission on Indian Finance and Currency yesterday issued ...	royal commission indian finance currency yesterday issued interim re p...
2 Mr. Pierce Lacy, of the firm of Cutler and Lacy, of Birmingham, has joined ...	pierce lacy firm cutler lacy birmingham joined board united glass b...
3 Mr. John Maynard Keynes has won a reputation as a "serious" writer. He ...	john maynard keynes won reputation serious writer writes books whic...
4 The N1lxMY-MIRD AxNUs:iL GmreRAL MEEWtfG of the National Mutual L...	nixmymird axnusil gmral meewtfg national mutual life assurance soci...
5 The Wzsa -EaGHmT A-NtrAL GEcEBaL M.rjirxo of the National Mutual Life...	wzsa eaghmt antral gecebal mrjirxo national mutual life assur ance socie...
6 ESSAYS IN PERSUASION. By JOHN' MAYNARD KEYNES. (Macmillan. 10s. 6...	essays persuasion john maynard keynes macmillan los d net these great...
7 ESSAYS IN BIOGRAPHY. By John Maynard Keynes. (Macmillan. 7s. 6d. n.) I....	essays biography john maynard keynes macmillan d n impressioxns sout...
8 By John Maynard Keynes Mr. Roosevelt has made himself the trustee for t...	john maynard keynes roosevelt made himself trustee those every cou...
9 THE LEAGUE OF NATIONS AND THE RULE OF LAW, 1918-1935. By Alfred ...	league nations rule law alfred zimmern mac millan d battle ground ...
10 The Arts Theatre of Cambridge, now in its second year tinder the chairma...	arts theatre cambridge now second year tinder chairmanship john ma...
11 HOW TO PAY FOR THE WAR. By JOHN MAYNARD KEYNES. Macmillan. Is. I...	how pay war john maynard keynes macmillan three articles published ...
12 GREEK COASTAL SHIPPING ALL VESSELS REMOVED OR DESTROYED FRO...	greek coastal shipping vessels removed destroyed naval correspondent...
13 3 lines 9s. (,ninimnu,n) RUTLAND GATE HOTEL, RUTLAND GARDENS. S.W. ...	lines ninimnun rutland gate hotel rutland gardens sw sloane off park be...

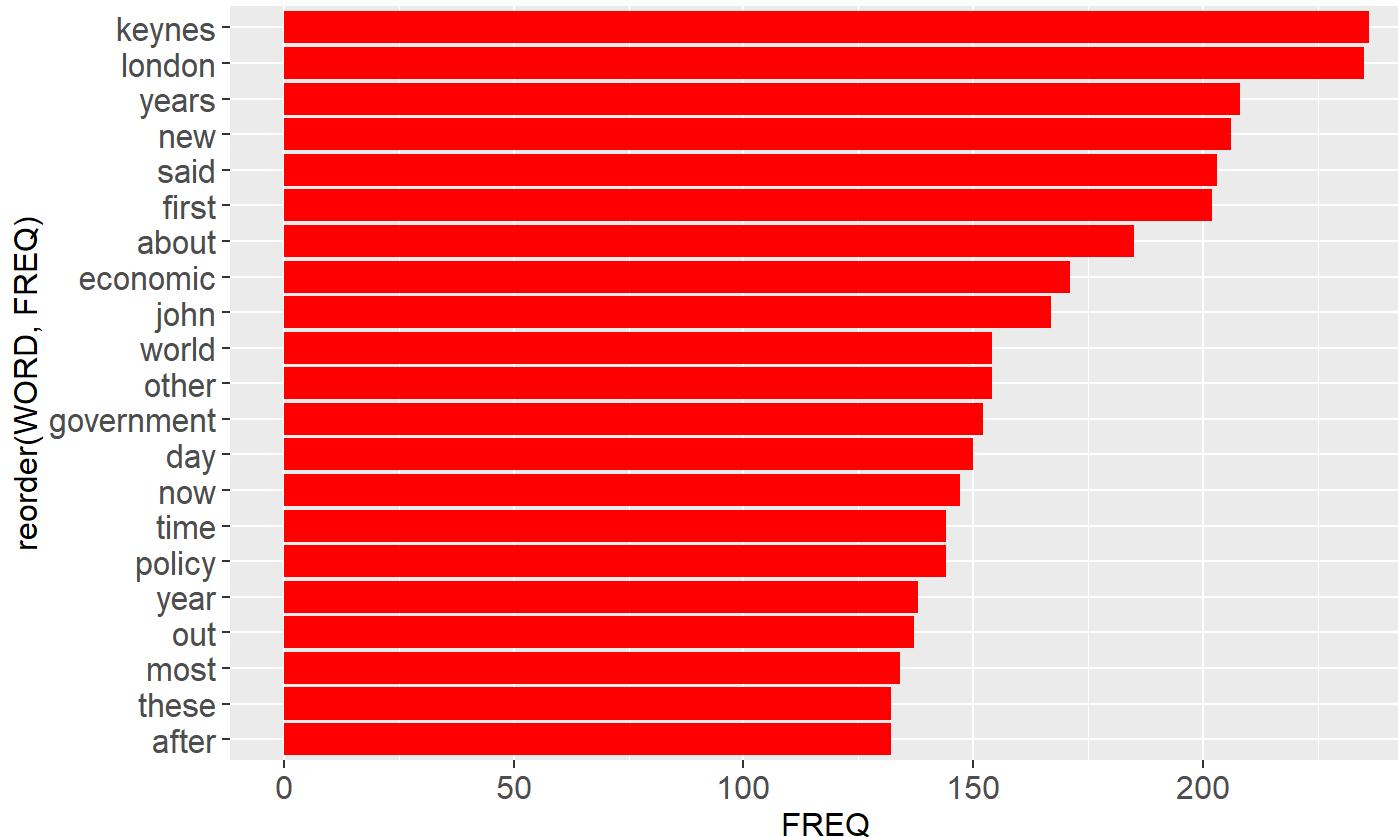
Original content

present now
many first economy
said work states maynard policy
london new government trade
much public may will general
way great con cent but born less long since
british even part well time bank
last sir world golding this must
1941 united tion year national
also people day inflation two war rate
keynes exchange years john
per economic life
money one can made the

Clean content

system economic
died government now
national second rates money
new british said years
also given office much rate even
less general present inflation september way first
life since year = like gold con india
two tax war bank trade policy long economy states
last ment exchange market great
may london claims currency must sir united
world keynes john day
monetary made people

Keynes Top 20 of (meaningful) words



STEMMING AND LEMMATIZATION

Stemming:

SnowballC library uses Porter's word stemming algorithm that collapses words to a common root. It supports many languages:

```
getStemLanguages()
```

```
## [1] "arabic"      "basque"       "catalan"      "danish"       "dutch"  
## [6] "english"     "finnish"      "french"       "german"       "greek"  
## [11] "hindi"        "hungarian"    "indonesian"   "irish"        "italian"  
## [16] "lithuanian"   "nepali"       "norwegian"    "porter"       "portuguese"  
## [21] "romanian"    "russian"      "spanish"      "swedish"     "tamil"  
## [26] "turkish"
```

Examples

```
wordStem("compulsary")
```

```
## [1] "compulsari"
```

```
wordStem("articles")
```

```
## [1] "articl"
```

```
wordStem("published")
```

```
## [1] "publish"
```

```
wordStem("November")
```

```
## [1] "Novemb"
```

Stem example text

```
## [1] "how to pay for the war by john\nmaynard keynes macmillan is\n\nin three articles published in tlhe tim
```

how to pay for the war by john maynard keynes macmillan Is in three articles published in the times last november mr j m keynes put forward what was expressly a first draft of proposals for compulsory savings in wartime

how to pay for the war by john maynard keyn macmillan is in three articl publish in the time last novemb mr j m keyn put forward what was expressli a first draft of propos for compulsori save in wartim

Lemmatization:

Lists on Github

<https://github.com/michmech/lemmatization-lists>

Make lemma dictionary

```
make_lemma_dictionary(TEXT2)
```

```
##          token      lemma
## 1         is        i
## 2    articles   article
## 3  published  publish
## 4     times     time
## 5  expressly  express
## 6 proposals proposal
## 7    savings    save
## 8      his     hi
## 9  reappears  appear
## 10 elaborates elaborate
## 11   earlier   early
## 12    varies    vary
## 13   comment    com
## 14      has     ha
## 15    widely    wide
## 16  includes  include
## 17 resources   source
## 18 allowances allowance
## 19    taxes     tax
## 20   leading    lead
```

Lemmatize (cleaned) example text

```
Lemmatext <- lemmatize_strings(TEXT2, dictionary = lexicon::hash_lemmas)
```

```
[1] "how to pay for the war by john maynard keynes macmillan be in  
three article publish in tthe time last november mr j be keynes put  
forward what be expressly a first draft of proposal for compulsory  
saving in wartime his plan now reappear in the form of a pamphlet in  
which mr keynes elaborate his early argument and vary it to meet the  
comment and suggestion which it have so widely provokedthe pamphlet  
include four appendix on the national incomethe extent of our resource  
abroad the cost of family allowance and the formula for the aggregate  
of defer pay and direct tax the pamphlet be the subject of a lead  
article in today issue"
```

Language detection

Several language detectors available

- Package cld = Compact Language Detector (Google)
- several versions (2, 3)

Use CLD2 to detect language in Good Reads reviews

```
GRreviewssc <- read.csv("CLDdemo2.csv", sep = ";")  
GRreviewssc$langdtcd <- detect_language(GRreviewssc$Text,  
plain_text = TRUE, lang_code = TRUE)
```

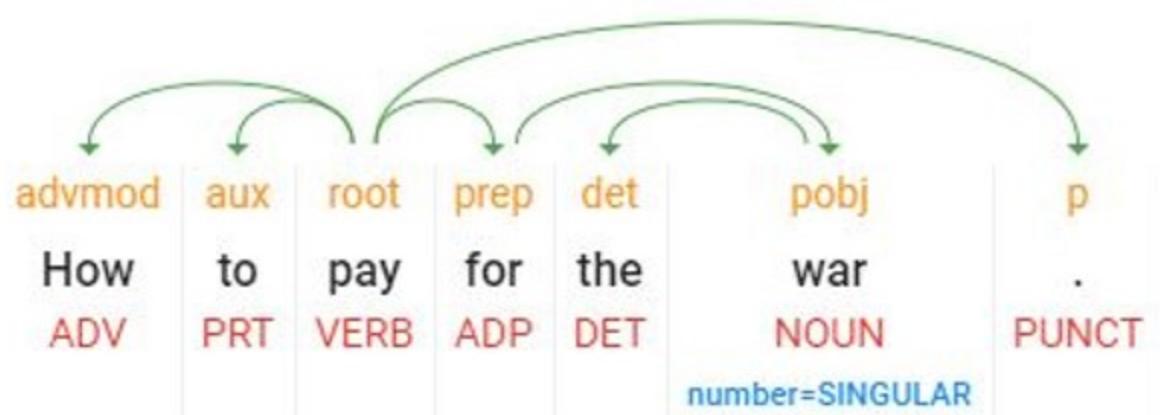
Result

	ID	Text	Language	langdtcd
50	7266	A bit sentimental, I still liked this book. It was simplistic but sweet.	NA	en
51	7267	Cerita'y sedih..tp ending'y bkin bingung dia bs reinkarnasi ^ ^	NA	id
52	7295	cerita yg sangat menyedihkan..betapa anjing merupakan hewan peliharaan...	NA	id
53	7310	ProÅŒitala sam dvije od trinaest priÅŒa i krenula na treÅŒu s najboljom na...	NA	hr
54	7314	Astonished at the poor language choice! Amazed that it won so many accl...	NA	en
55	7353	A beautiful book, but I never felt like underlining a sentence.	NA	en
56	7560	Kon me niet boeien.	NA	nl
57	7651	I really enjoy Tropper's cathartic writings.	NA	en
58	7652	Grappig, ontspannend, realistische familie dynamiek. Goede balans tussen ...	NA	nl
59	7769	An enjoyable read, somewhere between a beach book and a more serious ...	NA	en
60	7826	Ik vond het helaas langdradig en verwarrend	NA	nl
61	7837	Dwa z plusem, ale jednak dwa. Gdyby ta ksiÃ...Å¼ka byÅ,a sÅ,abiej ocenia...	NA	pl
62	7838	The Martian je prije svega propuÅŒtena prilika. PriÅŒa o astronautu koji je ...	NA	hr
63	7959	The Martian, romanzo di â€œfantascienzaâ€ di Andy Weir, Ã“ una strepit...	NA	it
64	7984	So so good! Now to see the film interpretation!	NA	en
65	7985	[Baca endingnya nggak jadi terharu, aku malah ngakak. Ya salahin Watney ...]	NA	id
66	7999	â€œI bought this on a drunken whimâ€ is a sentence that has immediate...	NA	en
67	8018	Admito no terminar el libro, solo me faltaron 40 pÃ¡ginas, pero me resultÃ...;	NA	es
68	8052	apa jadinya kalo punya boz paling kejam yang setiap permintaan gak ada ...	NA	id

POSTAGGING

Google Demo

Dependency Parse label Part of speech Lemma Morphology



Example text

```
## how to pay for the war by john
## maynard keynes macmillan is
##
## in three articles published in tlhe times last november mr j m
## keynes put forward what was expressly a first draft of proposals for
## compulsory savings in wartime his plan now reappears in the form
## of a pamphlet in which mr keynes elaborates his earlier argument and
## varies it to meet the comment and suggestion which it has so widely
## provokedthe pamphlet includes four appendices on the national incomethe
## extent of our resources abroad the cost of family allowances and
## the formula
## for the aggregate of deferred pay and direct taxes
## the
## pamphlet
## is the subject of a leading article in todays issue
```

how/WRB to/TO pay/VB for/IN the/DT war/NN by/IN john/NN
maynard/NN keynes/NNS macmillan/NN is/VBZ in/IN three/CD
articles/NNS published/VBN in/IN tlhe/DT times/NNS last/JJ
november/NN mr/NN j/NN m/NN keynes/NNS put/VBP
forward/RB what/WP was/VBD expressly/RB a/DT first/JJ
draft/NN of/IN proposals/NNS for/IN compulsory/JJ
savings/NNS in/IN wartime/NN his/PRP\$ plan/NN now/RB
reappears/VBZ in/IN the/DT form/NN of/IN a/DT pamphlet/NN
in/IN which/WDT mr/NN keynes/NNS elaborates/VBG his/PRP\$
earlier/JJR argument/NN and/CC varies/VBZ it/PRP to/TO
meet/VB the/DT comment/NN and/CC suggestion/NN
which/WDT it/PRP has/VBZ so/RB widely/RB provokedthe/VBN
pamphlet/NN includes/VBZ four/CD appendices/NNS on/IN
the/DT national/JJ incomethe/NN extent/NN of/IN our/PRP\$
resources/NNS abroad/RB the/DT cost/NN of/IN family/NN
allowances/NNS and/CC the/DT formula/NN for/IN the/DT
aggregate/JJ of/IN deferred/JJ pay/NN and/CC direct/JJ
taxes/NNS the/DT pamphlet/NN is/VBZ the/DT subject/NN
of/IN

Tagging of content field of Keynes dataframe (package: Udpipe)

```
ud_model <- udpipe_download_model(language = "english")
```

Part of speech tagging of the clean content field of the Keynes dataframe.

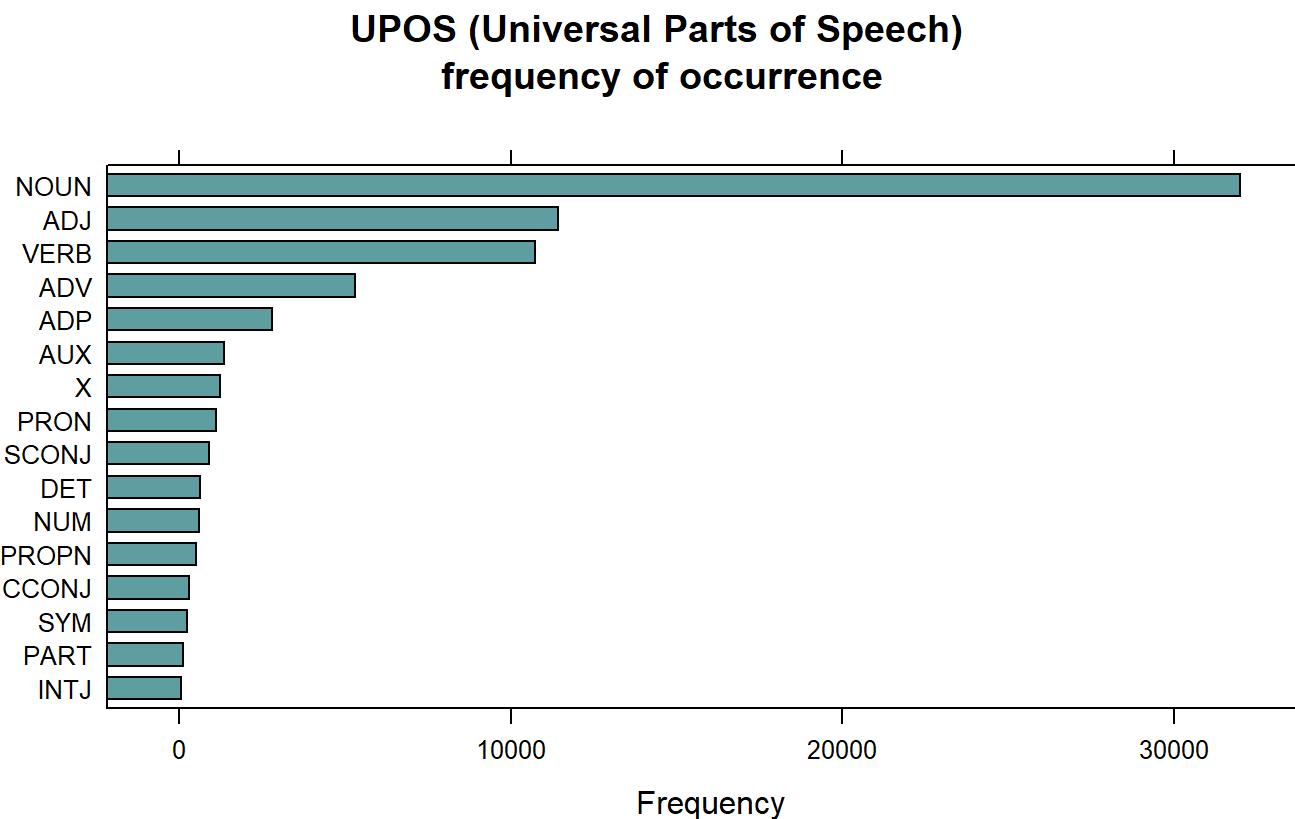
```
ud_english <- udpipe_load_model(ud_model$file_model)
cleanfieldPOS <- udpipe_annotate(ud_english, x = Keynes$clean_content, doc_id = Keynes$date.pub)
cleanfieldPOS <- as.data.frame(cleanfieldPOS)
```

This is what the resulting dataframe looks like:

▲	doc_id	paragraph_id	sentence_id	sentence	token_id	token	lemma	upos
1	September 12, 1913	1	1	The Royal Commission on Indian Finance and Currency yest...	1	The	the	DET
2	September 12, 1913	1	1	The Royal Commission on Indian Finance and Currency yest...	2	Royal	Royal	PROPN
3	September 12, 1913	1	1	The Royal Commission on Indian Finance and Currency yest...	3	Commission	Commission	PROPN
4	September 12, 1913	1	1	The Royal Commission on Indian Finance and Currency yest...	4	on	on	ADP
5	September 12, 1913	1	1	The Royal Commission on Indian Finance and Currency yest...	5	Indian	indian	ADJ
6	September 12, 1913	1	1	The Royal Commission on Indian Finance and Currency yest...	6	Finance	Finance	NOUN
7	September 12, 1913	1	1	The Royal Commission on Indian Finance and Currency yest...	7	and	and	CCONJ
8	September 12, 1913	1	1	The Royal Commission on Indian Finance and Currency yest...	8	Currency	currency	NOUN
9	September 12, 1913	1	1	The Royal Commission on Indian Finance and Currency yest...	9	yesterday	yesterday	NOUN
10	September 12, 1913	1	1	The Royal Commission on Indian Finance and Currency yest...	10	issued	issue	VERB

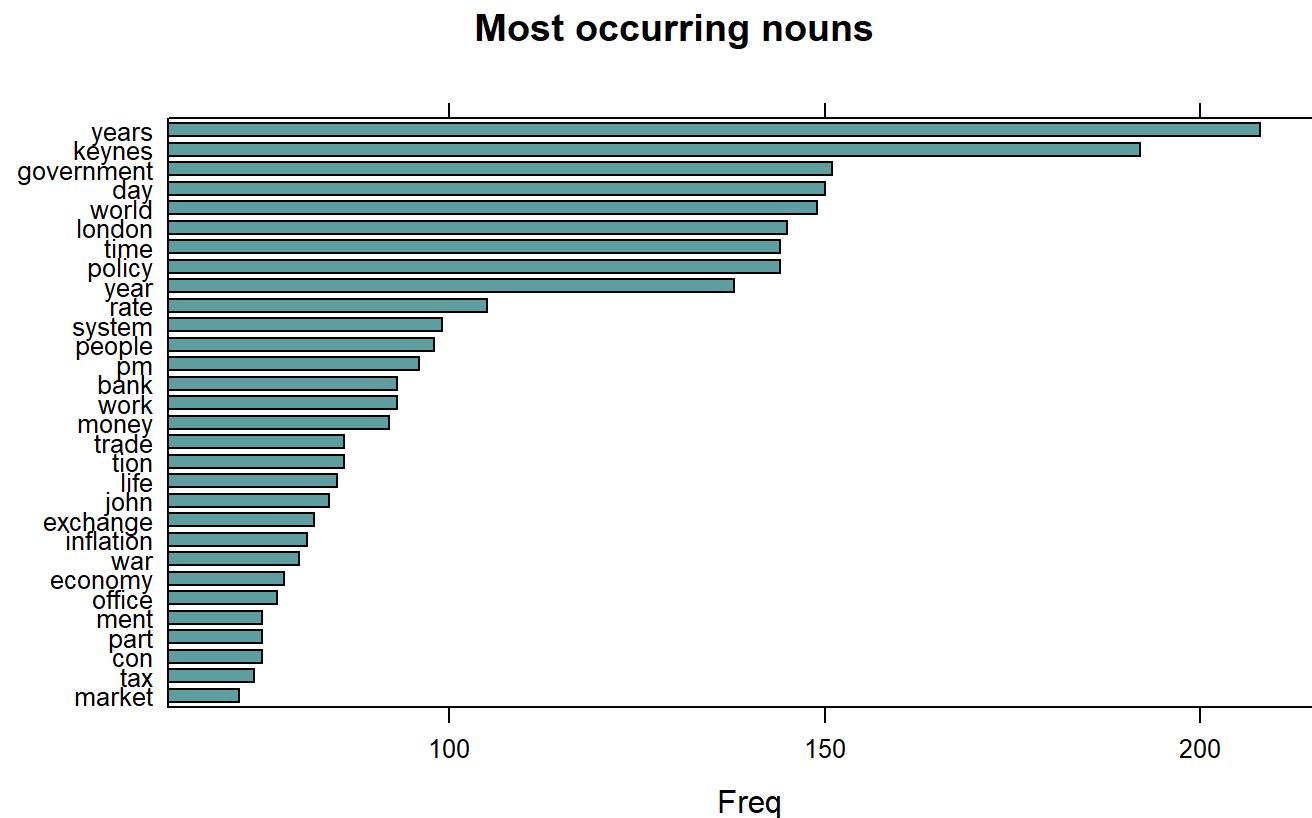
Results barchart

```
statsclean <- txt_freq(cleanfieldPOS$upos)
statsclean$key <- factor(statsclean$key, levels = rev(statsclean$key))
barchart(key ~ freq, data = statsclean, col = "cadetblue",
main = "UPOS (Universal Parts of Speech)\n frequency of occurrence",
xlab = "Frequency")
```

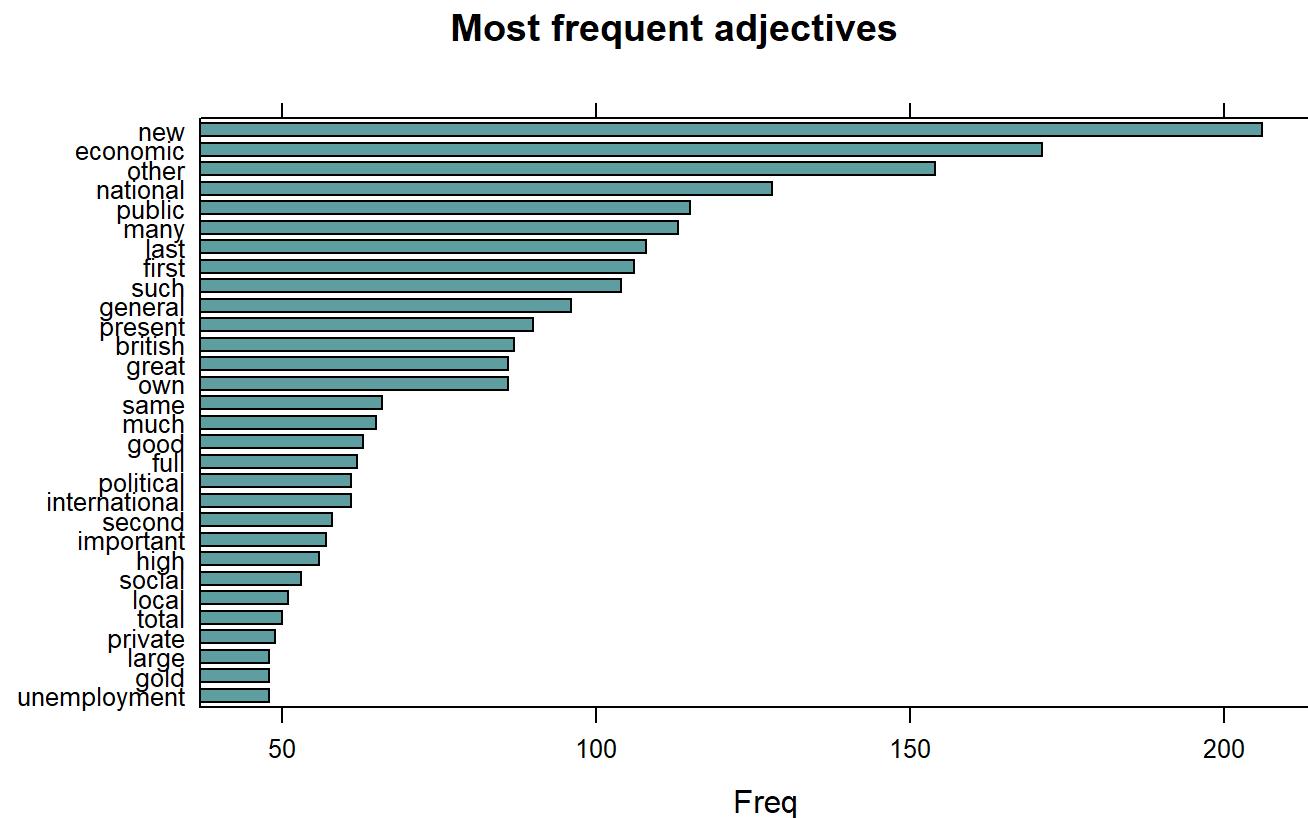


Which nouns are most important?

```
nounstat <- subset(cleanfieldPOS, upos %in% c("NOUN"))  
nounstat <- txt_freq(nounstat$token)
```



And adjectives:

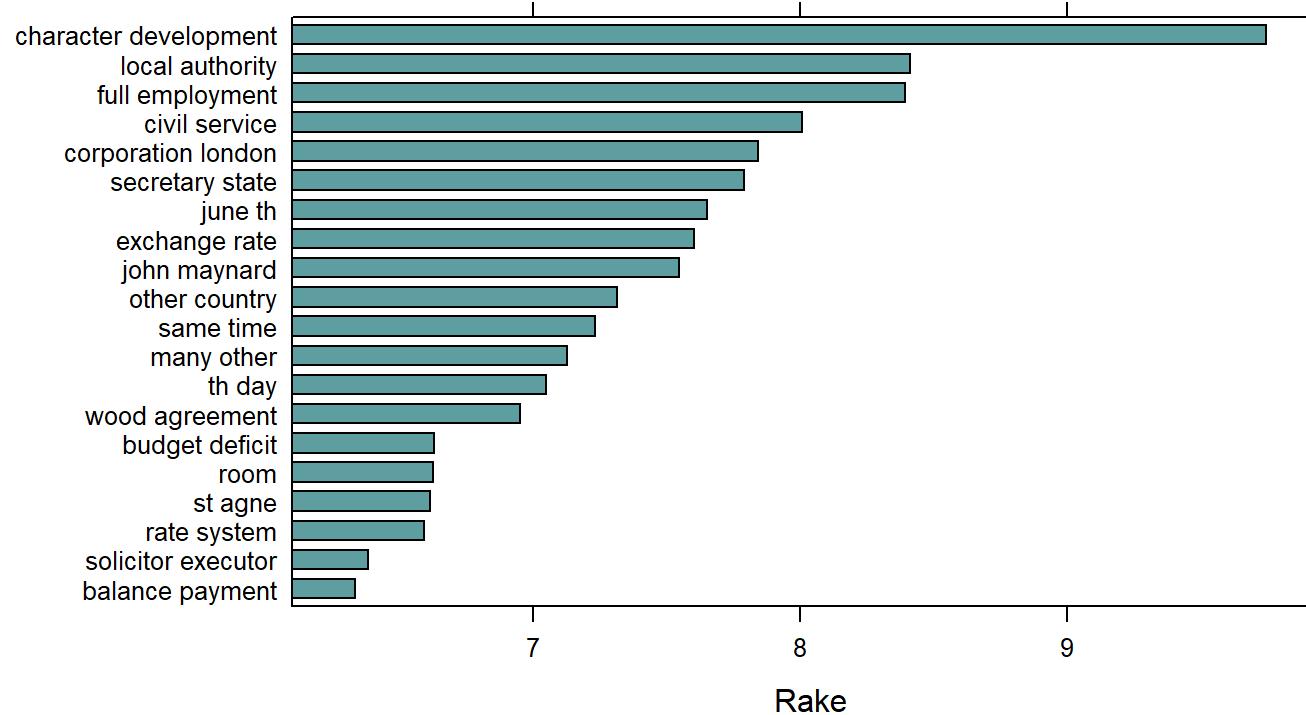


Keyword combinations in text

Identifying keywords with RAKE

```
keyrake <- keywords_rake(x = cleanfieldPOS, term = "lemma", group = "doc_id",  
                           relevant = cleanfieldPOS$upos %in% c("NOUN", "ADJ"))
```

Keywords identified by RAKE



NER = Named Entity Recognition

Tags for NER

Correct NER tagging requires a model for the language used. Wirtschafts universität in Vienna offers models on this website: <https://datacube.wu.ac.at/>.

The Wikipedia text on the first election debate with Biden and Trump

```
President <- readLines('Debat.txt')
President <- as.String(President)
```

Wrapper around library

Open NLP needs the file to be tokenized, either in words or in sentences and the result will be stored as a list.

```
sent_ann <- Maxent_Sent_Token_Annotator()
```

Annotated text to replace the original

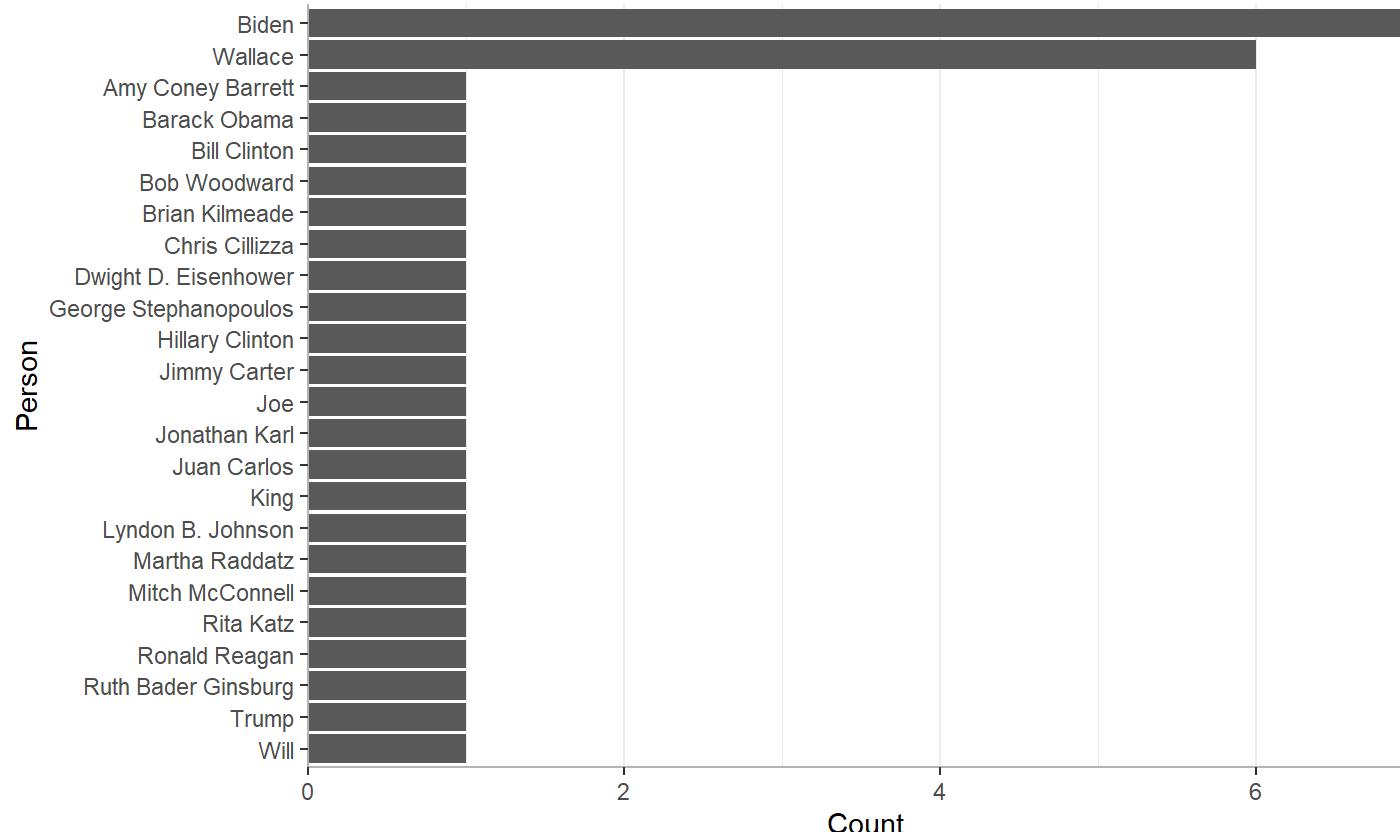
```
President_sent <- NLP::annotate(President, list(sent_ann))
```

```
President_nmbrd <- President[President_sent]
```

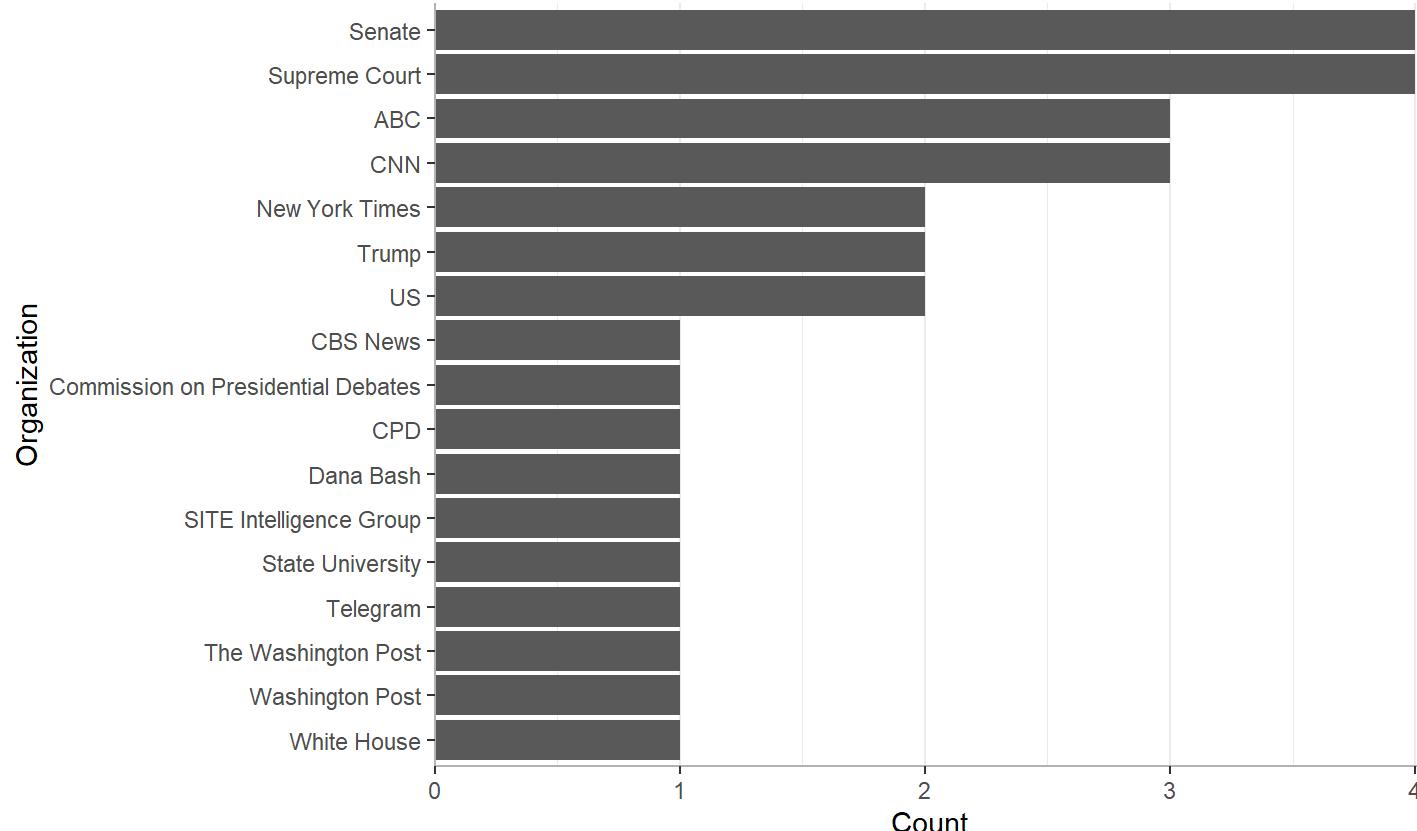
```
President_nmbrd
```

```
## [1] "President <- \"Entering into the debate, Biden had a significant and persistent lead in the polls."  
## [2] "Biden's lead was compounded by a funding shortage in Trump's campaign, with Biden's campaign donations  
## [3] "Since Biden's successful nomination in the Democratic primaries, Trump had attempted to cast doubt over  
## [4] "Trump called for Biden to be drug tested before the debate."  
## [5] "Biden mocked the idea.[20]"  
## [6] "Trump also claimed that Biden would use a hidden electronic earpiece for the debate, demanding that B  
## [7] "Again, Biden declined."  
## [8] "Running up to the debate, Trump made repeated claims that the election would be rigged by means of vot  
## [9] "When asked if he would commit to a peaceful transition of power, Trump said, we'll have to wait and see.  
## [10] "In the weeks leading up to the debate, Trump became part of various controversies."  
## [11] "Bob Woodward released his second book on the Trump presidency, based on 19 recorded interviews with T  
## [12] "In one recording made in February 2020, Trump indicated that he understood the severity of the COVID-19  
## [13] "Trump confirmed that he downplayed the severity of the pandemic, saying that I don't want to create a  
## [14] "The New York Times published an investigation into Trump's federal tax returns, which found that Trump  
## [15] "Additionally, they reported that his businesses lost money in most years."  
## [16] "A few days before the debate, the US reached the milestone of 200,000 deaths from COVID-19."  
## [17] "This number represented 20% of worldwide fatalities, despite the US having only 4% of the world's pop  
## [18] "Two weeks prior to the debate, Supreme Court Justice Ruth Bader Ginsburg died from cancer."  
## [19] "Ginsburg was one of four Supreme Court justices who are commonly considered liberal; the other five ju  
## [20] "The day after Ginsberg's funeral, Trump nominated conservative Amy Coney Barrett."  
## [21] "Senate Republicans, under the leadership of Senate Majority Leader Mitch McConnell, moved swiftly, pre  
## [22] "The move was controversial, since the same Senate Republicans had refused to consider a Supreme Court  
## [23] "Format and debate\nWill you shut up, man? redirects here."  
## [24] "For Why don't you shut up?, the phrase uttered by King Juan Carlos I of Spain to Hugo Chávez, see ¿Po  
## [25] "The debate was divided into six segments: Trump's and Biden's records, the Supreme Court, the COVID-19  
## [26] "Each was approximately 15 minutes in length; Wallace introduced each topic and gave each candidate two  
## [27] "The allotted time was generally not upheld; Trump repeatedly interrupted and criticized Biden during  
## [28] "On several occasions, Wallace pleaded with Trump to respect the rules and norms of the debate."
```

Persons mentioned in the text



And how about organizations?



Questions?

Practical