Pre-Processing Text

Text as Data

Amber Boydstun & Cory Struthers

Spring 2025

Contents

Introduction

After data has been gathered or collected, we load the text data with the corresponding metadata. We then review patterns. Patterns in the text can provide useful context about the types of analysis choices that may be (in)appropriate based on the data generation process (DGP) that we understand was responsible for the data as we receive them.

In this module, we'll need the following packages:

```r
# Load packages
require(tidyverse)
require(readxl)
require(quanteda)
library(quanteda.textstats)

# Set working directory
setwd("~/Library/CloudStorage/Box-Box/A_teaching/POL290G_TextAsData/modules/data/")
getwd() # view working directory
```

First we upload the raw data, which sometimes includes metadata about the text itself (e.g., the speaker of the text, source of the text, date).

```r
# Load data and with doc_id_keep
news_data <- read_xlsx("~/Library/CloudStorage/Box-Box/A_teaching/POL290G_TextAsData/modules/data/sampl

# Examine metadata (i.e., the variable names)
names(news_data)
```

```
## [1] "source_id"   "text"        "Article"     "Source"      "Year"
## [6] "title"       "section_sum" "section"     "keyword"
```

```r
# Read text
news_data$text[1]
```

```
## [1] "Border Fence Blocks Bears In Migration, Study Finds\r\r\n\r\r\nGILA BEND, Ariz. -- The much-bal
```

```r
# Check out source
news_data$Source[1]
```

```
## [1] "new york times"
```

Create a corpus

A corpus object saves text (character strings in a format that can be analyzed with text-as-data tools and pairs text with document-level variables (metadata) that tell us something about the text. Almost always, we're interested in understanding differences in the text according to one or more document-level variables. These metadata variables are called "docvars" (`docvars` in `quanteda`, which is a comprehensive text-as-data package in R).

We typically create a corpus from a data file that includes both text and docvars. Every corpus includes a "doc_id", which is either assigned by default in the format "text#" or by the researcher using the `docid_field` option in the `corpus` function.

```r
# Create the corpus and look at it
news_corp <- corpus(news_data, text_field = "text")
print(news_corp)
```

```
## Corpus consisting of 1,000 documents and 8 docvars.
## text1 :
## "Border Fence Blocks Bears In Migration, Study Finds GILA BEN..."
##
## text2 :
## "New Philly tour highlights historical sites for naturalizati..."
##
## text3 :
## "The one sure immigration fix What to do about the nation's 1..."
##
## text4 :
## "The honorable brother David Kaczynski, brother of the Unabom..."
##
## text5 :
## "O'Malley's record on immigration praised by fellow Democrats..."
##
## text6 :
## "ACROSS THE USA: NEWS FROM EVERY STATE Alabama Opelika - Anto..."
##
## [ reached max_ndoc ... 994 more documents ]
    # Confirming: 1,000 documents with 8 docvars
```

We can extract metadata using `docvars`, which stores the metadata in the columns of the text file we transformed to a corpus (so we can look at it). This is useful in a workflow where the corpus has been saved as an RDS file and we no longer rely on the original dataframe.

We can access docvars using the `$` operator, like we would for a dataframe.

```r
# View unique sources in bills corpus
unique(news_corp$Source)
```

```
##  [1] "new york times"          "philadelphia inquirer"
##  [3] "st. petersburg times"    "washington post"
##  [5] "usa today"               "tampa bay times"
##  [7] "daily news"              "st. louis post-dispatch"
##  [9] "saint paul pioneer press" "atlanta journal and constitution"
```

```
## [11] "palm beach post"                    "san jose mercury news"
## [13] "denver post"                         "herald-sun"
```
```r
# Store only the metadata (docvars)
metadata_news <- docvars(news_corp)
```

We can create additional docvars in the corpus as well.

```r
# Create a docvar for the corpus (and restore the metadata if we want), in this case as a character str
news_corp$newscountry <- as.character("United States")
head(news_corp$newscountry, 10)
```
```
##  [1] "United States" "United States" "United States" "United States"
##  [5] "United States" "United States" "United States" "United States"
##  [9] "United States" "United States"
```
```r
# Re-store only the metadata with this new docvar
metadata_news <- docvars(news_corp)
```

There are a number of ways to transform the corpus. First, we can subset the corpus based on docvars.

```r
# Subset to Washington Post & NY Times only
news_corp_wp_nyt <- corpus_subset(news_corp, Source %in% c("new york times", "washington post"))
ndoc(news_corp_wp_nyt)
```
```
## [1] 391
```
```r
# At this point (and any other point) we can make a dataframe if we want to view our data
news_df <- convert(news_corp_wp_nyt, to = "data.frame")
news_df$text[1]
```
```
## [1] "Border Fence Blocks Bears In Migration, Study Finds\n\n\n\nGILA BEND, Ariz. -- The much-ballyho
```

Second, we can reshape the corpus into different units: documents, paragraphs, and sentences.

```r
# Reshape to paragraphs
news_corp_wp_nyt_paragraphs <- corpus_reshape(news_corp_wp_nyt, to = "paragraphs")
ndoc(news_corp_wp_nyt_paragraphs)
```
```
## [1] 4709
```
```r
# Make sure the full text vs. paragraph reshape worked
news_df_para <- convert(news_corp_wp_nyt_paragraphs, to = "data.frame")
news_df_para$text[1]
```
```
## [1] "Border Fence Blocks Bears In Migration, Study Finds"
```

Third, we can split up segments of text based on *patterns*, including exact matches or regular expressions. For example, segmenting is often needed to split a document based on a speaker, in the case of a transcript. This is a really powerful feature of `quanteda` because it does the string splitting for you. All docvars are retained during these transformations.

```r
# Split at periods (into sentences) -- an example!
# "regex" for regular expressions; "fixed" for exact matching
news_corp_sentences <- corpus_segment(news_corp, pattern = ".", valuetype = "fixed",
                         extract_pattern = TRUE, pattern_position = "before")
ndoc(news_corp_sentences)
```
```
## [1] 13800
```
```r
head(news_corp_sentences)
```

```
## Corpus consisting of 6 documents and 10 docvars.
## text1.1 :
## "-- The much-ballyhooed border fence has not just made it mor..."
##
## text1.2 :
## "It has also become an obstacle, researchers say, for migrati..."
##
## text1.3 :
## "A study published in this month's edition of Biological Cons..."
##
## text1.4 :
## "Also fragmenting the bear habitat are the growing urban spra..."
##
## text1.5 :
## "Researchers used hair snags -- pieces of barbed wire set up ..."
##
## text1.6 :
## "They found significant genetic disparities between black bea..."
```

The corpus is now ready for analysis. Like an R dataframe, we can save the corpus as an RDS object to avoid repeating these steps in a workflow.

```
saveRDS(news_corp, "news_corp.RDS")
```

Tokenize corpus and reduce complexity

Most text analysis methods require tokenization of the corpus. Tokenization breaks a corpus into vectors of words, n-grams, or sentences. From a programming and English-language perspective, tokenization is simply splitting the text up by white space.

Tokenization is one of the most critical steps in pre-processing because the researcher makes a number of decisions that will affect analysis outcomes. Many of these decisions are based on pre-existing dictionaries that other researchers have developed (e.g., what counts as a stopword), and so it's important to examine those dictionaries to make sure we understand the decisions we're really making.

Tokenization decisions include:

- **Removing punctuation, URLs, symbols, numbers**: For many research questions, researchers are often not interested in how many commas, apostrophes, periods, symbols, URLS, or numbers appear within the text, and thus choose to delete them.

- **Lowercasing**: Replace all capital letters with lowercase letters. The idea is that the word "This" is no different than "this" for most projects and research questions.

- **Removing stopwords**: Stop words are common words used across documents that do not give much information about the task at hand. In English, common words such as "and", "the", and "that" may be removed from the documents to reduce the size and complexity of the feature set.

- **Generating n-grams**: We can incorporate elements of positional analysis by identifying *n-grams* or multi-word expressions (phrases). Analysts can identify specific n-grams to capture in a tokens document or convert the whole tokens object to n-grams (n-grams of 2 and 3 are most common). N-grams are important for bag-of-word approaches that would otherwise ignore the conceptual link between compound words like "death" and "penalty". Collocation analysis can be used to identify common n-grams in the text.

- **Stemming**: Many words and phrases are conjugated but hold the same meaning. For example, in most contexts, the word "immigrant" has the same relevance as "immigrants" or "immigration". Asterisks match zero or more characters in `quanteda`. For the concept "immigrant", for example, we can use "immig*". We can also use `tokens_wordstem` in `quanteda` to stem all words.
- **Maintaining document length**: Sometimes we need to retain information on terms (tokens) we remove, for example when applying methods that utilize information on word ordering. `quanteda` offers the `padding` option, which is defaulted to `FALSE`. When true, an empty string will remain where removed tokens previously existed.

`quanteda` provides all the functions we need to reduce the complexity of our text data following the above steps, using variations of `tokens()` from `quanteda package`.

```r
# Tokenize with no adjustments
toks_news <- tokens(news_corp)
print(toks_news[1:2])
```

```
## Tokens consisting of 2 documents and 9 docvars.
## text1 :
##  [1] "Border"    "Fence"     "Blocks"    "Bears"     "In"        "Migration"
##  [7] ","         "Study"     "Finds"     "GILA"      "BEND"      ","
## [ ... and 253 more ]
##
## text2 :
##  [1] "New"         "Philly"        "tour"          "highlights"
##  [5] "historical"  "sites"         "for"           "naturalization"
##  [9] "exam"        "With"          "their"         "collars"
## [ ... and 289 more ]
```

```r
# Use tokens_remove to drop stopwords
toks_news_subs <- tokens_remove(toks_news, stopwords("english"))
print(toks_news_subs[1:2])
```

```
## Tokens consisting of 2 documents and 9 docvars.
## text1 :
##  [1] "Border"    "Fence"     "Blocks"    "Bears"     "Migration" ","
##  [7] "Study"     "Finds"     "GILA"      "BEND"      ","         "Ariz"
## [ ... and 167 more ]
##
## text2 :
##  [1] "New"         "Philly"        "tour"            "highlights"
##  [5] "historical"  "sites"         "naturalization"  "exam"
##  [9] "collars"     "cold"          ","               "25"
## [ ... and 192 more ]
```

```r
# Stem words
toks_news_subs <- tokens_wordstem(toks_news_subs)

# We can instead pipe functions in one ordered step
toks_news_subs <- tokens(news_corp,
                         remove_url = TRUE,
                         remove_punct = TRUE,
                         remove_symbols = TRUE,
                         remove_numbers = TRUE) %>%
    tokens_select(pattern = stopwords("en"), selection = "remove", padding=TRUE) %>% # add padding to p
    tokens_tolower() %>% # note to_lower here
    tokens_wordstem()
```

```r
print(toks_news_subs[1:2])
```

```
## Tokens consisting of 2 documents and 9 docvars.
## text1 :
##  [1] "border" "fenc"   "block"  "bear"   ""       "migrat" "studi"  "find"
##  [9] "gila"   "bend"   "ariz"   ""
## [ ... and 225 more ]
##
## text2 :
##  [1] "new"       "philli"    "tour"      "highlight" "histor"    "site"
##  [7] ""          "natur"     "exam"      ""          ""          "collar"
## [ ... and 232 more ]
```

Oftentimes, words only meaningfully represent topics, sentiment, or other latent concepts when put together. For example, "death" on its own has an entirely different meaning than "death penalty". These phrases are called "n-grams" or "compound tokens" and must be included in the process of tokenization or will otherwise not be captured in the subsequent steps of analysis.

There are two approaches to dealing with phrases. One option is using `tokens_ngrams()`, where the user specifies the number of words that should be included in each token. This approach transforms the entire corpus into phrases.

```r
# Convert tokens to n-grams using first approach
toks_news_ngrams <- tokens_ngrams(toks_news_subs, n = 2:3) # specifying 2 and 3 n-grams
head(toks_news_ngrams[[1]], 15)
```

```
##  [1] "border_fenc"          "fenc_block"          "block_bear"
##  [4] "migrat_studi"         "studi_find"          "find_gila"
##  [7] "gila_bend"            "bend_ariz"           "much-ballyhoo_border"
## [10] "border_fenc"          "just_made"           "illeg_immigr"
## [13] "slip_across"          "unit_state"          "also_becom"
```

```r
head(toks_news_ngrams[[2]], 15)
```

```
##  [1] "new_philli"           "philli_tour"         "tour_highlight"
##  [4] "highlight_histor"     "histor_site"         "natur_exam"
##  [7] "peopl_includ"         "includ_new"          "new_immigr"
## [10] "immigr_troop"         "troop_behind"        "color_guid"
## [13] "independ_mall"        "mall_friday"         "friday_philadelphia"
```

While both efficient and useful when the user is uncertain about the scope of meaningful phrases, the approach also returns a very large tokens object without necessarily adding much information.

Alternatively, the user can apply `tokens_compund` to capture n-grams (certain phrases) more selectively.

```r
# Use token_compound to add n-grams selectively - remember to use stem terms if you've stemmed words
toks_news_comp <- tokens_compound(toks_news_subs,
                            pattern = phrase(c("climat *", "death penalt",
                                               "border fenc", "histor site")))
print(toks_news_comp[1:2])
```

```
## Tokens consisting of 2 documents and 9 docvars.
## text1 :
##  [1] "border_fenc"    "block"          "bear"           ""
##  [5] "migrat"         "studi"          "find"           "gila"
##  [9] "bend"           "ariz"           ""               "much-ballyhoo"
## [ ... and 223 more ]
##
```

```
## text2 :
##  [1] "new"         "philli"      "tour"        "highlight"   "histor_site"
##  [6] ""            "natur"       "exam"        ""            ""
## [11] "collar"      ""
## [ ... and 231 more ]
```

Finally, how might we identify important compound tokens when we don't know what they are?

Collocation analysis, which identifies words that are commonly seen together in a text, is an efficient way to locate common multi-word expressions.

Analysts can search for collocations with certain attributes, such as proper nouns (with capitalized words), or search for collocations more liberally. `textstat_collocations` can be used on tokens or corpus objects.

Let's try it first using our tokens with no pre-processing.

```
# Find all collocations in original tokens object
news_corp_coll <- textstat_collocations(toks_news, size = 2:3, min_count = 10)
news_corp_coll
```

```
##                        collocation count count_nested length     lambda
## 1                        more than   232          232      2  6.229968046
## 2                           in the  1281         1281      2  1.871461815
## 3                           of the  1444         1444      2  1.721097699
## 4                    united states   270          270      2  8.985432455
## 5                illegal immigrants   141          139      2  6.638607340
## 6                        have been   161          160      2  4.767438848
## 7                   climate change   126          122      2  8.288082144
## 8                          has been   144          144      2  4.895715075
## 9                          will be   156          156      2  4.113774740
## 10                       last year    95           87      2  5.746323573
## 11                   global warming   114          109      2  9.372864202
## 12                      gun control   101          100      2  6.938991703
## 13                       would be   147          147      2  3.884830608
## 14                    death penalty   241          240      2  9.033880960
## 15                       last week    78           69      2  6.281759410
## 16                     gay marriage    73           71      2  6.604963498
## 17                same-sex marriage    73           72      2  6.635591925
## 18                        he said   157           66      2  3.478148952
## 19                      should be   102          102      2  4.887959424
## 20                          to be   316          316      2  2.475406731
## 21                       years ago    61           61      2  7.253737849
## 22                         it was   138          137      2  3.385347744
## 23                         such as    88           88      2  5.091877082
## 24                          it is   163          161      2  3.044314188
## 25                     high school    49           49      2  7.157751842
## 26                    supreme court   160          155      2  8.664671102
## 27                tobacco companies    57           57      2  6.089085110
## 28                 attorney general    45           45      2  7.454940274
## 29                       had been    74           74      2  4.567451175
## 30                         did not    78           78      2  5.819343659
## 31                same-sex couples    44           44      2  6.903597283
## 32                         i don't    61           60      2  5.611165035
## 33                      white house    51           51      2  7.532027793
## 34                 second amendment    41           41      2  7.640231828
## 35                          one of   187          187      2  3.043398783
## 36                       this year    72           66      2  4.468592346
```

```
## 37                 capital punishment      55       55      2  9.333390138
## 38                        those who        66       66      2  4.646889462
## 39                       last month        49       47      2  6.084576755
## 40                   greenhouse gas        42       42      2  8.598693771
## 41                          said he        99       99      2  3.537074896
## 42                  tobacco industry       48       48      2  5.980287762
## 43                background checks        50       50      2  9.928753564
## 44                  law enforcement        40       39      2  6.201352656
## 45                        they were        79       79      2  3.703114337
## 46                           at the       327      327      2  1.931768455
## 47                         they are        95       94      2  3.309682453
## 48                           on the       435      435      2  1.604634057
## 49                              ` `        42       42      2  9.442007656
## 50                  greenhouse gases       46       44      2  9.575065468
## 51                     gas emissions       33       33      2  7.405583281
## 52                       smoking ban        41       41      2  5.236185068
## 53            naturalization service       34       34      2  9.062411060
## 54                          this is       108      108      2  3.016176199
## 55                          he was       105      105      2  3.010284754
## 56                       percent of       129      129      2  3.507726541
## 57                        this week        50       45      2  4.730010209
## 58                    death sentence       42       42      2  6.077964593
## 59                       people who        58       58      2  4.057639675
## 60                        the state       237      237      2  2.404131960
## 61                          i think        42       42      2  5.502659013
## 62                         there is        74       74      2  3.792510414
## 63                 rifle association       37       37      2  9.700664159
## 64                       the united       270      270      2  4.483740094
## 65              justice department        27       27      2  6.689125795
## 66                        next week        30       27      2  5.962017694
## 67                             as a       209      209      2  2.068166934
## 68                          that he       119      119      2  2.761384968
## 69                 president clinton       33       32      2  6.511383728
## 70                        and other       129      129      2  2.855650258
## 71                         gun shows       37       37      2  6.130094863
## 72                 concealed weapons       25       25      2  7.988929206
## 73                 immigration reform       46       46      2  6.728490666
## 74              illegal immigration       45       44      2  4.466552823
## 75                    district judge       29       29      2  5.854161469
## 76          undocumented immigrants       37       37      2  6.367895274
## 77                       lung cancer       26       25      2  8.947512407
## 78                  background check       24       24      2  8.353353375
## 79                         she said       58       20      2  3.904740709
## 80                      power plants       24       24      2  8.281231716
## 81                           do not       50       50      2  4.208351890
## 82                           is not       106      106      2  2.760835224
## 83                  president obama        29       29      2  6.109819627
## 84                        less than        33       33      2  5.661236198
## 85                         does not        46       46      2  5.989772945
## 86                           to get       113      113      2  3.388829652
## 87                        there are        52       51      2  3.986143580
## 88                        the first       171      170      2  3.266451503
## 89                      civil rights       24       24      2  6.202220588
## 90                       gay couples       25       25      2  5.965897977
```

```
## 91               as well    51    51   2  5.432941912
## 92                by the    298   298  2  1.663627702
## 93         national rifle    41    41   2  8.489684638
## 94              gun laws    33    32   2  4.973597609
## 95              could be    55    55   2  3.707290133
## 96             would have   75    75   2  3.078676742
## 97               can be    52    52   2  3.853869225
## 98               for the   429   429  2  1.313344039
## 99             don't think  23    23   2  5.984467627
## 100            health care  22    22   2  6.814263345
## 101        news conference  22    22   2  8.179500082
## 102                 if you  33    33   2  4.740204433
## 103                we have  60    60   2  3.401710444
## 104         north carolina  29    28   2  9.883630020
## 105     executive director  18    18   2  7.371397405
## 106             gun owners  30    30   2  6.206420277
## 107      same-sex marriages 21    21   2  6.817397977
## 108           gun violence  29    29   2  6.109163823
## 109             that would  106   106  2  2.509128400
## 110              the u.s    169   142  2  2.465578455
## 111          public health  26    26   2  5.242536112
## 112            debate over  26    26   2  5.433466122
## 113         president bush  24    24   2  5.964708089
## 114           donald trump  22    19   2  8.844954271
## 115               to make   95    95   2  3.557325726
## 116             each year   25    23   2  5.668239746
## 117          because they   40    40   2  4.060191902
## 118           city council  21    21   2  6.213319981
## 119            don't want   22    22   2  5.667741135
## 120          officials said 45    32   2  3.914549296
## 121             their own   32    32   2  4.911178468
## 122               well as   39    39   2  5.537947227
## 123        carry concealed  16    16   2  7.463133085
## 124            civil unions 27    27   2  8.991093635
## 125      federal government 27    27   2  4.972071783
## 126                 to do   100   100  2  2.900109275
## 127           charged with  40    40   2  5.661861152
## 128     customs enforcement 29    29   2  9.185344217
## 129             the death   191   191  2  2.041872317
## 130               at least  101   99   2  7.677401590
## 131             on tuesday  47    37   2  4.057606328
## 132                 in a    373   373  2  1.250440780
## 133        assault weapons  18    18   2  8.144123654
## 134          proposition 8  15    14   2  8.854969863
## 135           appeals court  26    26   2  5.269639310
## 136             five years   24    24   2  5.473619326
## 137         executive order  16    16   2  6.749471731
## 138                if they   42    42   2  3.721161309
## 139            three years   27    27   2  4.754352346
## 140             both sides   28    27   2  8.503075131
## 141             next month   19    17   2  5.809654086
## 142             civil union  16    15   2  6.718676926
## 143           border patrol  36    36   2  9.073563428
## 144        associated press 18    17   2  9.940234972
```

9

```
## 145                    this month   31    28   2  4.606597080
## 146    environmental protection   14    14   2  7.798344498
## 147                       part of   95    95   2  4.763049342
## 148              carbon emissions   19    19   2  5.753123070
## 149                        ban on   48    48   2  3.676064815
## 150                  public safety   19    19   2  6.646106003
## 151                    first time   27    27   2  4.598382915
## 152                      said they   60    60   2  3.013432502
## 153                        want to   82    82   2  3.919778203
## 154                       right to   81    81   2  3.446175986
## 155                        may be   40    40   2  3.852778807
## 156      constitutional amendment   15    15   2  7.131321599
## 157                   the country  126   126   2  3.014579340
## 158                       when he   41    41   2  3.637604800
## 159                  was convicted   35    35   2  4.439070663
## 160                majority leader   13    13   2  8.245538264
## 161                  this country   34    34   2  4.059762704
## 162              district attorney   18    18   2  5.718763816
## 163                    earlier this   25    25   2  5.667799147
## 164               bush administration   16    16   2  6.140438598
## 165                     along with   32    32   2  5.207809065
## 166                    long island   14    14   2  7.289975392
## 167                      it would   61    61   2  2.870561086
## 168                    row inmates   13    13   2  7.255465020
## 169                family members   16    16   2  6.025146941
## 170                         a new  119   119   2  2.148538671
## 171                     two years   30    30   2  4.125332026
## 172               supreme court's   16    16   2  6.764893251
## 173                       going to   86    86   2  4.555589193
## 174              general assembly   14    14   2  7.461982532
## 175                social security   14    14   2  6.709128345
## 176                  on wednesday   40    36   2  3.987199425
## 177                     two weeks   19    19   2  5.657013627
## 178                       you can   25    25   2  4.470637811
## 179                   capitol hill   14    14   2  9.997724508
## 180                      from the  232   232   2  1.526681724
## 181                       its own   21    21   2  5.050294196
## 182               without parole   17    17   2  7.787339712
## 183                 district court   24    24   2  4.648262630
## 184                    rather than   29    28   2  7.134690917
## 185              country illegally   14    14   2  6.411790711
## 186                        with a  154   154   2  1.763123890
## 187                  federal judge   24    24   2  4.497263601
## 188                       we are   46    46   2  3.179053437
## 189                  told reporters   16    15   2  7.605170279
## 190          undocumented workers   14    14   2  6.235877007
## 191                    don't know   16    15   2  5.654070903
## 192                   is expected   36    36   2  4.483632710
## 193                      over the  120   120   2  2.289972362
## 194              illegal immigrant   20    17   2  4.915169965
## 195                  young people   20    20   2  5.289546657
## 196                      how many   21    21   2  4.720816952
## 197                    green card   11    11   2  8.892575356
## 198                florida supreme   15    15   2  5.766543942
```

```
## 199              the national  107      107      2  2.897322926
## 200              department of   64       64      2  3.396044950
## 201           dioxide emissions   14       14      2  6.857786353
## 202              vice president   35       35      2  8.526573556
## 203                  six months   13       13      2  6.264183385
## 204              nation's largest  12       12      2  6.701955499
## 205              decide whether   13       13      2  6.667915830
## 206               illegal aliens   20       19      2  7.329289609
## 207                found guilty   13       13      2  6.643707324
## 208                      out of   84       84      2  2.494673750
## 209            immigration laws   24       24      2  4.386362620
## 210               mental illness   11       11      2  9.314547432
## 211                     we need   21       21      2  4.793615844
## 212            protection agency   11       11      2  7.186580377
## 213                     al gore   15       15      2 10.718365300
## 214                     must be   27       27      2  4.463677726
## 215              mexican border   14       12      2  5.980457952
## 216                      it's not   32       32      2  3.711160449
## 217            tobacco products   20       20      2  6.705296227
## 218                     are not   56       56      2  2.727116936
## 219                  even though   27       26      2  8.537496439
## 220                   york city   17       17      2  5.144300874
## 221                      so far   15       11      2  5.869756161
## 222             lethal injection   29       29      2 12.300153048
## 223                   death row   94       94      2  9.309338478
## 224                    with the  237      237      2  1.389584105
## 225             prison without   13       13      2  5.985591168
## 226                    10 years   19       19      2  4.885931627
## 227                   bear arms   16       16      2 11.208343441
## 228                    known as   27       26      2  5.644303281
## 229             united nations   16       16      2  5.513185356
## 230                  three days   14       14      2  5.675568509
## 231                 court ruling   20       20      2  4.665336564
## 232               tuesday night   11       10      2  6.541891652
## 233           tobacco settlement   17       17      2  5.311150717
## 234                    san jose   17       17      2  8.868939787
## 235                   he would   54       54      2  2.696815885
## 236                police chief   12       12      2  6.328710983
## 237                circuit court   18       18      2  5.991815002
## 238                   might be   26       26      2  4.277260387
## 239                   george w   10        6      2  9.114553544
## 240                    who are   55       55      2  2.645953752
## 241                    bay area   10       10      2  8.405062046
## 242                   weeks ago   11       10      2  6.611804167
## 243                  members of   56       56      2  3.426953436
## 244                    know how   14       14      2  5.503579754
## 245           cigarette makers   15       15      2  8.458095035
## 246               electric chair   13       13      2 10.882518782
## 247                    25 years   16       16      2  5.882541736
## 248                    the most  108      108      2  2.351017011
## 249                 natural gas   10        9      2  7.735282845
## 250                    20 years   16       16      2  5.535294957
## 251                    based on   31       31      2  5.088709111
## 252           marriage licenses   13       13      2  5.853503260
```

```
## 253             likely to   60    60   2  4.124150266
## 254             an effort   22    22   2  4.779407843
## 255          state supreme   21    21   2  4.345673724
## 256              1 million   13    13   2  6.029869022
## 257          immigration and   83    83   2  2.185095584
## 258          central america   10     9   2  6.928749494
## 259          prime minister   10    10   2 10.102431850
## 260              next year   18    18   2  4.652225022
## 261               say they   24    24   2  4.017767081
## 262          capital murder   13    11   2  5.520195721
## 263                  is no   42    42   2  3.059901142
## 264          lesbian couples   12    12   2  7.769648833
## 265             on thursday   32    26   2  3.775275497
## 266          president obama's   14    14   2  6.238704334
## 267          innocent people   17    17   2  6.262855718
## 268             legal status   13    12   2  5.766299052
## 269             being held   13    13   2  5.618500294
## 270              on monday   33    25   2  3.647499590
## 271                to take   60    60   2  3.047428903
## 272          human services   10    10   2  6.471786905
## 273             director of   54    54   2  4.125394883
## 274             court ruled   16    16   2  5.299657268
## 275                to help   58    58   2  3.180648793
## 276              number of   75    75   2  5.164971117
## 277            human rights   12    12   2  5.826918692
## 278              away from   22    22   2  4.893694949
## 279             high court   19    19   2  4.524242203
## 280           circuit judge   12    12   2  6.247420126
## 281               into the  103   103   2  2.150543022
## 282              when they   31    31   2  3.390140673
## 283          court's decision   11    11   2  6.331040961
## 284              new jersey   49    49   2  8.028315528
## 285              that they   72    72   2  2.241655557
## 286          first-degree murder   19    16   2  8.724497952
## 287              four years   17    16   2  4.737891824
## 288                look at   21    21   2  5.055394221
## 289               how much   15    15   2  4.937277817
## 290                who had   38    38   2  3.030962959
## 291               said she   27    27   2  3.773358630
## 292          on immigration   55    54   2  2.537757845
## 293             court judge   21    21   2  4.118392941
## 294             expected to   58    58   2  4.317666059
## 295          border security   13    13   2  5.330004068
## 296          married couples   10    10   2  6.578583916
## 297          driver's licenses   11    11   2  9.022624740
## 298             support for   35    35   2  3.415442358
## 299             has become   21    21   2  4.363597068
## 300           check system   10    10   2  6.526142620
## 301             more likely   19    19   2  4.497851399
## 302              would not   45    45   2  2.732796147
## 303          clinton administration   11    11   2  5.779448218
## 304          mentally retarded   12    12   2 10.509965355
## 305              the senate   83    82   2  2.826606615
## 306          president trump   14    14   2  5.139894402
```

```
## 307                because of   70     70   2  2.377483686
## 308                  50 years   14     14   2  5.484194601
## 309           states illegally   13     13   2  5.453053946
## 310                    to keep   53     53   2  4.010603262
## 311             police officer   10     10   2  6.901263531
## 312                    as much   27     27   2  3.706234145
## 313                  city hall   11     11   2  6.473639969
## 314               three times   12     12   2  5.406680671
## 315                     $ 100   13     13   2  5.771976558
## 316                  are being   27     27   2  3.555387076
## 317                    he did   26     26   2  3.681846644
## 318                11 million   12     12   2  5.507504863
## 319               do something  12     12   2  5.581746411
## 320                  that the  375    375   2  0.954721048
## 321                 months ago  10     10   2  5.936381041
## 322                her husband  13     12   2  6.734384031
## 323                 be allowed  20     20   2  4.708838523
## 324                   his wife  20     20   2  5.735572265
## 325                 there were  24     24   2  3.643355300
## 326                       at a 124    124   2  1.615051092
## 327                   so many   17     17   2  4.324313181
## 328                   he says   25     14   2  3.708416666
## 329                  web site   10     10   2 10.335621147
## 330                    he has   48     48   2  2.541335603
## 331            tobacco company  14     14   2  5.056370133
## 332                 in prison   41     41   2  3.438022810
## 333              immigrants who  27     27   2  3.402647787
## 334                     a few   43     43   2  3.769096643
## 335                 make sure   10     10   2  6.201133640
## 336                 on friday   28     25   2  3.666769956
## 337               sentenced to  62     62   2  4.867272199
## 338                 no matter   12     12   2  5.957841118
## 339              convicted of   46     46   2  3.263652274
## 340                   says he   20     19   2  4.022607048
## 341     republican presidential  10     10   2  5.836495607
## 342             federal appeals  15     15   2  4.623263210
## 343                     for a  191    191   2  1.266641380
## 344                 any other   19     19   2  4.021426933
## 345                 no longer   28     28   2  8.444094456
## 346              fossil fuels   12     12   2 11.122320667
## 347                  may have   27     27   2  3.416012264
## 348                  so much   14     14   2  4.707792353
## 349                   as part   24     24   2  3.777221432
## 350                   he had   38     38   2  2.811053366
## 351        obama administration  10     10   2  5.642987363
## 352                   vote on   27     27   2  3.627241204
## 353                   need to   47     47   2  3.230401282
## 354                  which is   46     46   2  2.559387015
## 355                 ruled that  28     28   2  4.342522849
## 356                take effect  10     10   2  5.659600121
## 357                  the bill  105    102   2  1.953140547
## 358               gay marriages  10     10   2  5.733809944
## 359           homeland security  40     40   2 10.806616438
## 360               teen smoking  12     12   2  6.253472623
```

```
## 361          cigarette smoke   11   11   2  5.269232458
## 362                   way to   50   50   2  2.913392107
## 363                     is a  164  164   2  1.343379220
## 364            oklahoma city   12   12   2  7.303443896
## 365            public places   10   10   2  6.058674043
## 366                 in which   42   42   2  3.059390131
## 367                 to carry   46   46   2  3.905120460
## 368                  who was   50   50   2  2.394532541
## 369               york times   10   10   2  5.531058769
## 370             an estimated   17   17   2  5.138655119
## 371           their families   16   16   2  4.618532104
## 372             recent years   14   14   2  4.714170857
## 373            life sentence   10   10   2  5.517064498
## 374                every day   11   10   2  5.218790968
## 375                  his own   20   20   2  3.981832506
## 376                  even if   15   15   2  4.393031587
## 377                    but i   25   25   2  3.433538176
## 378                  make it   23   23   2  3.593152032
## 379            san francisco   33   31   2 10.689410501
## 380           said yesterday   22   22   2  3.768117398
## 381                there was   29   29   2  3.144069905
## 382           federal courts   12   12   2  5.161733272
## 383                    do so   14   14   2  4.503672584
## 384                  life in   39   39   2  3.038885372
## 385             barack obama   15   12   2  9.582310898
## 386                  two men   13   13   2  4.738448482
## 387           spokesman for   24   24   2  4.488302366
## 388     enforcement officials   11   11   2  5.106803294
## 389                  a major   39   39   2  3.408481108
## 390               that's what   12   12   2  4.932065464
## 391                 focus on   25   25   2  5.462020126
## 392               smoking in   55   55   2  2.352548189
## 393                 who came   19   19   2  4.004664837
## 394                right now   13    9   2  4.615056847
## 395           million illegal   15   15   2  4.272028014
## 396                day after   14   14   2  4.477681069
## 397                know what   12   12   2  4.878342898
## 398                there's no   10    9   2  5.692717753
## 399                    if he   27   27   2  3.151183579
## 400                    a man   43   43   2  2.929430927
## 401                caused by   19   19   2  5.558369710
## 402              ban smoking   16   16   2  4.115492226
## 403            talking about   14   14   2  6.370877312
## 404            university of   40   40   2  3.329017771
## 405             months after   12   12   2  4.930534055
## 406                 not only   23   23   2  3.448497233
## 407                 it comes   17   17   2  4.557566272
## 408             other states   23   23   2  3.376455486
## 409                gun lobby   12   11   2  6.120274424
## 410                   but it   43   43   2  2.457207955
## 411                state law   24   24   2  3.292514368
## 412             in california   38   37   2  3.059246274
## 413                 light up   11   11   2  5.905188142
## 414         and naturalization   37   37   2  4.384695039
```

14

```
## 415            find out  12  12  2  4.861742513
## 416            who have  46  46  2  2.362969998
## 417            law firm  10  10  2  6.139578311
## 418          across the  63  63  2  4.151572664
## 419       kyoto protocol  13  13  2 11.375402115
## 420       county district  11  11  2  4.913858697
## 421             won't be  16  16  2  4.921234327
## 422              be able  16  16  2  5.065097185
## 423           40 percent  10  10  2  5.521798582
## 424       federal district  13  13  2  4.499974467
## 425           into effect  11  11  2  4.982846718
## 426            what they  24  24  2  3.226091076
## 427            same time  12  12  2  4.639393031
## 428           continue to  40  40  2  4.036328840
## 429            efforts to  39  39  2  3.840540496
## 430         gun ownership  15  15  2  7.073108252
## 431          would allow  16  16  2  4.150258237
## 432             year ago  10  10  2  5.014647532
## 433        what happened  10  10  2  5.973049409
## 434             of those  55  55  2  2.302533009
## 435             fact that  24  24  2  4.677616897
## 436              she was  27  27  2  3.024305710
## 437                  $ 2  10  10  2  5.344602964
## 438              we want  15  15  2  4.070097163
## 439              to find  38  38  2  3.579391265
## 440          years later  10  10  2  5.064769482
## 441             least one  13  13  2  4.395038076
## 442           their lives  13  13  2  4.785623725
## 443          most recent  11  10  2  4.786162264
## 444      legal immigrants  17  17  2  3.762643158
## 445              what he  24  24  2  3.156195975
## 446           they would  37  37  2  2.533805986
## 447              most of  59  59  2  2.165875520
## 448                if we  17  17  2  3.753273268
## 449           around the  51  51  2  3.233676240
## 450               a year  61  61  2  2.096062030
## 451            the issue  74  74  2  2.195141262
## 452            the right  72  72  2  2.248235212
## 453             is being  29  29  2  2.963301119
## 454              when it  28  28  2  2.906325369
## 455        death sentences  11  11  2  5.209945298
## 456                do it  23  23  2  3.225603781
## 457           the state's  83  83  2  4.625871595
## 458       concerned about  11  11  2  5.681237712
## 459                 $ 10  12  12  2  4.500559015
## 460            talk about  11  11  2  5.743762670
## 461            they have  43  43  2  2.320315727
## 462            could not  25  25  2  3.058476815
## 463          should have  25  25  2  3.073291426
## 464          marriage act  10  10  2  4.865415964
## 465          was arrested  16  16  2  4.292085770
## 466            found that  29  29  2  3.030969497
## 467                 $ 50  10  10  2  5.076293005
## 468      reynolds tobacco  10  10  2  6.130348335
```

```
## 469                  this case    21      21    2  3.318638768
## 470               is scheduled    19      19    2  4.124883014
## 471                     not be    43      43    2  2.296984442
## 472                  they will    34      34    2  2.581898402
## 473                  after the   102     102    2  1.632083946
## 474                  effort to    36      36    2  3.886301924
## 475                federal law    19      19    2  3.450433893
## 476                    but he     35      35    2  2.539476716
## 477                  the world    56      56    2  3.024956454
## 478              was sentenced    17      17    2  3.958920877
## 479                 where they    18      18    2  3.570693202
## 480                be executed    15      15    2  4.345705946
## 481                       $ 1    10      10    2  4.923792230
## 482                   the case    82      82    2  1.922470938
## 483                   15 years    10      10    2  5.015571846
## 484                for example    21       0    2  4.860351089
## 485         federal authorities    10      10    2  4.917692188
## 486                 lawyers for    23      23    2  3.387370431
## 487                    as they    39      39    2  2.369172192
## 488                 might have    17      16    2  3.772163589
## 489               an interview    13      13    2  5.169758373
## 490          immigrant children    10      10    2  4.694513259
## 491                   wants to    35      35    2  4.001494963
## 492               will continue    13      13    2  4.359938702
## 493                   not just    20      20    2  3.336523829
## 494                 days after    10      10    2  4.841336002
## 495              would require    12      12    2  4.966470901
## 496                   the same   109     109    2  5.326212568
## 497                  gun buyers    10      10    2  6.192739672
## 498              carbon dioxide    54      52    2 11.798057335
## 499                  that could    40      40    2  2.382706859
## 500             state attorney    14      14    2  3.963160453
## 501                  i didn't    12      12    2  4.432108115
## 502                 argued that    23      23    2  5.055468929
## 503                    plan to    34      34    2  3.292445242
## 504                  was found    20      20    2  3.363990310
## 505                  gun rights    14      14    2  3.915978690
## 506                    what we    15      15    2  3.747649244
## 507                    i have    33      33    2  2.524177077
## 508                  to reduce    34      34    2  3.898750393
## 509              believe that    22      22    2  3.572238116
## 510                  the house    74      74    2  2.000609807
## 511                   of them    59      58    2  2.006742388
## 512                   the past    55      55    2  3.740015694
## 513                  they had    28      28    2  2.726532840
## 514                 the second    55      54    2  2.641221430
## 515              global climate    11      11    2  4.354631066
## 516                     to go    38      38    2  2.802369409
## 517                  said that    74      74    2  1.680755532
## 518                    in this    71      71    2  1.755703009
## 519                    that it    67      67    2  1.768358072
## 520             not immediately    12      12    2  5.144786767
## 521                 allowed to    33      33    2  3.943222825
## 522                  were not    30      30    2  2.614132574
```

16

```
## 523            state department   14   14   2  3.873619252
## 524                    no one   17   16   2  3.462563093
## 525                  in april   27   24   2  3.576844586
## 526               this summer   11   11   2  5.052534275
## 527                big tobacco   10   10   2  4.695079875
## 528                    to see   35   35   2  2.986922751
## 529                   gun show   11   11   2  4.424949190
## 530                  only one   15   15   2  3.669032093
## 531                   it will   35   35   2  2.399773832
## 532                the border   66   66   2  2.115949158
## 533                    set up   11   11   2  4.354291820
## 534                  to death   81   81   2  1.648803817
## 535                     a lot   32   32   2  4.547387437
## 536             restrictions on   16   16   2  4.629832179
## 537                  to stay   32   32   2  3.441403591
## 538                   than $   18   18   2  3.319474732
## 539               during his   17   17   2  3.482192388
## 540                    he is   59   59   2  1.836231050
## 541                  in favor   26   26   2  4.000476333
## 542               our country   12   12   2  4.056529690
## 543              case against   11   11   2  4.241387771
## 544                  when you   14   14   2  3.762047328
## 545             more difficult   10   10   2  5.168464494
## 546                because it   22   22   2  3.005567765
## 547                    use of   36   36   2  2.693000825
## 548                carried out   16   16   2  7.984224976
## 549                 much more   14   14   2  3.785802231
## 550                   was not   41   41   2  2.181822425
## 551              professor at   12   12   2  5.147791390
## 552                 a federal   55   55   2  1.992375973
## 553                  deal with   13   13   2  4.403384965
## 554                  to allow   34   34   2  2.958210199
## 555                   live in   24   24   2  3.853805793
## 556                   to work   42   42   2  2.441666024
## 557                  during a   36   36   2  2.547730504
## 558            state university   12   12   2  4.094391964
## 559                    end of   30   30   2  3.199766794
## 560               the supreme   58   58   2  2.285470874
## 561               would make   17   17   2  3.399801280
## 562                   in fact   26    5   2  3.396480405
## 563               they should   20   20   2  3.106034573
## 564                  debate on   20   20   2  3.272230517
## 565                     i was   32   32   2  2.432839093
## 566                   a woman   30   30   2  3.010861484
## 567                10 percent   10   10   2  4.390631226
## 568             immigration law   20   20   2  3.064851771
## 569                 about how   15   15   2  3.568419229
## 570                     we do   16   16   2  3.430874759
## 571                  far more   10   10   2  4.653785488
## 572                 accused of   28   28   2  3.580248280
## 573                 new trial   15   15   2  3.581289545
## 574                   to seek   30   30   2  3.681861314
## 575                 new study   14   14   2  3.728345824
## 576              been charged   10   10   2  4.489576104
```

```
## 577                     i believe    11     11     2   4.295228772
## 578                     it easier    12     12     2   5.112122604
## 579                     action on    17     17     2   3.665241008
## 580                        is now    26     26     2   2.752373129
## 581                     failed to    31     31     2   4.177669665
## 582             immigration status    10     10     2   4.529122882
## 583                   are expected  14     14     2   3.823686889
## 584                  an additional  11     11     2   5.128686062
## 585                     on climate   26     26     2   2.739829616
## 586                    from mexico   13     12     2   4.102010570
## 587                     about the  107    107     2   1.407331181
## 588                       some of    60     60     2   1.840799563
## 589          immigration authorities 10    10     2   4.487002194
## 590                         as an    40     40     2   2.137515851
## 591                     new york   165    160     2  10.735460454
## 592                       how to     47     47     2   2.155153505
## 593                    checks for    17     17     2   3.817926482
## 594                       i can't    10     10     2   4.515800807
## 595                     limits on    16     16     2   5.296352780
## 596                     has taken    11     11     2   4.440910913
## 597                      where he    16     16     2   3.372979278
## 598                     in georgia   26     23     2   3.139205849
## 599                  philip morris   37     35     2  15.105053076
## 600                       plans to    31     31     2   4.326099156
## 601                       we know    11     11     2   4.099753851
## 602                     living in    22     22     2   3.853373032
## 603                     and then     25     25     2   3.344958232
## 604                      all over    15     15     2   3.417047936
## 605                     they can     20     20     2   2.985847041
## 606                       i'm not    12     12     2   4.055246343
## 607                  the american   66     66     2   1.922459677
## 608                      that we     32     32     2   2.428181062
## 609                       to pay     30     30     2   3.045777605
## 610                     he added     12      2     2   4.301097905
## 611                      the law     99     95     2   1.473233589
## 612              cigarette smoking   10     10     2   4.203502291
## 613                     to bring     28     28     2   3.809524137
## 614                     has come     14     14     2   3.606735211
## 615              immigration agents   10     10     2   4.299493222
## 616                   the measure    49     44     2   2.440112162
## 617                      of our     44     44     2   2.141788846
## 618                     it seems     11     10     2   5.027852895
## 619                     women who    12     12     2   3.870077006
## 620                      as many     22     22     2   2.828727831
## 621                    bill would    18     18     2   3.083526671
## 622                    first day     10     10     2   4.112121296
## 623                      to pass     28     28     2   3.986469085
## 624                     makes it     11     11     2   4.261802258
## 625                       try to     38     38     2   5.038616340
## 626                    wanted to     29     29     2   4.260217155
## 627                   should not     19     19     2   2.998034656
## 628                   before the     58     58     2   1.957439593
## 629                      we can     15     15     2   3.346198515
## 630                      had no     15     14     2   3.336774722
```

```
## 631             against the   64    64   2  1.821932601
## 632                  in new   67    67   2  1.625558220
## 633                 to give   28    28   2  3.157681474
## 634               time when   11    11   2  3.882400020
## 635               he didn't   13    13   2  3.812411082
## 636             marco rubio   11     6   2 13.412348296
## 637              the number   45    45   2  2.576714519
## 638             were charged  10    10   2  4.221580885
## 639              cannot be    11    11   2  4.362797837
## 640                 kind of   29    29   2  4.640825513
## 641                who want   15    15   2  3.375321971
## 642              during the   50    50   2  2.124464833
## 643              don't have   17    17   2  3.151857963
## 644                a former   29    29   2  2.719866429
## 645                 take up   11    11   2  3.843213202
## 646               we should  14    14   2  3.390594888
## 647                 man who   12    12   2  3.730003909
## 648                 to stop   28    28   2  3.001492260
## 649                  lot of   24    24   2  3.787798367
## 650             scheduled to  26    26   2  3.260461285
## 651             increase in   23    23   2  3.049640696
## 652           in washington   25    25   2  2.913031086
## 653                 it also   22    22   2  2.696410948
## 654         secondhand smoke  24    23   2 10.326490811
## 655               are trying  12    12   2  3.816445565
## 656               which has   22    22   2  2.669561803
## 657                  of his   90    90   2  1.360688042
## 658               they also   20    20   2  2.804984295
## 659                  to the  624   624   2  0.521941742
## 660              they don't   15    15   2  3.274605779
## 661                  i want   12    12   2  3.633725363
## 662                a decade   23    23   2  4.144642233
## 663             an immigrant  15    15   2  3.277577383
## 664               argue that  18    18   2  5.241918685
## 665             an increase   12    12   2  3.741589675
## 666                will not   28    28   2  2.344439807
## 667               to protect  33    33   2  4.874780920
## 668                favor of   23    23   2  3.693319580
## 669              on saturday  13    12   2  4.089337553
## 670                    of a  298   298   2  0.732864644
## 671              tobacco use   10    10   2  3.908790183
## 672              lighting up   10    10   2  7.421154541
## 673                   go to   31    31   2  2.554253034
## 674           been convicted  10    10   2  3.923461275
## 675             ensure that   15    15   2  4.553070729
## 676                 has not   30    30   2  2.237305117
## 677              law school   10    10   2  3.857766914
## 678                  when i   12    12   2  3.521651395
## 679                 used to   31    31   2  2.527396295
## 680                 tried to   24    24   2  3.400787180
## 681                 do what   11    11   2  3.661878042
## 682                   i see   10    10   2  3.938127317
## 683                 to live   25    25   2  3.165602906
## 684           the government  53    52   2  2.010385980
```

```
## 685          immigration system   11   11   2  3.703223905
## 686              scheduled for   15   15   2  3.539520706
## 687                   taxes on   12   12   2  4.297426115
## 688                      as he   33   33   2  2.119131223
## 689                  the study   48   40   2  2.159090509
## 690                   but they   27   27   2  2.333576203
## 691                 the city's   45   45   2  4.207178322
## 692                 sandy hook   12   12   2 14.006555529
## 693                 instead of   22   21   2  3.816745131
## 694                   bars and   18   18   2  3.926740407
## 695                   trying to   72   72   2  6.438561183
## 696               an agreement   10   10   2  4.101094775
## 697                     can do   11   11   2  3.618716299
## 698             civil liberties   15   15   2 10.121359850
## 699                    new law   19   19   2  2.760986410
## 700                   group of   30   30   2  2.476735784
## 701                     led by   10   10   2  4.335864390
## 702                    at home   15   15   2  3.180178866
## 703              evidence that   17   17   2  3.239873182
## 704              million people   12   12   2  3.462738551
## 705                  under the   49   49   2  1.979705900
## 706                 demand for   13   13   2  4.105880730
## 707                  out there   10    9   2  3.776384433
## 708                 in january   19   18   2  3.909440476
## 709                   agreed to   23   23   2  3.444139551
## 710                 claim that   14   14   2  4.093953497
## 711                  before he   14   14   2  3.202940908
## 712               an important   10   10   2  4.027877540
## 713                     to die   23   23   2  3.671500045
## 714                   on death   35   35   2  2.034219898
## 715                   a second   28   28   2  2.527966939
## 716                     into a   43   43   2  1.903196383
## 717                  by lethal   10   10   2  4.314985416
## 718                   they did   14   14   2  3.221267821
## 719                 to address   23   23   2  3.376672024
## 720                     i know   10    7   2  3.805352066
## 721                    all but   10   10   2  3.738297726
## 722                       by a   97   97   2  1.225055942
## 723                  court has   20   20   2  2.641119798
## 724               clinton said   13    8   2  3.397792756
## 725                   the next   48   48   2  2.073638322
## 726                  bill that   29   29   2  2.260354894
## 727                  when she   10   10   2  3.732216738
## 728          immigration policy   10   10   2  3.758168957
## 729                 hearing on   15   15   2  3.228697153
## 730                   we're not   10   10   2  3.878576800
## 731                 to prevent   23   23   2  4.007996015
## 732                   food and   17   17   2  3.992315913
## 733                    back to   31   31   2  2.379833224
## 734                     and i   36   36   2  2.078757394
## 735                     in st   21   21   2  2.992987360
## 736                   the city   63   63   2  1.689366724
## 737                    a jury   27   27   2  2.529133196
## 738                 refused to   27   27   2  4.700577598
```

```
## 739                    comment on    11     11    2  4.581266279
## 740                        at age    10     10    2  4.008916843
## 741                       role in    17     17    2  4.040358278
## 742                     in recent    23     23    2  2.748919462
## 743                         i had    18     18    2  2.736949508
## 744                         a day    32     32    2  2.240277218
## 745                       man and    24     24    2  2.599227130
## 746                    his family    12     12    2  3.434797428
## 747                   which would    20     19    2  2.593452836
## 748                       but it's   11     11    2  3.622122227
## 749                       at risk    10     10    2  3.981132665
## 750                       pay for    15     15    2  3.266953109
## 751                     will have    29     29    2  2.155880315
## 752                    the nation    38     36    2  3.939553236
## 753             attorney's office    11     10    2  9.891511091
## 754                       on both    17     17    2  2.939021571
## 755                    whether he    12     12    2  3.367286617
## 756                      they say    15      8    2  3.011221018
## 757                     the white    38     38    2  2.473044036
## 758                       who has    28     27    2  2.182805677
## 759                       we will    19     18    2  2.640801110
## 760                     they want    13     13    2  3.248091717
## 761                       in 2000    18     17    2  3.502925556
## 762                 emissions from    13     13    2  3.256331442
## 763                     noted that    13     13    2  4.497595032
## 764                       call for    14     14    2  3.380363974
## 765                 compared with    10     10    2  5.894793745
## 766                         in an    71     70    2  1.391589558
## 767                       much as    14     14    2  3.114117029
## 768                       up with    21     21    2  2.502821382
## 769                     reason for    13     13    2  3.585281336
## 770                      who were    24     24    2  2.329034596
## 771                opportunity to    22     22    2  4.162709200
## 772                      she had    13     13    2  3.146148768
## 773                   to consider    22     22    2  3.140120564
## 774                  said tuesday    14     12    2  3.105511061
## 775                      he also    20     19    2  2.559398502
## 776                   the tobacco    87     87    2  1.343884432
## 777                    of appeals    28     28    2  2.407769554
## 778                       the new   115    115    2  1.145295825
## 779                    believed to    21     21    2  3.942741766
## 780                  and customs    30     30    2  5.590537200
## 781                 immigrants are    19     19    2  2.602441549
## 782                      become a    23     23    2  2.632798676
## 783                the university    40     40    2  2.238820820
## 784                       a long    26     26    2  2.467065432
## 785                      to meet    21     21    2  3.287446664
## 786                      just as    17     17    2  2.752173845
## 787                         i am    36     36    2  9.175622395
## 788                       the INS    32     32    2  3.150638122
## 789                      a small    20     20    2  3.069185275
## 790                     is likely    15     15    2  3.058276575
## 791                    center for    14     14    2  3.276079762
## 792                      to fight    23     23    2  2.857295820
```

```
## 793            government has  12   12   2  3.262544984
## 794              a spokesman   19   19   2  3.245810339
## 795           american people  11   11   2  3.361551588
## 796                  cost of   19   19   2  3.506103356
## 797               the report   45   42   2  2.009579940
## 798             federal court  14   14   2  2.975842560
## 799                  rest of   19   19   2  3.873856502
## 800                  on july   12   12   2  3.559824779
## 801               this issue   13   13   2  3.103937775
## 802          elementary school 13   13   2  9.422501694
## 803               to provide   21   21   2  3.130242092
## 804                   to try   21   21   2  3.130242092
## 805                of course   19    2   2  3.395875099
## 806                  be made   14   14   2  3.008735054
## 807               the federal  75   75   2  1.422528624
## 808                have made   14   14   2  3.002379326
## 809                  to come   29   29   2  2.313085648
## 810                   a good   21   21   2  2.818434612
## 811                chance to   20   20   2  3.263647600
## 812            yesterday that  17   17   2  2.863283993
## 813                  the end   35   35   2  2.436672948
## 814                come from   12   12   2  3.247121651
## 815                  said it   32   32   2  1.947183499
## 816                  many of   44   44   2  1.749512010
## 817                   what i   10   10   2  3.445235547
## 818              has already   10   10   2  3.552241826
## 819                  in 1995   16   13   2  3.942476535
## 820              in congress   27   27   2  2.274359409
## 821                to create   20   20   2  3.954390053
## 822                to become   26   26   2  2.471984181
## 823                shows that  14   14   2  3.262108553
## 824            suggested that  12   12   2  4.611342908
## 825                 told the   42   42   2  1.941877886
## 826                to decide   20   20   2  3.198002263
## 827              authority to   19   19   2  3.567168275
## 828             state senate   10   10   2  3.439548145
## 829                apply for   11   11   2  3.897050508
## 830              in november   16   15   2  3.524708443
## 831          the legislature   30   29   2  3.254055761
## 832                attack on   10   10   2  3.978896706
## 833               people are   20   20   2  2.429015842
## 834                  the NRA   32   32   2  3.769721590
## 835             eligible for   18   18   2  6.190842977
## 836               they could   17   17   2  2.633375278
## 837     immigration officials  11   11   2  3.259560287
## 838                 need for   16   16   2  2.856310314
## 839             the nation's   81   81   2  5.791113111
## 840                result of   18   18   2  3.981407368
## 841                  half of   19   19   2  3.092080280
## 842                  to kill   19   19   2  3.672424890
## 843                  many as   18   18   2  2.558058692
## 844                    if it   19   19   2  2.477710043
## 845                not going   11   11   2  3.302053829
## 846                  men and   19   19   2  2.755636821
```

22

```
## 847                 of homeland   18   18   2  3.343075772
## 848                clara county   13   13   2  9.092073801
## 849                     them to   40   40   2  1.817808052
## 850                of thousands   20   20   2  2.889677465
## 851                   source of   21   21   2  4.702730657
## 852                  would help   11   11   2  3.263496769
## 853                 marriage is   19   19   2  2.510829102
## 854                     it does   10    9   2  3.523958514
## 855             citizenship for   13   13   2  3.224198389
## 856                   has made   12   12   2  3.106585757
## 857                     used by   12   12   2  3.161614686
## 858               released from   10   10   2  3.494448168
## 859                 to continue   22   22   2  2.717192506
## 860                      it can   17   17   2  2.602647559
## 861                      in san   20   20   2  2.676831199
## 862                   on behalf   15   15   2  6.174499582
## 863                  that there   22   22   2  2.352077295
## 864                     in 2002   15   15   2  3.679073249
## 865                     to vote   31   31   2  2.095541692
## 866                  millions of   19   19   2  4.461671531
## 867                  majority of   18   18   2  3.146538756
## 868                  because he   14   14   2  2.823050333
## 869                 response to   24   24   2  4.894802322
## 870                  looking for   10   10   2  4.271051928
## 871                      it has   28   28   2  2.000623654
## 872                     long as   11   11   2  3.256240352
## 873                    for more   36   36   2  1.778616170
## 874                    that she   20   20   2  2.461933841
## 875                thousands of   58   58   2  6.665081876
## 876                    create a   15   14   2  3.553156208
## 877                    but even   11   11   2  3.218862592
## 878                     a path   16   16   2  3.589121379
## 879                     path to   19   19   2  4.230500436
## 880                 chairman of   17   17   2  4.016656542
## 881                   work with   12   12   2  3.079247448
## 882                   among the   37   37   2  2.001520547
## 883                    the jury   38   38   2  2.063809487
## 884                  to ensure   19   19   2  4.206535855
## 885                     they do   15   15   2  2.696724737
## 886                     made it   12   12   2  3.030118251
## 887              the department   44   43   2  1.845683716
## 888                 to regulate   23   23   2  4.829009793
## 889                     but mr   11    0   2  3.172534966
## 890                   member of   18   18   2  4.408870293
## 891                   are going   11   11   2  3.180853176
## 892                   board of   19   19   2  2.825769750
## 893                 centers for   10   10   2  4.858866429
## 894                  focused on   13   13   2  6.092643795
## 895                      a way   29   29   2  2.066200796
## 896                     who is   40   40   2  1.644599288
## 897                    in their   58   58   2  1.393922537
## 898                    the air   30   30   2  2.516368295
## 899                   penalty is   17   17   2  2.560826009
## 900                    what is   24   24   2  2.128748128
```

```
## 901               just one   10   10   2  3.231136717
## 902               came from   11   11   2  3.162805462
## 903              their home   10   10   2  3.266780185
## 904                in texas   19   18   2  2.643576741
## 905                 came to   26   26   2  2.253089909
## 906              intended to  18   18   2  4.177714576
## 907              said monday  12   11   2  3.013719634
## 908             opposition to 17   17   2  3.458670458
## 909              in december  14   14   2  3.889834307
## 910                a result   16   16   2  3.235448323
## 911                 after a   46   46   2  1.563759942
## 912             whether they  10   10   2  3.243964157
## 913                we could   12   12   2  2.937218377
## 914                  to win   17   17   2  3.635397296
## 915                 to save   17   17   2  3.563933585
## 916              a statement  16   16   2  3.193770910
## 917               warming is  13   13   2  2.951972125
## 918               on illegal  21   21   2  2.261265137
## 919           responsible for 14   14   2  5.945508198
## 920                bureau of  17   17   2  4.353142961
## 921                 about $   13   13   2  2.800676682
## 922               which was   21   21   2  2.219311593
## 923                 he told   12   12   2  2.975770931
## 924              a separate   15   15   2  3.978443709
## 925                 needs to  17   17   2  4.010772241
## 926                a recent   21   21   2  2.475747150
## 927               attempt to  18   18   2  4.302882462
## 928         federal immigration 13  13   2  2.790865420
## 929              of dollars   16   16   2  4.045172195
## 930                show that  14   14   2  2.908885555
## 931                  a bill   40   40   2  1.674719817
## 932                 a third   15   15   2  3.412010502
## 933                his home   11   11   2  3.063859066
## 934                 to sell   17   17   2  3.217628868
## 935               of capital  23   23   2  2.340086788
## 936                  to put   25   25   2  2.252268690
## 937               in florida  22   22   2  2.318173407
## 938                 a child   16   16   2  3.042521073
## 939              the problem   28   28   2  2.539394457
## 940               in america  20   17   2  2.451937890
## 941              account for   10   10   2  5.169030636
## 942              qualify for   10   10   2  5.169030636
## 943                 used in   20   20   2  2.412911527
## 944               clear that  12   12   2  3.186194538
## 945                  such a   29   29   2  1.954565428
## 946               a similar   18   18   2  2.711612409
## 947                level of   15   15   2  3.804004294
## 948                 say that  20   20   2  2.300065614
## 949                this bill  14   14   2  2.646377029
## 950              los angeles   23   22   2 15.225563486
## 951               women and   19   19   2  2.465762379
## 952               on whether  15   15   2  2.635070729
## 953            immigrants from 17   17   2  2.407429350
## 954              decision to   27   27   2  2.094975897
```

```
## 955               benefits to   19   19   2  2.694118889
## 956               many people   11   11   2  2.956497168
## 957                   but now   13    9   2  2.748815649
## 958                 a hearing   22   22   2  2.315102626
## 959                   that if   23   23   2  2.115153839
## 960               police said   12   10   2  2.880931867
## 961                 effect on   10   10   2  3.342815338
## 962                   oil and   12   12   2  3.869845517
## 963                  and drug   17   17   2  2.718768493
## 964                 a special   15   15   2  3.131079817
## 965              supporters of   17   17   2  2.826959010
## 966                   seem to   16   16   2  4.063020033
## 967                   you are   15   15   2  2.528681440
## 968                is charged   11   11   2  3.117247042
## 969                   down on   14   14   2  2.685084871
## 970                  the idea   25   25   2  2.809945987
## 971                dangers of   16   16   2  4.437250577
## 972                   form of   16   16   2  4.437250577
## 973                  to marry   16   16   2  4.039065220
## 974                  issue of   26   26   2  2.069891688
## 975                     as it   31   31   2  1.751307657
## 976                  the rest   25   25   2  3.626766136
## 977                   but not   25   25   2  1.940737805
## 978               his support   11   11   2  2.935786662
## 979               necessary to   16   16   2  4.188187919
## 980                involved in   15   15   2  5.088674092
## 981                     up to   46   46   2  1.497402179
## 982            and restaurants   15   15   2  2.933570037
## 983                    in his   64   64   2  1.235376247
## 984              concluded that   11   11   2  5.274843456
## 985               emissions by   11   11   2  2.948458317
## 986                    due to   15   15   2  3.399550550
## 987                 passage of   14   14   2  3.828132019
## 988                   care for   10   10   2  3.332684856
## 989                  is trying   11   11   2  3.047635220
## 990                  to build   15   15   2  3.674110939
## 991              to determine   15   15   2  3.674110939
## 992                prepared to   15   15   2  3.889127233
## 993                   to hear   15   15   2  3.442295085
## 994                    to buy   15   15   2  3.375598964
## 995                  to avoid   15   15   2  3.865175669
## 996                 a growing   16   16   2  2.768041775
## 997                      i do   11    8   2  2.848905123
## 998               and unusual   13   13   2  4.261503540
## 999                the future   26   25   2  2.466497348
## 1000             in manhattan   12   11   2  3.741018223
## 1001              people from   18   18   2  2.240737987
## 1002                  for gun   29   29   2  1.787194776
## 1003                  was one   19   19   2  2.181727934
## 1004                  in 2007   12   10   2  4.015469164
## 1005                 in front   13   13   2  3.186853822
## 1006                   of all   42   42   2  1.529886786
## 1007            their children   10   10   2  2.983901746
## 1008               the editor   25   25   2  3.863164749
```

```
## 1009              time to   33    33    2  1.756728990
## 1010               set for   13    13    2  2.737770017
## 1011              aimed at   18    18    2  7.881193044
## 1012              vote for   14    14    2  2.617982605
## 1013              it could   16    16    2  2.360695212
## 1014             access to   15    15    2  4.125525496
## 1015               ways to   15    15    2  4.125525496
## 1016          a compromise   13    13    2  3.450463637
## 1017      the constitution   23    22    2  2.878427185
## 1018               wake of   14    14    2  4.164618438
## 1019               even as   13    13    2  2.619119752
## 1020          opponents of   14    14    2  3.209050262
## 1021               have to   93    93    2  1.005054305
## 1022               in 2004   12    12    2  3.463367692
## 1023         an immigration 19    19    2  2.148898484
## 1024               rate of   15    15    2  2.932089681
## 1025         companies have 10    10    2  2.996601580
## 1026              gays and   11    11    2  4.171920562
## 1027                him to   27    27    2  1.957072246
## 1028              on april   10    10    2  3.133103246
## 1029            many other   10    10    2  2.924745411
## 1030               became a   14    14    2  2.892423739
## 1031             appears to  14    14    2  3.631229090
## 1032             like this   10    10    2  2.930200727
## 1033            and lesbian  18    18    2  5.339670934
## 1034         officials have 13    13    2  2.588390165
## 1035             have come   10    10    2  2.972197000
## 1036              to prove   14    14    2  3.698257297
## 1037               to raise  14    14    2  3.698257297
## 1038              of their   62    62    2  1.216684417
## 1039                say it   11    11    2  2.796875003
## 1040              a chance   15    15    2  2.788746442
## 1041              a little   17    17    2  2.529359313
## 1042              hands of   13    13    2  3.856603860
## 1043               terms of  13    13    2  3.856603860
## 1044           declined to   14    14    2  3.933524191
## 1045          the justices   22    22    2  3.401308034
## 1046               to send   14    14    2  3.798345502
## 1047                up for   24    24    2  1.912909238
## 1048            in october  12    11    2  4.394972881
## 1049                 is so   15    15    2  2.433128901
## 1050             china and   11    11    2  3.449743676
## 1051           santa clara  15    15    2 13.283342195
## 1052             las vegas  11    11    2 15.021795387
## 1053        something that  11    11    2  3.005978570
## 1054             while he   10    10    2  2.890346416
## 1055                on gun   25    25    2  1.853943255
## 1056                have a   74    74    2  1.090628148
## 1057       immigration bill 11    11    2  2.741578359
## 1058          the proposal  28    25    2  2.149855916
## 1059              the best   22    22    2  2.834873217
## 1060              the fact   26    26    2  2.292114451
## 1061            of killing  17    17    2  2.537577447
## 1062              sides of  13    13    2  3.304502185
```

```
## 1063              required to    14     14    2  3.155772192
## 1064               when the      65     65    2  1.215915909
## 1065              report said    10     10    2  2.919483402
## 1066                 in 1986     11     11    2  3.831801418
## 1067               changes in    12     12    2  3.062684485
## 1068                 said the   133    133    2  0.829873541
## 1069                 to limit    14     14    2  4.034743773
## 1070                in august    11      7    2  3.931889575
## 1071                 a lawyer    15     15    2  2.705350682
## 1072                enough to    17     17    2  2.560615026
## 1073                     on a   109    109    2  0.887967261
## 1074                this time    12     10    2  2.607605325
## 1075                 a single    12     12    2  3.974138823
## 1076                linked to    15     15    2  4.435689910
## 1077                    a key    14     14    2  2.845135869
## 1078                 order to    19     19    2  2.336098279
## 1079              a different    13     13    2  3.034265507
## 1080                 in 1992     11      8    2  4.168287750
## 1081               the agency    28     28    2  2.087320804
## 1082                  lack of    13     13    2  4.236107664
## 1083                 dozens of    18     18    2  5.171029254
## 1084                 that has    41     41    2  1.424238029
## 1085                 but also    14     13    2  2.399562956
## 1086              the company    26     24    2  2.215138655
## 1087                  guns in    17     17    2  2.338445191
## 1088              people have    17     17    2  2.176081552
## 1089                 to begin    14     14    2  3.029160201
## 1090                victims of    14     14    2  2.865240901
## 1091               the public    51     51    2  1.397333451
## 1092                 a series    14     14    2  4.758907787
## 1093           for citizenship    12     12    2  2.704689079
## 1094                the ruling    35     32    2  1.760710916
## 1095                 to minors    13     13    2  3.535683035
## 1096                   to bear    13     13    2  3.726747765
## 1097                effects of    12     12    2  3.679397276
## 1098               in newtown    11     10    2  4.311393292
## 1099                 phase of    12     12    2  3.779485462
## 1100          legislation that    14     14    2  2.484171443
## 1101                on smoking    21     21    2  1.961451023
## 1102                     a $     37     37    2  1.523685785
## 1103                  as long    10     10    2  2.855883663
## 1104                could have    15     15    2  2.294300303
## 1105                 an issue    11     11    2  2.678594579
## 1106               the largest    24     24    2  2.326320267
## 1107              through the    35     35    2  1.680143692
## 1108                 in 1997     11     11    2  3.320942904
## 1109                  may not    11     11    2  2.667419336
## 1110               outside the    21     21    2  2.415600988
## 1111                  build a    11     11    2  3.309454455
## 1112                 to close    15     15    2  2.731982269
## 1113                 his life    10     10    2  2.788712934
## 1114                 NEW YORK    10      8    2 12.810514178
## 1115               licenses to    14     14    2  2.874841348
## 1116                 is going    12     12    2  2.607150505
```

```
## 1117          approach to  13    13   2  3.139745797
## 1118            amount of  14    14   2  4.675458244
## 1119        the atmosphere  26    26   2  4.412540277
## 1120              case is  14    14   2  2.384037238
## 1121              to leave  13    13   2  3.174645956
## 1122            palm beach  17    17   2 12.497058793
## 1123              risk of  13    13   2  2.950829051
## 1124              made by  11    11   2  2.652498801
## 1125       the legislation  38    37   2  1.622311307
## 1126             he could  15    15   2  2.258757388
## 1127           defense of  17    17   2  2.369623068
## 1128             which he  14    14   2  2.321614593
## 1129               but we  12    12   2  2.516852253
## 1130               at gun  18    18   2  2.059691220
## 1131            to impose  15    15   2  4.612413804
## 1132            after all  10     2   2  2.729254065
## 1133                was a  77    77   2  1.016828784
## 1134            impact of  12    12   2  3.227383787
## 1135              side of  12    12   2  3.227383787
## 1136              to move  17    17   2  2.420222751
## 1137              a great  13    13   2  2.841338257
## 1138        immigrants were  11    11   2  2.598575318
## 1139            school in  19    19   2  2.108682806
## 1140              in 1993  10     8   2  3.840719855
## 1141              so they  10     9   2  2.724135529
## 1142          on same-sex  14    14   2  2.365573537
## 1143              in 2001  10     9   2  3.740631698
## 1144           an american  11    11   2  2.613826617
## 1145            about his  17    17   2  2.098515040
## 1146              issued a  12    12   2  2.905276929
## 1147       throughout the  19    19   2  4.124726424
## 1148           about what  10    10   2  2.706441062
## 1149            in arizona  11    10   2  3.115072055
## 1150             front of  13    13   2  2.869141564
## 1151               idea of  13    13   2  2.869141564
## 1152              with his  26    26   2  1.696693437
## 1153           failure to  12    12   2  3.784820155
## 1154           and others  16    14   2  2.358342303
## 1155               lead to  15    15   2  2.599341129
## 1156              to offer  14    14   2  2.772411873
## 1157               of its  39    39   2  1.436610101
## 1158              to deny  12    12   2  3.649647970
## 1159        a constitutional  14    14   2  2.612166575
## 1160              seeks to  12    12   2  3.909988041
## 1161               sort of  15    15   2  4.993626180
## 1162               to ask  13    13   2  2.957562464
## 1163               to quit  13    13   2  2.957562464
## 1164                at $  15    15   2  2.207291655
## 1165                at an  24    24   2  1.746575425
## 1166            the courts  23    23   2  2.255159906
## 1167            month that  11    11   2  2.726889453
## 1168        restaurants and  10    10   2  3.200309319
## 1169              fear of  14    14   2  2.638434486
## 1170             a better  16    16   2  2.358052442
```

```
## 1171      a national   26   26   2  1.764366210
## 1172        along the   22   22   2  2.177150897
## 1173        since the   39   39   2  1.502636253
## 1174   the republican   36   36   2  1.626844253
## 1175       numbers of   13   13   2  2.793624556
## 1176            up in   32   32   2  1.546969826
## 1177        his first   12   12   2  2.444587054
## 1178   to citizenship   19   19   2  2.164281286
## 1179          sent to   14   14   2  2.689100486
## 1180           to cut   13   13   2  2.909929668
## 1181       planned to   12   12   2  3.184013094
## 1182          a crime   17   17   2  2.239124854
## 1183        slaying of   11   11   2  4.075450325
## 1184         not have   24   24   2  1.719149659
## 1185        of alcohol   11   11   2  3.331964682
## 1186          in june   13   11   2  2.594260887
## 1187        year from   10   10   2  2.653424786
## 1188          stay in   12   12   2  2.675507576
## 1189         to reach   12   12   2  3.097546161
## 1190         would do   11   11   2  2.506380700
## 1191       rights and   18   18   2  2.074476165
## 1192      industry is   10   10   2  2.711403885
## 1193           to end   18   18   2  2.190929200
## 1194      violation of   12   12   2  4.526723501
## 1195   a controversial   11   11   2  3.060186411
## 1196      with their   22   22   2  1.777087508
## 1197          in 1998   10   10   2  3.174198635
## 1198       opposed to   12   12   2  3.007068157
## 1199       decided to   13   13   2  2.763230281
## 1200            a big   14   14   2  2.490040293
## 1201       to discuss   12   12   2  4.196210662
## 1202         you have   13   13   2  2.297337931
## 1203       called for   11   11   2  2.589077287
## 1204          loss of   11   11   2  4.242509137
## 1205          none of   11   11   2  4.242509137
## 1206      for illegal   19   19   2  1.931683392
## 1207        seemed to   11   11   2  3.701296561
## 1208        return to   11   11   2  3.489977982
## 1209            as if   13   13   2  2.301945782
## 1210         law that   22   22   2  1.794955965
## 1211        in denver   10    9   2  3.121550202
## 1212         to stand   12   12   2  2.983126316
## 1213         a dozen   11   11   2  4.526594227
## 1214        and local   17   17   2  2.148058881
## 1215          the way   39   39   2  1.483768414
## 1216       there will   10   10   2  2.574155644
## 1217         ready to   11   11   2  3.315615109
## 1218         comes to   15   15   2  2.414573821
## 1219     immigrants in   30   30   2  1.543057547
## 1220      smoking and   28   28   2  1.596286236
## 1221       the latest   19   19   2  3.904652402
## 1222         not been   15   15   2  2.104210967
## 1223         lives in   10   10   2  2.963407059
## 1224     the question   23   23   2  2.094787746
```

```
## 1225              cruel and    13     13    2  5.556515941
## 1226      the environment    17     17    2  3.149649395
## 1227           history of    11     11    2  3.029425040
## 1228             of these    19     19    2  2.026878499
## 1229         of california    20     19    2  1.965148928
## 1230           hundreds of    36     36    2  6.700716993
## 1231           states and    27     27    2  1.615982993
## 1232             to block    11     11    2  3.945631488
## 1233             of which    21     21    2  1.902769457
## 1234             agree to    12     12    2  2.856771725
## 1235             risks of    10     10    2  3.504729250
## 1236          designed to    18     18    2  5.401518464
## 1237          friday that    10     10    2  2.707684205
## 1238           to enforce    11     11    2  3.214710753
## 1239            the kyoto    17     17    2  2.898320214
## 1240           whether to    24     24    2  1.766675362
## 1241              in bars    11     11    2  2.765356017
## 1242         of secondhand    10     10    2  3.401189612
## 1243               led to    12     12    2  2.811304607
## 1244            to follow    11     11    2  3.143247042
## 1245              to turn    11     11    2  3.143247042
## 1246              of whom    10     10    2  3.971758117
## 1247               all of    38     38    2  1.357680870
## 1248              carry a    14     14    2  2.334824911
## 1249             a person    14     14    2  2.381213847
## 1250          congress to    22     22    2  1.845484052
## 1251               than a    39     39    2  1.323482066
## 1252            to extend    11     11    2  4.112690319
## 1253         aldape guerra    10      7    2 16.029445077
## 1254          a temporary    11     11    2  2.831904176
## 1255            that will    34     34    2  1.391170167
## 1256            needed to    11     11    2  3.037964400
## 1257          the dangers    17     17    2  3.629242145
## 1258          as children    11     10    2  2.418168698
## 1259         from smoking    14     14    2  2.130159642
## 1260       possibility of    10     10    2  4.151380052
## 1261              able to    35     35    2  6.566529692
## 1262              but the   108    108    2  0.805803834
## 1263               to ban    30     30    2  1.534473177
## 1264           of illegal    31     31    2  1.491184902
## 1265              hard to    12     12    2  2.726137313
## 1266            san diego    15     14    2 11.288602594
## 1267             were more    13     13    2  2.179273097
## 1268             series of    15     15    2  5.330103144
## 1269              to show    17     17    2  2.118821941
## 1270         the industry    32     32    2  1.593484879
## 1271              have no    12     12    2  2.273400640
## 1272          decision on    11     11    2  2.417036030
## 1273          monday that    10     10    2  2.618315852
## 1274           percent in    25     25    2  1.630004936
## 1275             jeb bush    11      8    2 11.291813948
## 1276          the justice    26     26    2  1.813698893
## 1277          in response    10     10    2  2.853258023
## 1278             remain in    10     10    2  2.784335390
```

```
## 1279               to receive   12    12    2  2.702195450
## 1280                   on his   31    31    2  1.421035337
## 1281                   a plan   16    16    2  2.122524866
## 1282                 to blame   11    11    2  4.313365761
## 1283                also said   14    13    2  2.093204871
## 1284                  to curb   16    16    2  5.262869130
## 1285                murder and   15    15    2  2.139369289
## 1286                  a judge   25    25    2  1.643984790
## 1287               brought to   14    14    2  2.372136598
## 1288                   is one   20    20    2  1.762793493
## 1289           the democratic   23    23    2  1.956607897
## 1290                 round of   10    10    2  4.352055475
## 1291              the mexican   20    20    2  2.178106887
## 1292                 seems to   11    11    2  2.870896087
## 1293                   of her   32    32    2  1.440316920
## 1294               charges in   11    11    2  2.567509319
## 1295                asked for   10    10    2  2.536158544
## 1296             campaign for   10    10    2  2.536158544
## 1297                   took a   14    14    2  2.237158236
## 1298                 parts of   14    14    2  5.263254363
## 1299               version of   14    14    2  5.263254363
## 1300               received a   10    10    2  2.803300792
## 1301                 like the   39    39    2  1.351626807
## 1302                know that   11    11    2  2.428303448
## 1303                killing a   13    13    2  2.334014954
## 1304                     be a   66    66    2  0.974490457
## 1305                  the GOP   18    18    2  2.357921370
## 1306              committed to   12    12    2  2.610610196
## 1307               for nearly   10    10    2  2.529530701
## 1308                easier to   10    10    2  3.147535582
## 1309                 steps to   10    10    2  3.147535582
## 1310               arrested in   11    11    2  2.543693982
## 1311                to repeal   10    10    2  3.200566047
## 1312                of carbon   19    19    2  1.894030253
## 1313              states from   12    12    2  2.210531002
## 1314              coalition of   10    10    2  2.938295955
## 1315                 after he   12    12    2  2.199331111
## 1316               the debate   33    33    2  1.511442047
## 1317                 apply to   11    11    2  2.773248131
## 1318           legislation to   19    19    2  1.898553052
## 1319                 having a   10    10    2  2.730532038
## 1320              the current   19    19    2  2.191550866
## 1321                  are now   10    10    2  2.402709968
## 1322                a lawsuit   13    13    2  2.330423009
## 1323                 a number   16    16    2  2.041916072
## 1324                 said in a   24     0    3  2.789861754
## 1325                   on how   13    13    2  2.131142627
## 1326                   on its   18    18    2  1.803579146
## 1327                to remain   12    12    2  2.550945492
## 1328                  had not   14    14    2  2.019087422
## 1329               control and   16    16    2  1.986904604
## 1330                  to join   14    14    2  5.133379794
## 1331               to overturn   14    14    2  5.133379794
## 1332                to defend   10    10    2  4.222255267
```

```
## 1333         a comprehensive   10    10    2  2.782439221
## 1334                 face a     13    13    2  2.263170099
## 1335              on their      26    26    2  1.493002701
## 1336            behind the      16    16    2  2.312150790
## 1337             they said      23    13    2  1.575454448
## 1338            appeared to     10    10    2  2.946850657
## 1339           attention to     10    10    2  2.946850657
## 1340                has no      10    10    2  2.357309807
## 1341                act of      13    13    2  2.312240995
## 1342               for all      20    20    2  1.699660301
## 1343              the last      58    58    2  1.069210550
## 1344               under a      20    20    2  1.750295676
## 1345               also be      13    13    2  2.069814025
## 1346             the earth      15    15    2  2.705357824
## 1347           the economy      18    18    2  2.206670728
## 1348          government is     11    11    2  2.274583349
## 1349               we were     11    11    2  2.226175710
## 1350             the group      35    31    2  1.409135915
## 1351              guilty of     11    11    2  2.562753503
## 1352              what has      11    11    2  2.223625606
## 1353                as the     135   134    2  0.669100455
## 1354              woman in      11    11    2  2.411606133
## 1355       the government's    14    14    2  3.172709623
## 1356          for instance     10     0    2  6.407981813
## 1357           the district    28    28    2  1.600883803
## 1358            a cigarette    17    17    2  1.913168347
## 1359               to drive    10    10    2  2.864070077
## 1360            the biggest    14    14    2  3.061479070
## 1361               by some    13    13    2  2.049344061
## 1362               moved to    11    11    2  2.602603641
## 1363               to apply    11    11    2  2.618675106
## 1364                pass a    10    10    2  2.599177235
## 1365              murder of    17    17    2  1.919160800
## 1366           measure that    10    10    2  2.417584697
## 1367               today in    11    11    2  2.391192573
## 1368                a very    16    16    2  1.967319068
## 1369           contribute to    10    10    2  4.497509727
## 1370              for many    16    14    2  1.861895388
## 1371            tuesday that    10    10    2  2.403690934
## 1372               with him    10    10    2  2.313633389
## 1373       the president's     17    17    2  4.248296106
## 1374          the president    49    47    2  1.141540697
## 1375              to combat    10    10    2  4.473574442
## 1376               to enact    10    10    2  4.473574442
## 1377           democrats and    12    12    2  2.244295273
## 1378              a permit    10    10    2  2.625234770
## 1379               to life    23    23    2  1.616811289
## 1380               if this    10    10    2  2.271728928
## 1381                it may    10    10    2  2.288697531
## 1382              behind a    10    10    2  2.539448600
## 1383              jury that    11    11    2  2.251302972
## 1384             chance of     11    11    2  2.465889493
## 1385             willing to    21    21    2  6.063057473
## 1386             change is     10    10    2  2.316164081
```

```
## 1387                      home in   14   14   2  2.031202362
## 1388                     for some   18   17   2  1.724003787
## 1389                   and they're  10   10   2  2.543178315
## 1390                   ruling that  11   11   2  2.240545529
## 1391                   in virginia  12   12   2  2.232299473
## 1392                      to tell   10   10   2  2.705831831
## 1393                       to let   11   11   2  2.503147978
## 1394                 the country's  16   16   2  4.189386778
## 1395                    the entire  16   16   2  4.189386778
## 1396                      for his   33   33   2  1.261876688
## 1397                       by u.s   14   13   2  1.914379779
## 1398                  the governor  19   19   2  1.984902190
## 1399                      the boy   14   14   2  2.638597631
## 1400                      that mr   12   11   2  2.124596747
## 1401                      the top   20   20   2  1.909793718
## 1402                      gay and   18   18   2  1.748105075
## 1403                    author of   18   17   2  6.018336569
## 1404                    the worst   13   13   2  3.226349906
## 1405                     the wake   20   20   2  4.743203777
## 1406                      work in   15   15   2  1.924128316
## 1407                in restaurants  10   10   2  2.467539155
## 1408                    thanks to   11   11   2  4.925100449
## 1409                     the word   14   13   2  2.571901339
## 1410                      he will   18   18   2  1.670775502
## 1411                     the door   13   13   2  2.889862916
## 1412                      come to   18   18   2  1.790173618
## 1413                     and even   16   16   2  1.867038152
## 1414                     seek the   16   16   2  2.092488207
## 1415                  of congress   20   20   2  1.673355047
## 1416                    influx of   10   10   2  4.939851595
## 1417                     years of   27   27   2  1.420172260
## 1418                     to have   78   78   2  0.825540754
## 1419                 crossing the  11   11   2  3.285817340
## 1420                     issue is   10   10   2  2.251327034
## 1421                   think that   11   11   2  2.178321412
## 1422                   of same-sex  21   21   2  1.616107310
## 1423                     to know   17   17   2  1.837161136
## 1424                superior court  22   22   2  9.957449250
## 1425                      if the   55   55   2  1.003270203
## 1426                     of texas   14   14   2  2.015645929
## 1427                     a civil   15   15   2  1.916681032
## 1428                  about their   12   12   2  1.995716842
## 1429                    to start   12   12   2  2.250143867
## 1430                     of this   47   47   2  1.047100705
## 1431                  to increase   15   15   2  1.949541149
## 1432                 execution of   14   14   2  1.997197088
## 1433                   that might   10   10   2  2.257363446
## 1434                      a news   16   16   2  1.832982834
## 1435              the possibility   13   13   2  3.737190282
## 1436              the researchers   15   15   2  2.265951822
## 1437                   without a   17   17   2  1.751929657
## 1438                  the results   13   12   2  2.638533735
## 1439                      was no   11   11   2  2.074918666
## 1440                    with some   12   12   2  1.983878625
```

```
## 1441             to legalize   10   10   2  4.810051425
## 1442             to recover    10   10   2  4.810051425
## 1443                much of     19   19   2  1.667919723
## 1444                 is just    12   12   2  1.999269472
## 1445                the truth   12   12   2  3.149320056
## 1446            the agency's    14   14   2  4.060037405
## 1447                the bush    21   21   2  1.746132335
## 1448                about it    14   14   2  1.827103461
## 1449                to force    13   13   2  2.096290632
## 1450                   get a    22   22   2  1.513790118
## 1451                the full    15   15   2  2.222461792
## 1452                of being    23   23   2  1.494239870
## 1453                the hands   13   13   2  2.567069854
## 1454                  make a    19   19   2  1.628578481
## 1455                close to    11   11   2  2.300280079
## 1456                in order    13   13   2  1.993322541
## 1457                status of   10   10   2  2.406017680
## 1458                he should   10   10   2  2.138200374
## 1459                 say the    31   31   2  1.322354091
## 1460              the result   15   13   2  2.180784178
## 1461                  at all    12   12   2  1.943504363
## 1462                murder in   13   13   2  1.959522794
## 1463                  it had    14   14   2  1.795589835
## 1464                money to    14   14   2  1.954301778
## 1465                the board   20   20   2  1.757290695
## 1466               of america   15   15   2  1.857699995
## 1467               between a    16   16   2  1.758260159
## 1468                of human    12   12   2  2.116048462
## 1469              the average   14   14   2  2.292291888
## 1470              charges of    11   11   2  2.221393603
## 1471                  a deal    11   11   2  2.191021207
## 1472                  it's a    22   22   2  1.475743158
## 1473                 hope to    11   11   2  2.243917656
## 1474              from other   13   13   2  1.843571065
## 1475                   be no   10   10   2  2.095622528
## 1476                 that is    59   57   2  0.880967971
## 1477                north of    12   12   2  2.088996880
## 1478                time for    13   13   2  1.860225130
## 1479                found in    15   15   2  1.775952425
## 1480                  to run    10   10   2  2.376314655
## 1481                 a large    10   10   2  2.288710647
## 1482                the only    38   38   2  1.175165745
## 1483              address the   14   14   2  2.082671785
## 1484             a political    15   15   2  1.802223499
## 1485                of three    19   19   2  1.592039163
## 1486                 the FBI    15   15   2  4.463274752
## 1487               a problem    11   11   2  2.148900287
## 1488                the time    47   47   2  1.043062788
## 1489                 held in    10   10   2  2.203748864
## 1490                the main    11   11   2  3.376034406
## 1491                for them    17   17   2  1.600458193
## 1492                  in all    27   24   2  1.293117390
## 1493             execution in   11   11   2  2.066242422
## 1494                making a    11   11   2  2.085481609
```

```
## 1495              called the    20    20    2  1.623641411
## 1496             whether the    27    27    2  1.365244575
## 1497             the council    20    20    2  1.679562037
## 1498            the internet    11    11    2  2.829471029
## 1499              the middle    12    12    2  2.490040004
## 1500             of marriage    28    28    2  1.273427556
## 1501            the proposed    19    19    2  1.727625223
## 1502                said his    19    19    2  1.485103352
## 1503                  find a    11    11    2  2.068809855
## 1504              new york's    15    15    2  9.213141434
## 1505                 way for    10    10    2  2.065124755
## 1506                  a drug    11    11    2  2.088868126
## 1507           the courtroom    11    11    2  2.729382653
## 1508              failing to    12    12    2  5.519454410
## 1509                voted to    12    12    2  2.022719174
## 1510              tuesday to    14    14    2  1.846471998
## 1511                are more    14    14    2  1.704708759
## 1512              to support    20    20    2  1.509579938
## 1513                in north    10    10    2  2.150232491
## 1514               for those    14    14    2  1.722919467
## 1515               million in    18    18    2  1.546440861
## 1516              the planet    11    11    2  2.638405957
## 1517              the verdict    11     9    2  2.638405957
## 1518            according to   125   125    2  8.940824115
## 1519                   up on    15    15    2  1.643857720
## 1520                 here in    13    13    2  1.823837664
## 1521                are also    10    10    2  1.982796879
## 1522                a report    16    16    2  1.651444607
## 1523                unable to    11    11    2  5.435930816
## 1524           companies that    10    10    2  2.028413050
## 1525              the matter    13    13    2  2.173131143
## 1526                  she is    13    12    2  1.747732763
## 1527               the globe    10    10    2  3.284993828
## 1528                years in    20    20    2  1.440853700
## 1529                the cost    14    14    2  2.035542534
## 1530                 and his    50    50    2  0.902375813
## 1531             the outcome    10    10    2  3.485669441
## 1532                 with no    10    10    2  1.955191831
## 1533               that many    15    15    2  1.626714773
## 1534                is about    19    19    2  1.432179059
## 1535           new hampshire    11    11    2  8.908709304
## 1536                   is an    25    25    2  1.247108612
## 1537             the court's    14    14    2  1.998496344
## 1538            the minnesota    14    14    2  1.998496344
## 1539                 is also    14    14    2  1.653721404
## 1540                   had a    33    33    2  1.099949716
## 1541             penalty for    11    11    2  1.873771040
## 1542                   to 20    11    11    2  2.021764092
## 1543                 has had    11    11    2  1.834619536
## 1544               house and    17    17    2  1.531122199
## 1545               the judge    36    36    2  1.115688950
## 1546              the airport    11    11    2  2.406589590
## 1547       the administration    23    21    2  1.432804182
## 1548                in effect    10     4    2  2.040001154
```

```
## 1549                  what was   11   11   2   1.826701717
## 1550                 of justice   13   13   2   1.797738687
## 1551                  the media   12   11   2   2.193749599
## 1552                    year in   15   15   2   1.618815587
## 1553               the european   10   10   2   2.638342076
## 1554               the so-called  10    9   2   2.638342076
## 1555               children and   13   13   2   1.740483649
## 1556               of immigrants  35   35   2   1.053670887
## 1557                    to deal   11   11   2   1.977791474
## 1558                    bill is   12   12   2   1.738918418
## 1559                 became the   13   13   2   1.920141866
## 1560                 of tobacco   33   33   2   1.079977165
## 1561                the program   24   22   2   1.365961297
## 1562                 seeking to   11   11   2   1.959609195
## 1563                  until the   19   19   2   1.518340498
## 1564             immigrants and   23   23   2   1.276919458
## 1565                 of florida   15   15   2   1.623199938
## 1566                     put to   15   15   2   1.634684935
## 1567              decision that   10   10   2   1.912150569
## 1568                   guns and   11   11   2   1.863109518
## 1569                 protect the   12   12   2   1.999689543
## 1570                    or more   10   10   2   1.856794359
## 1571                  billion in   11   11   2   1.856249616
## 1572               emissions of   15   15   2   1.606344692
## 1573                   the head   12   12   2   2.096101293
## 1574               the decision   29   28   2   1.207243340
## 1575                the subject   10   10   2   2.463978853
## 1576                 filled with   12   12   2   8.468822170
## 1577                  the court   65   64   2   0.780720507
## 1578                  of mexico   10   10   2   2.003918798
## 1579                   here and   12   12   2   1.755576089
## 1580                 the former   22   22   2   1.402075821
## 1581                    to look   10   10   2   2.024859803
## 1582                   and it's   11   11   2   1.859783929
## 1583                   laws and   11   11   2   1.828609281
## 1584                  of people   35   35   2   1.019825706
## 1585                 a criminal   11   11   2   1.863723382
## 1586                      a gun   33   33   2   1.046025764
## 1587                  the final   12   12   2   2.050634001
## 1588                 the amount   10   10   2   2.387012894
## 1589                 benefits of   10   10   2   1.974361049
## 1590                 health and   15   15   2   1.544920656
## 1591                 from death   13   13   2   1.600593492
## 1592                      in my   12   12   2   1.728404778
## 1593                  after his   10   10   2   1.797857903
## 1594                 leaders and   10   10   2   1.883181040
## 1595                 and former   12   12   2   1.727244782
## 1596                that global   11   11   2   1.742793453
## 1597               to death for   13    0   3   4.080655733
## 1598               said that the   21    0   3   1.922702228
## 1599                a republican   15   13   2   1.529381042
## 1600        the immigration and   26    0   3   1.768060232
## 1601                   the suit   12   11   2   1.965466357
## 1602                  that some   16   16   2   1.433990369
```

```
## 1603            to appeal   11    11   2  1.837528900
## 1604           citizens to   10    10   2  1.932361155
## 1605              out that   16    16   2  1.428245147
## 1606               a young   10    10   2  1.887494313
## 1607                in our   18    18   2  1.368314254
## 1608               days of   10    10   2  1.895092764
## 1609                or not   12    12   2  1.609059066
## 1610             the mayor   14    14   2  1.743110307
## 1611             the rules   16    16   2  1.602485963
## 1612           people with   11    11   2  1.679494274
## 1613            this state   10    10   2  1.750214780
## 1614                to say   21    21   2  1.276086255
## 1615            support of   16    16   2  1.460364212
## 1616          the colorado   12    12   2  1.925456105
## 1617                at its   10    10   2  1.752526213
## 1618                like a   18    18   2  1.351780254
## 1619               a group   16    16   2  1.446689246
## 1620             the trial   28    26   2  1.148672366
## 1621              the lead   14    14   2  1.715325825
## 1622              but this   12    12   2  1.589045185
## 1623                 us to   12    12   2  1.697767963
## 1624              said in an   12     0   3  3.558585115
## 1625             the story   12    12   2  1.886984906
## 1626              one of the  120     0   3  3.593708355
## 1627            the florida   23    23   2  1.268625291
## 1628              cases in   10    10   2  1.801219959
## 1629                in her   21    21   2  1.231828767
## 1630                a life   16    16   2  1.425547541
## 1631           the process   17    17   2  1.502607553
## 1632       the environmental   14    14   2  1.688292234
## 1633               all the   44    44   2  0.872833135
## 1634               the old   10    10   2  2.127482028
## 1635              near the   11    11   2  1.882882058
## 1636               off the   17    17   2  1.451073854
## 1637                 me to   10    10   2  1.848081837
## 1638             the third   11    11   2  1.967183588
## 1639             the words   11    10   2  1.967183588
## 1640                day of   14    14   2  1.520347112
## 1641               week to   13    13   2  1.588840040
## 1642           between the   22    22   2  1.249356764
## 1643                on may   10    10   2  1.713789599
## 1644          and children   16    16   2  1.398619412
## 1645              judge in   16    16   2  1.384169460
## 1646              that you   12    12   2  1.575759859
## 1647           lawyers and   10    10   2  1.760280498
## 1648             while the   28    28   2  1.087956534
## 1649              the risk   13    13   2  1.729579624
## 1650               a state   35    35   2  0.932065491
## 1651          the southern   11    11   2  1.923693559
## 1652              from his   16    16   2  1.336884660
## 1653               case of   17    17   2  1.346779379
## 1654               said in   50    50   2  0.770231954
## 1655                  up a   24    24   2  1.116401582
## 1656            is not the   13     0   3  2.154871152
```

```
## 1657                        by an    17    17    2   1.288185321
## 1658                  evidence of    10    10    2   1.769896522
## 1659                    for years    16    14    2   1.332531366
## 1660                  in addition    22    13    2   7.549719058
## 1661                       in may    17    15    2   1.317601405
## 1662                     have not    16    16    2   1.320094760
## 1663                the execution    20    20    2   1.308847700
## 1664                       made a    14    14    2   1.451161705
## 1665                    state and    28    28    2   1.015516613
## 1666                    year that    10    10    2   1.679622916
## 1667              the governor's    10    10    2   4.584296482
## 1668                   and their    38    38    2   0.871649250
## 1669                     laws in    10    10    2   1.707546844
## 1670                     and we    19    19    2   1.230346973
## 1671               for immigrants    17    17    2   1.271857630
## 1672                     the job    15    15    2   1.518614466
## 1673                   the front    11    11    2   1.842005692
## 1674                      of gay    16    16    2   1.350636788
## 1675               of cigarettes    12    12    2   1.569334458
## 1676                   to smoke    16    15    2   1.354274916
## 1677                     as one    11    11    2   1.551004623
## 1678             the immigration    66    66    2   0.676972746
## 1679                 states that    14    14    2   1.391940896
## 1680                 similar to    10    10    2   1.733643023
## 1681                   who will    12    12    2   1.478802011
## 1682                  people in    24    24    2   1.073123915
## 1683                    that i    12    12    2   1.499511066
## 1684                 man and a    16     0    3   3.171968640
## 1685                  of global    16    16    2   1.336492813
## 1686                 though the    13    13    2   1.579469442
## 1687                 leaders of    11    11    2   1.622793603
## 1688            of undocumented    10    10    2   1.708128871
## 1689                   court of    31    31    2   0.948567686
## 1690                    to use    15    15    2   1.384170866
## 1691                 is not a    15     0    3   2.063664501
## 1692                 the floor    10    10    2   1.921610303
## 1693               the shooting    14    14    2   1.539842719
## 1694          the united states   250     0    3   2.463683336
## 1695                 the early    12    12    2   1.682860041
## 1696                    him in    12    12    2   1.497376019
## 1697               convicted in    10    10    2   1.642631088
## 1698                    a u.s    26    23    2   1.018534577
## 1699                  now that    11    11    2   1.533671308
## 1700               this is the    18     0    3   2.313689389
## 1701                of georgia    10    10    2   1.676620743
## 1702                  that are    32    32    2   0.902089714
## 1703                   for an    25    25    2   1.016597678
## 1704                 and most    19    19    2   1.186719983
## 1705                 the 1996    10    10    2   1.876143011
## 1706                the price    10     9    2   1.876143011
## 1707                reduce the    12    12    2   1.611250920
## 1708                 action to    10    10    2   1.680985058
## 1709                the center    19    18    2   1.271760552
## 1710              some of the    36     0    3   3.436239590
```

```
## 1711                 in public   17        17       2  1.244327459
## 1712                  the sept   12        12       2  1.652549774
## 1713              of greenhouse  11        11       2  1.571248292
## 1714              the ordinance  10        10       2  1.832652981
## 1715                   came in   10        10       2  1.601580936
## 1716                 of health   17        17       2  1.239721970
## 1717                     of my   12        12       2  1.479851550
## 1718                 change in   11        11       2  1.517928759
## 1719                to america   12        12       2  1.487794916
## 1720               the lawsuit   14        13       2  1.473136590
## 1721               supposed to   18        18       2  7.010965862
## 1722                    in its   22        22       2  1.066609120
## 1723          court of appeals   25         0       3  7.677403796
## 1724                   so that   12        12       2  1.412707089
## 1725                the ballot   10        10       2  1.790975367
## 1726              the missouri   10        10       2  1.790975367
## 1727            the washington   17        17       2  1.303547281
## 1728                saying the   12        12       2  1.544544984
## 1729                   in what   19        19       2  1.137586256
## 1730                   in many   17        17       2  1.202804531
## 1731                  behalf of  15        15       2  6.939550511
## 1732                a decision   14        14       2  1.331482674
## 1733              president of   23        23       2  1.036692628
## 1734                    just a   17        17       2  1.191302556
## 1735                     to me   11         9       2  1.519830192
## 1736                   tens of   14        14       2  6.872701730
## 1737                 have said   18        18       2  1.128307127
## 1738               members and   10        10       2  1.542835632
## 1739                 a man and   20         0       3  2.779638180
## 1740               because the   34        34       2  0.856853699
## 1741              and lesbians   11        11       2  6.808873284
## 1742                  issue in   11        11       2  1.452561895
## 1743                   the two   39        39       2  0.808116636
## 1744                   of life   15        15       2  1.259869962
## 1745                   but his   13        13       2  1.295500485
## 1746                 of murder   14        14       2  1.294995621
## 1747               despite the   10        10       2  1.639995405
## 1748               addition to   13        13       2  6.695174487
## 1749              the majority   12        12       2  1.513412380
## 1750                  down the   19        19       2  1.142224389
## 1751                   by law    10         9       2  1.455265097
## 1752              the world's   26        26       2  6.609784524
## 1753                   of any   17        17       2  1.158920241
## 1754               the couple   10         8       2  1.675447726
## 1755             immigrants to  26        26       2  0.935454866
## 1756                the legal   32        32       2  0.875827186
## 1757                for legal   10        10       2  1.458246564
## 1758            of immigration  34        34       2  0.808969984
## 1759     the federal government 25         0       3  3.457773981
## 1760                  the ban   32        31       2  0.870559094
## 1761                a handful   10        10       2  6.632695301
## 1762               of business 10        10       2  1.505938459
## 1763                the obama   16        16       2  1.259794677
## 1764                is one of  18         0       3  3.235922834
```

```
## 1765              a time when   11      0   3  7.311474878
## 1766              the defense   14     14   2  1.351746221
## 1767              amendment to   10     10   2  1.508632874
## 1768              the companies  23     23   2  1.029272604
## 1769              including a   12     12   2  1.345706334
## 1770                  get the   28     28   2  0.907965360
## 1771              in order to   10      0   3  4.290818272
## 1772                  tend to   11     11   2  6.534547847
## 1773                 the line   11     11   2  1.539680557
## 1774                   of gun   30     30   2  0.848400155
## 1775                  case in   12     12   2  1.323209609
## 1776                  for its   13     13   2  1.248795391
## 1777             congress and   10     10   2  1.445361956
## 1778                 that was   33     33   2  0.789978011
## 1779               smoking is   10     10   2  1.409996115
## 1780                 power to   10     10   2  1.480252689
## 1781                   from a   49     49   2  0.652287420
## 1782             the military   10     10   2  1.605233631
## 1783            the potential   10     10   2  1.605233631
## 1784       the tobacco industry  39      0   3  2.513299421
## 1785                    of an   46     46   2  0.673645225
## 1786                   law is   10     10   2  1.399881113
## 1787                the party   14     14   2  1.314348852
## 1788                for their   20     20   2  1.001154098
## 1789                and women   10     10   2  1.444260051
## 1790                   a week   15     15   2  1.178657375
## 1791              country and   10     10   2  1.428615788
## 1792                  is that   48     48   2  0.650073787
## 1793                the crime   16     16   2  1.214996318
## 1794                   when a   23     23   2  0.942837432
## 1795                  in part   10      8   2  1.432849546
## 1796           the smoking ban  14      0   3  2.770521896
## 1797                  let the   10     10   2  1.533208015
## 1798                   in any   14     14   2  1.202544512
## 1799               the reason   12     12   2  1.414546872
## 1800                 make the   23     23   2  0.969025250
## 1801                   a home   12     12   2  1.297421299
## 1802                 state of   32     32   2  0.789697377
## 1803                 to study   13     13   2  1.245742962
## 1804               and three   12     12   2  1.275024240
## 1805                 has said   14     14   2  1.146762288
## 1806                 asked to   10     10   2  1.412635676
## 1807                  the day   24     24   2  0.943272292
## 1808              marriage in   13     13   2  1.208903399
## 1809                 that his   26     26   2  0.847717076
## 1810                 year and   11     11   2  1.310745613
## 1811                 the need   18     18   2  1.096145115
## 1812             the need for   10      0   3  2.655684005
## 1813                 trial of   10     10   2  1.395600741
## 1814                 before a   14     14   2  1.164142367
## 1815                   of us   10     10   2  1.394670247
## 1816        the international   15     15   2  1.197205497
## 1817                  the rev   15     10   2  6.072722499
## 1818                 the high   19     19   2  1.049831576
```

```
## 1819              to same-sex   15   15   2   1.128791548
## 1820             of execution   10   10   2   1.371538367
## 1821                and many    16   16   2   1.071694226
## 1822                  to act    11   11   2   1.311685028
## 1823             and federal    20   20   2   0.955829618
## 1824                 in some    18   18   2   1.000950900
## 1825               to life in   16    0   3   2.446299639
## 1826        state and federal   11    0   3   6.306874040
## 1827             of immigrant   13   13   2   1.185759111
## 1828                  a vote    12   12   2   1.227037338
## 1829                 him and    10   10   2   1.326103121
## 1830                 said at    15   15   2   1.064039748
## 1831                 with it    15   15   2   1.064065700
## 1832                 be more    10   10   2   1.289651495
## 1833                death for   15   15   2   1.063624814
## 1834                 a death    30   30   2   0.767598561
## 1835                 of some    21   21   2   0.913425753
## 1836              the system    20   18   2   0.981487201
## 1837               a program   10   10   2   1.319012871
## 1838             the economic   11   11   2   1.354912026
## 1839                stop the    11   11   2   1.320815455
## 1840                and they    32   32   2   0.733185632
## 1841                  of two    20   20   2   0.931324576
## 1842               that were    19   19   2   0.936668112
## 1843         president of the   15    0   3   2.238097537
## 1844                  set to    10   10   2   1.325046574
## 1845              the general   18   18   2   1.029051367
## 1846              and health    13   13   2   1.136903887
## 1847             consider the   10   10   2   1.370022400
## 1848         immigrants in the   20    0   3   1.915040124
## 1849                  a case    17   17   2   0.993008696
## 1850            including the    19   19   2   0.972948550
## 1851                and more    29   29   2   0.756840223
## 1852                  a high    10   10   2   1.289870443
## 1853         a spokesman for    17    0   3   4.782821180
## 1854                   of mr    11   11   2   1.229180001
## 1855              of climate    16   16   2   1.014916811
## 1856                for this    17   17   2   0.962240334
## 1857            the fact that   22    0   3   3.554268718
## 1858                that can    11   11   2   1.188985428
## 1859               penalty in   11   11   2   1.201957359
## 1860               the fight    12   12   2   1.242038264
## 1861                asked the   13   13   2   1.163626483
## 1862                that this   19   19   2   0.906709620
## 1863               the panel    10   10   2   1.365243616
## 1864                the real    10   10   2   1.365243616
## 1865             will have to   11    0   3   1.689778774
## 1866             in favor of    23    0   3   6.487133031
## 1867                rights to   12   12   2   1.157930883
## 1868                  on all    10   10   2   1.218128102
## 1869                in court    24   24   2   0.805476248
## 1870                to think    11   11   2   1.206829666
## 1871                 only to    18   18   2   0.934590146
## 1872              the coming    10   10   2   1.338921389
```

```
## 1873                    was an   13   13   2   1.060368944
## 1874                     on an   18   18   2   0.906785303
## 1875               companies to  11   11   2   1.191720238
## 1876                     be an   12   12   2   1.098021470
## 1877              by immigration 10   10   2   1.197713625
## 1878                    as his   14   14   2   1.018108091
## 1879                 the union   12   12   2   1.203188594
## 1880                 on tobacco  11   11   2   1.142264033
## 1881            a federal judge  19    0   3   2.490596174
## 1882                   time in   12   12   2   1.105655208
## 1883                 to report   12   12   2   1.125509298
## 1884                 million to  16   16   2   0.968458024
## 1885                   been a    23   23   2   0.798760030
## 1886               put to death  12    0   3   3.874680520
## 1887                   with an   13   13   2   1.036333943
## 1888                   the man   19   18   2   0.922143693
## 1889               it comes to   13    0   3   3.226266218
## 1890                 this is a   24    0   3   1.847839208
## 1891                   the st   12   10   2   1.165791224
## 1892            the authorities  11   11   2   1.219846232
## 1893                     to $    23   23   2   0.794125259
## 1894                 to those    18   18   2   0.896914946
## 1895          said yesterday that 10   0   3   2.656439067
## 1896                 of such    12   12   2   1.093980880
## 1897            of marriage act  10    0   3   6.247652387
## 1898                 court in    20   20   2   0.837218200
## 1899                 a month    10   10   2   1.189735079
## 1900                  a city    15   15   2   0.969923804
## 1901             to ban smoking  12    0   3   2.359438017
## 1902        the american people  10    0   3   3.398576270
## 1903             the community   12   12   2   1.147603987
## 1904                 a study    10   10   2   1.181109617
## 1905            the population   10   10   2   1.263871449
## 1906                   not a    38   38   2   0.604138808
## 1907               become the   13   13   2   1.072620725
## 1908                 a system   10   10   2   1.172557838
## 1909             and a woman    14    0   3   2.288383685
## 1910                  why the   10   10   2   1.211235180
## 1911               from their   10   10   2   1.120370537
## 1912                with this   10   10   2   1.110885577
## 1913                   to gay   13   13   2   1.011068286
## 1914               the family   16   16   2   0.947488134
## 1915                  for one   12   12   2   1.014251973
## 1916                  at his    12   12   2   1.008478338
## 1917               and murder   10   10   2   1.122374018
## 1918           the white house  34    0   3   2.327293023
## 1919                the money   14   14   2   1.001652050
## 1920       on illegal immigration 13  0   3   2.507814217
## 1921                  law in    15   15   2   0.909518391
## 1922                 house in   11   11   2   1.058739247
## 1923             marriage and   11   11   2   1.056241143
## 1924              in front of   13    0   3   7.047716726
## 1925                 of what    18   18   2   0.832137535
## 1926                 got the    10   10   2   1.157636372
```

```
## 1927                the power   14    14    2  0.988313600
## 1928                   of new   33    33    2  0.610080708
## 1929                the order   14    12    2  0.975150597
## 1930                the times   14    14    2  0.975150597
## 1931             and that the   14     0    3  1.060285596
## 1932               of american   14    14    2  0.926996353
## 1933                  of guns   10    10    2  1.093774156
## 1934                   over a   17    17    2  0.830795602
## 1935               the end of   30     0    3  5.569293219
## 1936                  time of   14    14    2  0.918550352
## 1937         the presidential   10    10    2  1.150087971
## 1938             to deal with   10     0    3  3.702159100
## 1939        the death penalty  168     0    3  1.983672675
## 1940               the effort   11    11    2  1.082340811
## 1941                to change   15    15    2  0.881843376
## 1942                 rights of   10    10    2  1.072272709
## 1943              the georgia   12    12    2  1.028786019
## 1944                 the deal   10    10    2  1.128805654
## 1945                 the news   16    16    2  0.882570994
## 1946            president and   14    14    2  0.885420302
## 1947          circuit court of   14     0    3  5.097432394
## 1948                 says the   16    16    2  0.853487537
## 1949               two of the   10     0    3  4.841183832
## 1950           the high court   12     0    3  2.110814689
## 1951                 a ban on   10     0    3  2.915240440
## 1952               was one of   17     0    3  2.268538195
## 1953           republican of   11    11    2  0.987267372
## 1954                   and he   30    30    2  0.595492713
## 1955               at a news   11     0    3  2.665448579
## 1956                  put the   13    13    2  0.933611009
## 1957                   a time   16    16    2  0.809691804
## 1958              a bill that   14     0    3  1.379371240
## 1959              officials in   11    11    2  0.962128547
## 1960                   by his   12    12    2  0.901783928
## 1961           state and local   10     0    3  4.789518159
## 1962                 the plan   13    11    2  0.930124273
## 1963                  like to   15    15    2  0.828763354
## 1964                in federal   16    16    2  0.792766879
## 1965                allow the   11    11    2  0.989596048
## 1966                  city of   13    13    2  0.878627310
## 1967                  and now   12    11    2  0.902541926
## 1968              many of the   16     0    3  1.906591980
## 1969                   to any   14    14    2  0.836596389
## 1970                  down to   10    10    2  0.983293546
## 1971              the climate   25    25    2  0.643403378
## 1972                  a way to   18     0    3  1.392196509
## 1973              who came to   11     0    3  1.733770845
## 1974              the killing   10    10    2  1.028697603
## 1975                 the move   10    10    2  1.028697603
## 1976           bill that would   16     0    3  2.274848055
## 1977               most of the   31     0    3  4.372347842
## 1978               to control   12    12    2  0.892978635
## 1979                  bill to   17    17    2  0.745487009
## 1980                 in other   19    19    2  0.696015690
```

43

```
## 1981                            up the    32    32    2    0.551412392
## 1982               the debate over    11     0    3    1.743372171
## 1983                        and its    17    17    2    0.731505939
## 1984    immigration and naturalization    36     0    3    4.925597229
## 1985                    to congress    11    11    2    0.907849538
## 1986                     the clinton    11    11    2    0.945335606
## 1987                        but as    11    11    2    0.872198218
## 1988                       on this    12    11    2    0.837193659
## 1989                    tobacco and    16    16    2    0.730892793
## 1990                      the face    11    11    2    0.929457338
## 1991                  in recent years    12     0    3    2.326511455
## 1992                      a public    13    13    2    0.812764638
## 1993                        in how    10    10    2    0.917905826
## 1994                         and a   158   156    2    0.233670561
## 1995                     the other    43    43    2    0.456988906
## 1996                  the settlement    10    10    2    0.955250460
## 1997                    the police    16    15    2    0.745703815
## 1998                  in the face    10     0    3    4.693720918
## 1999    and naturalization service    34     0    3    4.713069859
## 2000                    that have    25    25    2    0.565091770
## 2001                      is the   136   136    2    0.250411274
## 2002                   and public    12    12    2    0.816072523
## 2003                  to the editor    24     0    3    4.751456568
## 2004                     first of    15    15    2    0.729787439
## 2005                between a man    13     0    3    2.904747452
## 2006                    of federal    18    18    2    0.665147016
## 2007                     the texas    11    11    2    0.883280540
## 2008                   without the    15    15    2    0.743035542
## 2009            the justice department    22     0    3    2.162565100
## 2010                  the committee    10    10    2    0.920454506
## 2011                       year to    11    11    2    0.840113549
## 2012                         it as    12    12    2    0.781388385
## 2013                      that had    13    13    2    0.755208437
## 2014                     the trump    11    11    2    0.868349970
## 2015                    as many as    16     0    3    2.156026034
## 2016                      keep the    10    10    2    0.885201289
## 2017                      as well as    39     0    3    5.393530540
## 2018                       is more    11    11    2    0.797106584
## 2019                       where a    10    10    2    0.845909677
## 2020               gay and lesbian    18     0    3    6.769657789
## 2021                    the new law    10     0    3    1.317253389
## 2022                   for tobacco    10    10    2    0.826411944
## 2023                      take the    15    15    2    0.704838690
## 2024             a news conference    15     0    3    3.326783350
## 2025                        a ban    12    12    2    0.763766169
## 2026                       see the    10    10    2    0.859046301
## 2027               in an interview    12     0    3    3.183690863
## 2028                       two of    17    17    2    0.638151666
## 2029                  the campaign    13    13    2    0.757851976
## 2030        the clinton administration    10     0    3    4.338860260
## 2031                   at gun shows    16     0    3    4.109199979
## 2032                  the people who    10     0    3    1.123320257
## 2033                        out a    14    14    2    0.684023511
## 2034                    of smoking    18    18    2    0.606033688
```

```
## 2035                   a right   10    10    2  0.805897836
## 2036             for immigration   12    12    2  0.718011598
## 2037                    only a   11    11    2  0.756806276
## 2038                    are at   10    10    2  0.778635354
## 2039                  the four   12    11    2  0.751055901
## 2040                 the woman   11    11    2  0.783162648
## 2041            to the federal   10     0    3  1.280068659
## 2042                 people to   22    22    2  0.529575988
## 2043                    a law   19    19    2  0.565521656
## 2044             many of them   13     0    3  3.557463423
## 2045                   and it   34    34    2  0.422630192
## 2046                    of $   17    17    2  0.596683879
## 2047           because of the   19     0    3  1.453757742
## 2048             life in prison   28     0    3  3.600844367
## 2049             ban on smoking   12     0    3  2.274219782
## 2050                 where the   18    18    2  0.585726194
## 2051             all over the   10     0    3  1.602050234
## 2052               have to be   13     0    3  3.461735289
## 2053                  use the   12    12    2  0.715561049
## 2054                  and two   13    13    2  0.667520071
## 2055          in prison without   13     0    3  3.665092075
## 2056                 the local   12    12    2  0.715323060
## 2057                years and   10    10    2  0.745342925
## 2058                 against a   11    11    2  0.708919320
## 2059                as part of   24     0    3  3.414967283
## 2060                to stay in   11     0    3  2.157787649
## 2061                   in one   16    16    2  0.581109539
## 2062         legislation that would   10     0    3  3.530500530
## 2063                    is at   16    16    2  0.572550141
## 2064                as long as   10     0    3  3.828924741
## 2065                 but that   20    20    2  0.513710455
## 2066               do not have   11     0    3  2.126462075
## 2067              was the first   19     0    3  1.587544694
## 2068            the report said   10     0    3  3.319710407
## 2069                   up and   13    13    2  0.626205732
## 2070                  a legal   11    11    2  0.680213940
## 2071               a year ago   10     0    3  3.345991636
## 2072          the mexican border   10     0    3  3.349017756
## 2073                and state   21    21    2  0.486047180
## 2074                  the gun   42    42    2  0.353762075
## 2075      and customs enforcement   29     0    3  3.937070563
## 2076              as a result   14     0    3  3.666704868
## 2077            smoking in bars   10     0    3  3.668926674
## 2078                 the home   15    15    2  0.587335310
## 2079                 and some   13    12    2  0.600029149
## 2080               to the u.s   22     0    3  0.906731304
## 2081                at a time   11     0    3  1.206870319
## 2082         members of congress   14     0    3  3.220547266
## 2083           a federal appeals   11     0    3  1.457233898
## 2084       the bush administration   15     0    3  3.223060930
## 2085                 death in   17    17    2  0.516848043
## 2086             the hands of   13     0    3  4.331403396
## 2087               the health   19    19    2  0.505707740
## 2088                   of how   10    10    2  0.669435142
```

```
## 2089           the florida supreme   13    0    3  1.973015852
## 2090              seek the death     12    0    3  3.189845287
## 2091              the first time     27    0    3  2.985468639
## 2092                 a new study    12    0    3  1.623242765
## 2093                      what a   14   14    2  0.548849271
## 2094                     law and   11   11    2  0.615217538
## 2095                   a path to   14    0    3  1.819853746
## 2096             sentenced to death  40    0    3  2.978751612
## 2097                       the $   31   31    2  0.376257158
## 2098                    states to   15   15    2  0.524210123
## 2099               of the people   11    0    3  0.911141752
## 2100                 time of the   10    0    3  1.197408106
## 2101       immigration and customs  30    0    3  4.060813617
## 2102                     a house   10   10    2  0.613134795
## 2103                has become a   12    0    3  1.107432093
## 2104                   to have a   15    0    3  0.618757866
## 2105                 marriage to   11   11    2  0.585606601
## 2106                  the use of   13    0    3  2.813842296
## 2107                     was the   83   83    2  0.218483669
## 2108               the cigarette   12   12    2  0.573718934
## 2109             was convicted of   21    0    3  0.863910696
## 2110                  and those   10   10    2  0.591959278
## 2111                 the cost of   14    0    3  2.834830315
## 2112                is the first   14    0    3  2.689216025
## 2113                  it was the   14    0    3  0.629189684
## 2114                 the appeals   12   12    2  0.553607915
## 2115                 the risk of   10    0    3  1.786724696
## 2116              when it comes   12    0    3  2.882833892
## 2117                      not to   39   39    2  0.299237600
## 2118            path to citizenship  14    0    3  2.929597571
## 2119                on death row   35    0    3  3.721651685
## 2120       department of homeland   17    0    3  3.016810842
## 2121                    the back   13   13    2  0.517783232
## 2122                to seek the   14    0    3  1.363131527
## 2123                     how the   17   17    2  0.444608814
## 2124               in response to   10    0    3  2.725756646
## 2125                in this case   11    0    3  1.300668240
## 2126                     out the   24   24    2  0.371474118
## 2127              support for the   12    0    3  0.989505558
## 2128            and lesbian couples  12    0    3  3.122009047
## 2129                  end of the   16    0    3  1.582870083
## 2130                      of law   18   18    2  0.414122361
## 2131                  a group of   11    0    3  1.001891861
## 2132                   those in   10   10    2  0.543122801
## 2133                   a million   10   10    2  0.542612320
## 2134                     the use   13   13    2  0.491451163
## 2135                  that more   12   12    2  0.486019543
## 2136                 officials to   11   11    2  0.510122790
## 2137                 of the new   21    0    3  0.627852205
## 2138                 people and   13   13    2  0.465795900
## 2139           the national rifle   37    0    3  1.760484349
## 2140                    and her   11   11    2  0.499484602
## 2141              people who are   11    0    3  2.418589340
## 2142                     in u.s   13   13    2  0.457560031
```

46

```
## 2143               one that   10    10   2  0.506387871
## 2144          on behalf of   15     0   3  3.991038962
## 2145             up to the   11     0   3  0.634843277
## 2146        of thousands of   18     0   3  2.537951478
## 2147         in addition to   13     0   3  3.914479268
## 2148               and the  319   319   2  0.092029735
## 2149                a more   21    21   2  0.339230669
## 2150              of a man   12     0   3  0.867541270
## 2151            the county   21    21   2  0.347216229
## 2152           sides of the   11     0   3  1.500724430
## 2153            rest of the   13     0   3  2.315809630
## 2154               a court   18    18   2  0.357704884
## 2155            the time of   10     0   3  0.903520047
## 2156                out in   10    10   2  0.469036238
## 2157            this is not   21     0   3  1.027994710
## 2158          in the future   12     0   3  2.203595063
## 2159        was found guilty   10     0   3  2.345206761
## 2160            a result of   10     0   3  2.231679726
## 2161            more than a   24     0   3  0.762348144
## 2162          in the senate   24     0   3  2.109437599
## 2163          and president   10    10   2  0.460741866
## 2164            the rest of   19     0   3  2.237054004
## 2165        increase in the   10     0   3  1.329120929
## 2166               about a   18    18   2  0.335115092
## 2167               and one   14    14   2  0.378818946
## 2168        that the state   11     0   3  0.999187087
## 2169               were in   17    17   2  0.339360759
## 2170          over the next   16     0   3  1.232560953
## 2171       the university of   30     0   3  0.722804413
## 2172               to be in   10     0   3  0.598382997
## 2173          is expected to   33     0   3  0.886238721
## 2174                had to   21    21   2  0.297698356
## 2175               and all   12    11   2  0.388081920
## 2176          in the united   99     0   3  1.914616888
## 2177        chairman of the   12     0   3  2.053978097
## 2178           the right to   38     0   3  0.440614731
## 2179           of the most   24     0   3  0.787835709
## 2180         members of the   27     0   3  0.736242811
## 2181                out to   14    14   2  0.353514402
## 2182            the people   37    37   2  0.223465410
## 2183               of many   10    10   2  0.410973665
## 2184        argued that the   13     0   3  1.250649978
## 2185   the country illegally   11     0   3  1.174005763
## 2186             a new trial   12     0   3  0.840610441
## 2187       bars and restaurants   11     0   3  2.029871594
## 2188            the women   10     9   2  0.409437033
## 2189               and not   27    27   2  0.244250236
## 2190                but in   25    25   2  0.251648634
## 2191                  if a   13    13   2  0.345407075
## 2192          of president   11    11   2  0.374733584
## 2193          would have to   13     0   3  0.447559927
## 2194            came to the   10     0   3  0.612746303
## 2195            a series of   14     0   3  2.262498611
## 2196       spokesman for the   14     0   3  0.851406096
```

```
## 2197        on climate change   22     0   3   1.824465203
## 2198          wake of the        10     0   3   1.842684066
## 2199          out of the         37     0   3   0.381884048
## 2200              to a          184   184   2   0.089145429
## 2201       is scheduled to       15     0   3   0.775473489
## 2202        percent of the       44     0   3   0.519363319
## 2203         to bear arms        12     0   3   2.488382922
## 2204    convicted of killing     10     0   3   1.727943637
## 2205       the execution of      10     0   3   0.755754390
## 2206          that their        12    12   2   0.323438838
## 2207        for more than        18     0   3   0.816425876
## 2208         in the last        18     0   3   0.478738149
## 2209     to carry concealed      14     0   3   1.247249184
## 2210         is likely to        15     0   3   1.630645144
## 2211            what the        24    24   2   0.232015187
## 2212            of other        16    16   2   0.272320104
## 2213          and tobacco       13    13   2   0.296964527
## 2214         on the issue        21     0   3   0.518802670
## 2215        lawyers for the      10     0   3   0.601949740
## 2216         the california      10    10   2   0.340067901
## 2217           as much as        13     0   3   0.936662016
## 2218         food and drug       11     0   3   1.599069490
## 2219          an effort to       17     0   3   0.597613483
## 2220            a lot of        21     0   3   0.724408228
## 2221            and new        19    19   2   0.224844261
## 2222            those of        11    11   2   0.291690663
## 2223           the rights       13    13   2   0.274248885
## 2224         to make sure        10     0   3   1.503342632
## 2225         member of the       11     0   3   1.423122701
## 2226          by the u.s        12     0   3   0.790962171
## 2227         on gun control      11     0   3   1.159110797
## 2228            of time        10    10   2   0.238344149
## 2229            it's the        15    15   2   0.198090183
## 2230             of u.s        13    13   2   0.208956041
## 2231           the guns        10    10   2   0.239930322
## 2232          by the end        13     0   3   0.551112521
## 2233         for the first       17     0   3   0.610939543
## 2234          of the case        10     0   3   0.507932295
## 2235       of climate change     14     0   3   0.598345892
## 2236         believed to be      12     0   3   1.046384773
## 2237          who want to        11     0   3   0.450042287
## 2238          more than $        13     0   3   1.008632208
## 2239             to two        11    11   2   0.201209531
## 2240        a death sentence     17     0   3   0.351405701
## 2241          to carry a        12     0   3   0.491825205
## 2242         both sides of       11     0   3   0.630398575
## 2243          of the bill        11     0   3   0.399731782
## 2244    of homeland security     18     0   3   1.063263740
## 2245            court to        16    16   2   0.157305709
## 2246            state to        19    19   2   0.144643544
## 2247            to this        22    22   2   0.133929329
## 2248      of greenhouse gases     11     0   3   0.966206697
## 2249           the murder       11    11   2   0.186051386
## 2250         likely to be        18     0   3   0.863511065
```

48

```
## 2251                 for the death  14       0   3  0.303385392
## 2252          executive director of  13       0   3  0.494854332
## 2253                 to the state    15       0   3  0.241328376
## 2254               and immigration   15      14   2  0.150046376
## 2255                        do the   15      15   2  0.149569110
## 2256                       to them   12      12   2  0.158105938
## 2257                  the children   13      12   2  0.143672508
## 2258                         has a   22      22   2  0.105912956
## 2259              to the united     33       0   3  0.319774883
## 2260                  in the next   11       0   3  0.339737775
## 2261             sentenced to life  10       0   3  0.692477289
## 2262       greenhouse gas emissions  32       0   3  0.748758388
## 2263           the supreme court's  11       0   3  0.343927906
## 2264                 at least one   13       0   3  0.746522083
## 2265                  in the wake   13       0   3  0.695455844
## 2266          background checks for  14       0   3  0.628266588
## 2267                 by the state   12       0   3  0.265720891
## 2268             cruel and unusual   12       0   3  0.774783995
## 2269             the border patrol   14       0   3  0.565839236
## 2270                     the smoke   10      10   2  0.117753821
## 2271                      it was a   19       0   3  0.107444298
## 2272                  in the past   27       0   3  0.301891695
## 2273                from death row   11       0   3  0.577663727
## 2274                 in the nation  14       0   3  0.491594350
## 2275                to ensure that   11       0   3  0.252021429
## 2276                    we have to   13       0   3  0.110097446
## 2277               around the world  15       0   3  0.437542358
## 2278                 more likely to  14       0   3  0.184782311
## 2279                      the vote   10      10   2  0.087543166
## 2280                     to try to   17       0   3  0.208935575
## 2281               that it would    10       0   3  0.095973841
## 2282              director of the    32       0   3  0.118853580
## 2283                 in the house    13       0   3  0.095987251
## 2284                      than in    10      10   2  0.050470709
## 2285                       said on   10      10   2  0.045359419
## 2286                      to some    10      10   2  0.039676240
## 2287           for illegal immigrants 15      0   3  0.064707667
## 2288              united states from  10       0   3  0.154979616
## 2289                  at the time    19       0   3  0.031608226
## 2290               global warming is  10       0   3  0.126930496
## 2291                        to gun   15      15   2  0.017450521
## 2292                  the wake of    14       0   3  0.103347875
## 2293                 of the house    10       0   3  0.028949585
## 2294                 be allowed to   16       0   3  0.026235939
## 2295                     to their   22      22   2  0.007785494
## 2296                  over the past  12       0   3  0.040109928
## 2297                        to one   13      13   2  0.006239213
## 2298                         to an   27      27   2  0.003737093
## 2299                   he said he    13       0   3  0.006867163
## 2300              of the american   13       0   3  0.003401133
## 2301 environmental protection agency  10       0   3 -0.028718921
## 2302        million illegal immigrants  15       0   3 -0.041538256
## 2303                       for it   10      10   2 -0.009177974
## 2304               the dangers of   14       0   3 -0.039935385
```

```
## 2305                in the u.s   17        0   3 -0.037164509
## 2306                 and that    69       67   2 -0.008889928
## 2307         of the immigration  12        0   3 -0.029211728
## 2308              author of the  10        0   3 -0.134922314
## 2309            gays and lesbians 10        0   3 -0.187800611
## 2310              the new jersey  10        0   3 -0.128446227
## 2311                   than the  25       25   2 -0.020925081
## 2312              in new jersey   13        0   3 -0.176465675
## 2313                   part of a  25        0   3 -0.065874839
## 2314            death penalty in  11        0   3 -0.195240366
## 2315                there is no   28        0   3 -0.065401447
## 2316                      are a   23       23   2 -0.032227327
## 2317                would be a    12        0   3 -0.062215472
## 2318                      be in   22       22   2 -0.037120074
## 2319           tens of thousands  12        0   3 -0.399813541
## 2320                 going to be  13        0   3 -0.289233929
## 2321                 vote on the  13        0   3 -0.098985141
## 2322        of capital punishment 10        0   3 -0.319319128
## 2323                   people of  11       11   2 -0.064685910
## 2324                 to live in   10        0   3 -0.138342928
## 2325                the first of  13        0   3 -0.219305623
## 2326                not going to  10        0   3 -0.233504528
## 2327               the same time  12        0   3 -0.411162391
## 2328                   to make a  13        0   3 -0.151676008
## 2329                  the school  10       10   2 -0.095530642
## 2330               on both sides  14        0   3 -0.309889416
## 2331                    in state  11       11   2 -0.093089486
## 2332                  all of the  10        0   3 -0.136956187
## 2333                      to new  19       19   2 -0.073817955
## 2334                  in the new  12        0   3 -0.138306117
## 2335   national rifle association 36        0   3 -0.446169166
## 2336        federal appeals court  11        0   3 -0.264298830
## 2337                    death of  13       13   2 -0.097946939
## 2338        illegal immigrants in  10        0   3 -0.255112714
## 2339                    only the  11       11   2 -0.114208949
## 2340                     for at   10       10   2 -0.117371713
## 2341                    of state  14       14   2 -0.103660350
## 2342         state supreme court  21        0   3 -0.591355754
## 2343                    to court  14       14   2 -0.108477124
## 2344               new york city  17        0   3 -0.894116467
## 2345             earlier this year 12        0   3 -0.649544099
## 2346               according to a  31        0   3 -0.984324961
## 2347              said they would  14        0   3 -0.210349922
## 2348                 in new york   41        0   3 -0.836707680
## 2349                      so the  11       11   2 -0.153559041
## 2350                      or in   11       11   2 -0.152888816
## 2351                     was in   23       23   2 -0.110259828
## 2352          by lethal injection  10        0   3 -1.147596593
## 2353                 has been a   12        0   3 -0.251732625
## 2354                     to his   25       25   2 -0.115061468
## 2355                said he would  23        0   3 -0.218045393
## 2356               in immigration  10       10   2 -0.182367702
## 2357                of gun control 10        0   3 -0.378835380
## 2358                      but a   19       19   2 -0.146260029
```

50

```
## 2359            we need to   12         0   3 -0.342638897
## 2360                  are in   18        18   2 -0.155560909
## 2361            are trying to  12         0   3 -1.044776518
## 2362         that would have   17         0   3 -0.244963628
## 2363             at the same   16         0   3 -1.143053773
## 2364        supreme court has  11         0   3 -1.154546791
## 2365          to protect the   10         0   3 -0.655288725
## 2366           according to the 31        0   3 -1.573730675
## 2367         death row inmates  13         0   3 -1.696835505
## 2368     florida supreme court  14         0   3 -1.377202575
## 2369              the change   10        10   2 -0.271274224
## 2370             there is a   12         0   3 -0.366631406
## 2371                have the   54        54   2 -0.120834530
## 2372            for the state   11         0   3 -0.455623718
## 2373           of the national  26         0   3 -0.407814720
## 2374                   or a   11        11   2 -0.274556461
## 2375                  is on   13        13   2 -0.253872711
## 2376          death penalty is  16         0   3 -0.922234786
## 2377               to do it   12         0   3 -0.424219114
## 2378          in this country  19         0   3 -0.398517174
## 2379             of the death  12         0   3 -0.472259587
## 2380           are expected to  11         0   3 -0.728267944
## 2381      the nation's largest  12         0   3 -1.538319061
## 2382       the tobacco companies 14        0   3 -0.493003781
## 2383                and who   12        12   2 -0.287660901
## 2384            that they are   12         0   3 -0.396624298
## 2385               which the   20        20   2 -0.228491149
## 2386              the smoking   18        18   2 -0.248544268
## 2387              in the same   12         0   3 -1.569316140
## 2388              the new york  23         0   3 -1.814140438
## 2389             to immigrants  10        10   2 -0.345283212
## 2390       of same-sex marriage 11         0   3 -0.597600170
## 2391            new york times  10         0   3 -2.448559780
## 2392                 and an   15        15   2 -0.287747424
## 2393                and was   22        22   2 -0.245031763
## 2394              to get the   15         0   3 -0.472168600
## 2395                about to   13        13   2 -0.320476944
## 2396               that it is   13         0   3 -0.412859057
## 2397            i don't want   13         0   3 -1.104712050
## 2398              of new york   10         0   3 -2.014204934
## 2399          of carbon dioxide 10         0   3 -2.062997827
## 2400               to people   10        10   2 -0.377206874
## 2401                we want to  10         0   3 -0.761333396
## 2402             a number of   14         0   3 -0.951805346
## 2403                 the one   20        20   2 -0.294329493
## 2404      superior court judge   14         0   3 -2.720051979
## 2405                and has   13        13   2 -0.367696200
## 2406          at the university  15         0   3 -0.664982898
## 2407                 be the   50        50   2 -0.194580472
## 2408              the case of   10         0   3 -0.689694315
## 2409         the department of   27         0   3 -0.538823463
## 2410     illegal immigrants from 10        0   3 -0.961471488
## 2411                to get a   12         0   3 -0.623041823
## 2412                of death   10        10   2 -0.436513224
```

```
## 2413             that he is   12      0   3 -0.540747292
## 2414        was sentenced to   14      0   3 -1.038829352
## 2415             the penalty   11     11   2 -0.455801602
## 2416       santa clara county   13      0   3 -3.625091216
## 2417           i don't think   21      0   3 -1.250067454
## 2418                   it to   27     27   2 -0.302330075
## 2419           those who are   10      0   3 -0.945716790
## 2420                to about   10     10   2 -0.525114763
## 2421            to new york   11      0   3 -2.885916870
## 2422              be able to   15      0   3 -2.849865605
## 2423     the second amendment   28      0   3 -0.869862342
## 2424      across the country   18      0   3 -1.596971824
## 2425            in the world   10      0   3 -1.182170598
## 2426        death penalty for   10      0   3 -1.720448960
## 2427                 were to   11     11   2 -0.544067860
## 2428                 to be a   22      0   3 -0.521448928
## 2429                  that a   53     53   2 -0.265183265
## 2430           the state of   21      0   3 -0.748069796
## 2431                 not the   36     36   2 -0.325791350
## 2432              to be the   15      0   3 -0.609380655
## 2433                  or the   23     23   2 -0.410628578
## 2434         of the nation's   21      0   3 -2.104639071
## 2435            don't want to   13      0   3 -1.125795458
## 2436                   it a   16     16   2 -0.493976629
## 2437  carbon dioxide emissions   13      0   3 -4.374450199
## 2438               the states   18     18   2 -0.479624675
## 2439      of illegal immigrants   17      0   3 -0.895843363
## 2440             part of the   35      0   3 -1.091418508
## 2441            to make the   11      0   3 -0.984564435
## 2442        the united nations   12      0   3 -1.707254479
## 2443           that he would   11      0   3 -0.932256870
## 2444                the year   11     11   2 -0.706154794
## 2445                  he and   12     12   2 -0.670748169
## 2446                 had the   14     14   2 -0.652659075
## 2447                   is in   25     25   2 -0.500449881
## 2448               of the u.s   11      0   3 -1.174907692
## 2449           the number of   39      0   3 -1.442920292
## 2450             of the state   13      0   3 -1.113127147
## 2451           of the united   16      0   3 -1.801311033
## 2452                  said a   13     13   2 -0.712124439
## 2453        the state supreme   13      0   3 -1.299553358
## 2454                   is to   43     43   2 -0.407728635
## 2455             he did not   16      0   3 -1.390782633
## 2456                   it in   10     10   2 -0.830851520
## 2457         united states and   11      0   3 -2.074916157
## 2458                and said   14     14   2 -0.722623478
## 2459                it is a   15      0   3 -0.871938435
## 2460    united states illegally   12      0   3 -3.392639637
## 2461                 and is   27     26   2 -0.543598452
## 2462                  of it   11     11   2 -0.886894448
## 2463                 has to   11     11   2 -0.910821468
## 2464                the more   19     19   2 -0.713176389
## 2465        the supreme court   47      0   3 -1.739860917
## 2466                 and as   10     10   2 -0.982312331
```

```
## 2467                   were the   14   14   2 -0.847624556
## 2468              who have been   10    0   3 -1.586813422
## 2469                     and by   11   11   2 -0.992090954
## 2470               that he was   13    0   3 -1.225732617
## 2471               the issue of   11    0   3 -1.535899585
## 2472            of the state's   13    0   3 -2.522894142
## 2473              in the state   40    0   3 -1.006215767
## 2474                    are the   29   29   2 -0.670804818
## 2475                     are to   11   11   2 -1.091496559
## 2476              in the first   14    0   3 -1.839014880
## 2477                   it is not   16    0   3 -1.381029968
## 2478    carry concealed weapons   11    0   3 -3.825837242
## 2479                     was to   11   11   2 -1.289387882
## 2480                      as to   12   12   2 -1.280746490
## 2481                    has the   13   13   2 -1.288691452
## 2482                     and on   10   10   2 -1.526289459
## 2483                    and for   13   13   2 -1.491888776
## 2484                    of that   32   32   2 -0.971105925
## 2485               it would be   12    0   3 -2.054335183
## 2486                    in that   16   16   2 -1.408459040
## 2487                    that in   12   12   2 -1.619759083
## 2488                      on to   15   15   2 -1.465162123
## 2489               in the case   15    0   3 -2.109053175
## 2490            in the country   43    0   3 -1.917698103
## 2491                    and in   34   34   2 -1.217457838
## 2492                    to that   16   16   2 -1.774653865
## 2493                     in and   10   10   2 -2.330764290
## 2494                    and to   35   35   2 -1.642665767
##               z
## 1    59.736417112
## 2    55.318686322
## 3    54.604630727
## 4    50.061683071
## 5    49.648238782
## 6    46.578827500
## 7    46.309106994
## 8    45.878577613
## 9    42.938779734
## 10   42.525035324
## 11   40.673234322
## 12   40.597231746
## 13   40.284006073
## 14   40.219790328
## 15   39.488085271
## 16   39.291605375
## 17   39.257373904
## 18   38.161578308
## 19   37.651006760
## 20   36.720043507
## 21   35.650145793
## 22   35.269768481
## 23   35.185938502
## 24   34.628193728
## 25   34.530426785
```

```
## 26    34.422290733
## 27    33.871422585
## 28    33.789066165
## 29    33.752889057
## 30    33.740577779
## 31    33.158252129
## 32    33.021808462
## 33    32.748455253
## 34    32.730320250
## 35    32.492274906
## 36    32.379587486
## 37    32.295721696
## 38    32.031615334
## 39    31.892444009
## 40    31.748988761
## 41    31.401161848
## 42    31.274639474
## 43    30.289140960
## 44    29.747593780
## 45    29.496655422
## 46    29.368613740
## 47    29.274759950
## 48    29.172333278
## 49    28.917540759
## 50    28.727403109
## 51    28.691109806
## 52    28.654842822
## 53    28.473380975
## 54    28.330863491
## 55    28.228351492
## 56    28.182379155
## 57    28.017779227
## 58    27.840687273
## 59    27.662621353
## 60    27.632948989
## 61    27.593007083
## 62    27.536791508
## 63    27.479170547
## 64    27.409094136
## 65    27.188397812
## 66    27.112049118
## 67    27.022328253
## 68    26.986017193
## 69    26.970861642
## 70    26.917225834
## 71    26.729528682
## 72    26.602812112
## 73    26.523720872
## 74    26.521784551
## 75    26.416170427
## 76    26.413284436
## 77    26.385174185
## 78    26.146944731
## 79    26.088858723
```

```
## 80    26.087904181
## 81    26.057661972
## 82    26.057307543
## 83    25.673831340
## 84    25.441047925
## 85    25.424283671
## 86    25.424208845
## 87    25.414825505
## 88    25.264308265
## 89    25.186897205
## 90    25.129287360
## 91    25.098158326
## 92    24.977609447
## 93    24.793403503
## 94    24.684758130
## 95    24.667871941
## 96    24.660262073
## 97    24.652623723
## 98    24.368026153
## 99    24.284189848
## 100   24.210570569
## 101   24.190522069
## 102   24.177457198
## 103   24.136841500
## 104   24.102718485
## 105   24.054026619
## 106   23.996621485
## 107   23.773239053
## 108   23.761205461
## 109   23.630290949
## 110   23.628762552
## 111   23.621352347
## 112   23.594070110
## 113   23.460644150
## 114   23.459836724
## 115   23.436615751
## 116   23.397178865
## 117   23.305943705
## 118   23.290949344
## 119   23.151388651
## 120   23.124054603
## 121   23.108297344
## 122   22.994654587
## 123   22.954574997
## 124   22.883936387
## 125   22.855134323
## 126   22.777210621
## 127   22.745752147
## 128   22.650317622
## 129   22.629611310
## 130   22.536538657
## 131   22.466944650
## 132   22.436495763
## 133   22.431278235
```

```
## 134   22.429911267
## 135   22.384389885
## 136   22.335497308
## 137   22.284760083
## 138   22.278307776
## 139   22.115853456
## 140   22.109964936
## 141   22.109843801
## 142   22.102538028
## 143   22.073880246
## 144   22.043152992
## 145   22.032124717
## 146   22.031823433
## 147   21.970543339
## 148   21.919848576
## 149   21.916658859
## 150   21.836001561
## 151   21.792099623
## 152   21.775849433
## 153   21.752591750
## 154   21.727440757
## 155   21.725850247
## 156   21.639191531
## 157   21.638968424
## 158   21.590175014
## 159   21.553679051
## 160   21.532342076
## 161   21.411404551
## 162   21.394009981
## 163   21.297762888
## 164   21.268174760
## 165   21.145822847
## 166   21.042582738
## 167   21.034793088
## 168   21.002410796
## 169   20.977743576
## 170   20.969632373
## 171   20.906026440
## 172   20.780146658
## 173   20.777622710
## 174   20.771849636
## 175   20.750027259
## 176   20.637750050
## 177   20.631951950
## 178   20.584737975
## 179   20.583190064
## 180   20.564557171
## 181   20.486724153
## 182   20.446848536
## 183   20.376624421
## 184   20.334336506
## 185   20.273256454
## 186   20.272551033
## 187   20.230526452
```

```
## 188   20.171938991
## 189   20.165567157
## 190   20.161547477
## 191   20.079556271
## 192   20.073247987
## 193   20.045102687
## 194   19.983592276
## 195   19.922501668
## 196   19.889538187
## 197   19.871943599
## 198   19.849038342
## 199   19.827684859
## 200   19.785839262
## 201   19.780723260
## 202   19.756394675
## 203   19.721568110
## 204   19.712652218
## 205   19.660147959
## 206   19.611415972
## 207   19.605406519
## 208   19.562253536
## 209   19.510772927
## 210   19.476835173
## 211   19.464741163
## 212   19.409978738
## 213   19.397622957
## 214   19.338220098
## 215   19.335338771
## 216   19.332432190
## 217   19.323809197
## 218   19.307485439
## 219   19.296324898
## 220   19.255201493
## 221   19.250337546
## 222   19.210284783
## 223   19.194422059
## 224   19.164913353
## 225   19.117743315
## 226   19.068687960
## 227   19.005653131
## 228   18.965089745
## 229   18.957041303
## 230   18.953429487
## 231   18.869215318
## 232   18.826144236
## 233   18.779570219
## 234   18.776191351
## 235   18.762867258
## 236   18.711776901
## 237   18.696205403
## 238   18.677624099
## 239   18.673534052
## 240   18.620489803
## 241   18.578585405
```

```
## 242   18.570570850
## 243   18.551888190
## 244   18.532564231
## 245   18.531116921
## 246   18.512293613
## 247   18.510322043
## 248   18.492105449
## 249   18.477000069
## 250   18.458539192
## 251   18.370498582
## 252   18.343536188
## 253   18.338625329
## 254   18.261369493
## 255   18.233233684
## 256   18.226728488
## 257   18.209410670
## 258   18.207541197
## 259   18.186320648
## 260   18.178146238
## 261   18.161648216
## 262   18.115424671
## 263   18.099224574
## 264   18.094652511
## 265   18.090559237
## 266   18.087217571
## 267   18.077135810
## 268   18.073126313
## 269   18.052874501
## 270   18.050392733
## 271   18.048983021
## 272   18.026341683
## 273   18.020143645
## 274   18.018068169
## 275   17.999836676
## 276   17.995847307
## 277   17.987142073
## 278   17.943760373
## 279   17.942247483
## 280   17.916430654
## 281   17.871562865
## 282   17.855195564
## 283   17.851684222
## 284   17.792728413
## 285   17.786571458
## 286   17.769131497
## 287   17.768115959
## 288   17.718374829
## 289   17.706998833
## 290   17.693113868
## 291   17.692420420
## 292   17.681683111
## 293   17.658008977
## 294   17.657005368
## 295   17.654114741
```

```
## 296   17.626187821
## 297   17.616648288
## 298   17.562045992
## 299   17.540537584
## 300   17.514495141
## 301   17.490252948
## 302   17.464766731
## 303   17.448966631
## 304   17.448093136
## 305   17.383856969
## 306   17.330087100
## 307   17.291942379
## 308   17.291553359
## 309   17.272305579
## 310   17.271111615
## 311   17.253692213
## 312   17.247889956
## 313   17.237217648
## 314   17.168450996
## 315   17.136520355
## 316   17.093310597
## 317   17.060031125
## 318   17.046668463
## 319   17.044840573
## 320   17.024719414
## 321   17.002665767
## 322   16.960629056
## 323   16.959180033
## 324   16.959016092
## 325   16.875140464
## 326   16.847352875
## 327   16.825009041
## 328   16.821958421
## 329   16.809311189
## 330   16.801889557
## 331   16.795688154
## 332   16.775964613
## 333   16.773298066
## 334   16.761969523
## 335   16.720718445
## 336   16.705725033
## 337   16.683493443
## 338   16.675548944
## 339   16.638511274
## 340   16.601534638
## 341   16.593499586
## 342   16.583447891
## 343   16.541486459
## 344   16.539212904
## 345   16.538389401
## 346   16.531108122
## 347   16.525194571
## 348   16.500228503
## 349   16.484553738
```

```
## 350    16.469763683
## 351    16.464312001
## 352    16.430227753
## 353    16.378954189
## 354    16.377519052
## 355    16.350518575
## 356    16.347791912
## 357    16.331018210
## 358    16.324146747
## 359    16.316043250
## 360    16.283556191
## 361    16.281889734
## 362    16.268394861
## 363    16.251766307
## 364    16.238216126
## 365    16.237105412
## 366    16.222115556
## 367    16.213290630
## 368    16.209607367
## 369    16.201111003
## 370    16.188770502
## 371    16.179520621
## 372    16.166811887
## 373    16.157985714
## 374    16.136468371
## 375    16.117000281
## 376    16.112732667
## 377    16.093886631
## 378    16.053350405
## 379    16.026031889
## 380    16.011640044
## 381    15.947647181
## 382    15.936868783
## 383    15.928688891
## 384    15.859356635
## 385    15.854834378
## 386    15.846842263
## 387    15.822499689
## 388    15.801617580
## 389    15.798771186
## 390    15.795691594
## 391    15.783619082
## 392    15.775997374
## 393    15.771046493
## 394    15.742486685
## 395    15.721954317
## 396    15.694813825
## 397    15.681746766
## 398    15.669285145
## 399    15.662526633
## 400    15.658671714
## 401    15.647165459
## 402    15.632653847
## 403    15.629378459
```

```
## 404   15.611363000
## 405   15.566507927
## 406   15.548939302
## 407   15.522327068
## 408   15.516939193
## 409   15.478819714
## 410   15.472124808
## 411   15.460816137
## 412   15.441751780
## 413   15.419172367
## 414   15.414655614
## 415   15.401914674
## 416   15.390728022
## 417   15.387518916
## 418   15.367689919
## 419   15.364412691
## 420   15.346626642
## 421   15.255320655
## 422   15.251128016
## 423   15.234737719
## 424   15.214591381
## 425   15.165470622
## 426   15.144299325
## 427   15.134319412
## 428   15.110842083
## 429   15.101123801
## 430   15.057955892
## 431   15.050184713
## 432   14.998604135
## 433   14.994338256
## 434   14.979068440
## 435   14.917694169
## 436   14.910215048
## 437   14.881034917
## 438   14.880364396
## 439   14.880062481
## 440   14.872951960
## 441   14.862717493
## 442   14.861180821
## 443   14.853919794
## 444   14.836464976
## 445   14.833845692
## 446   14.833646347
## 447   14.824217974
## 448   14.824102552
## 449   14.809676974
## 450   14.796735747
## 451   14.772626378
## 452   14.770118162
## 453   14.746142026
## 454   14.742510588
## 455   14.716017585
## 456   14.704215073
## 457   14.701492641
```

```
## 458  14.693240401
## 459  14.675082235
## 460  14.664420186
## 461  14.650842052
## 462  14.635571621
## 463  14.598438210
## 464  14.579504798
## 465  14.573728203
## 466  14.550276752
## 467  14.548135166
## 468  14.540189481
## 469  14.522530444
## 470  14.513159409
## 471  14.509118765
## 472  14.503306658
## 473  14.487443951
## 474  14.484505835
## 475  14.482268632
## 476  14.474773667
## 477  14.462541491
## 478  14.451097175
## 479  14.431519767
## 480  14.426162889
## 481  14.307949569
## 482  14.291596690
## 483  14.291367379
## 484  14.277505116
## 485  14.269833975
## 486  14.239093738
## 487  14.228054438
## 488  14.221770814
## 489  14.214626167
## 490  14.205149147
## 491  14.183077094
## 492  14.180792930
## 493  14.154148097
## 494  14.141601750
## 495  14.115171473
## 496  14.107504892
## 497  14.106120532
## 498  14.083013276
## 499  14.075177039
## 500  14.060881942
## 501  14.054011753
## 502  14.028292977
## 503  14.025433474
## 504  14.013228510
## 505  13.998337934
## 506  13.988558174
## 507  13.975790877
## 508  13.967834394
## 509  13.965943288
## 510  13.951111851
## 511  13.931426574
```

```
## 512   13.924105143
## 513   13.908348157
## 514   13.899247103
## 515   13.896646745
## 516   13.888004240
## 517   13.854432998
## 518   13.839451112
## 519   13.836666976
## 520   13.836558971
## 521   13.831190963
## 522   13.831125494
## 523   13.806651864
## 524   13.797356580
## 525   13.784020485
## 526   13.760455851
## 527   13.737583384
## 528   13.725100531
## 529   13.719158359
## 530   13.709588816
## 531   13.704225092
## 532   13.658260462
## 533   13.657286585
## 534   13.654191049
## 535   13.652361093
## 536   13.640963284
## 537   13.635338179
## 538   13.619994482
## 539   13.600381304
## 540   13.590855347
## 541   13.584795563
## 542   13.578458409
## 543   13.570325838
## 544   13.569442242
## 545   13.544977479
## 546   13.537026642
## 547   13.518994944
## 548   13.511521180
## 549   13.503734447
## 550   13.498543746
## 551   13.495305323
## 552   13.486527540
## 553   13.480160248
## 554   13.477488189
## 555   13.469703868
## 556   13.463371513
## 557   13.433900126
## 558   13.407378596
## 559   13.396968250
## 560   13.387548031
## 561   13.365234054
## 562   13.362677140
## 563   13.311624500
## 564   13.304444099
## 565   13.295761647
```

```
## 566   13.291803206
## 567   13.288869418
## 568   13.286175253
## 569   13.284115943
## 570   13.275889862
## 571   13.270113393
## 572   13.252360833
## 573   13.238780664
## 574   13.236300626
## 575   13.235905782
## 576   13.230591070
## 577   13.223002476
## 578   13.205668455
## 579   13.200273940
## 580   13.175419696
## 581   13.154594820
## 582   13.149361700
## 583   13.133092579
## 584   13.132098602
## 585   13.120261687
## 586   13.105661054
## 587   13.089709869
## 588   13.081126951
## 589   13.076037035
## 590   13.073915489
## 591   13.070358316
## 592   13.057369669
## 593   13.038375736
## 594   13.037358571
## 595   12.983522620
## 596   12.976487837
## 597   12.973140518
## 598   12.972627687
## 599   12.952454298
## 600   12.932864037
## 601   12.932164989
## 602   12.908228196
## 603   12.867581770
## 604   12.864823553
## 605   12.857536993
## 606   12.855177844
## 607   12.834211228
## 608   12.825169717
## 609   12.818205472
## 610   12.817208752
## 611   12.801282038
## 612   12.782659393
## 613   12.749116337
## 614   12.743968070
## 615   12.721435999
## 616   12.719126377
## 617   12.702480738
## 618   12.696664703
## 619   12.684129512
```

```
## 620   12.673194687
## 621   12.646980996
## 622   12.632239070
## 623   12.622258447
## 624   12.620716828
## 625   12.618064201
## 626   12.615163278
## 627   12.611446517
## 628   12.607871791
## 629   12.592598347
## 630   12.555504219
## 631   12.550684898
## 632   12.544745543
## 633   12.540702424
## 634   12.538801441
## 635   12.504384119
## 636   12.487398478
## 637   12.473910986
## 638   12.468860604
## 639   12.442500850
## 640   12.433183420
## 641   12.428982007
## 642   12.396375168
## 643   12.392929899
## 644   12.381134478
## 645   12.352475733
## 646   12.337178414
## 647   12.325441067
## 648   12.322752356
## 649   12.287122957
## 650   12.264850345
## 651   12.264818071
## 652   12.253813554
## 653   12.213450604
## 654   12.203225493
## 655   12.186246137
## 656   12.183484398
## 657   12.164837775
## 658   12.154191345
## 659   12.144294666
## 660   12.137639259
## 661   12.131133614
## 662   12.127189578
## 663   12.125497198
## 664   12.097683331
## 665   12.084157934
## 666   12.073628117
## 667   12.062152437
## 668   12.046434825
## 669   12.038042205
## 670   12.007730000
## 671   11.990558086
## 672   11.983362967
## 673   11.974707492
```

```
## 674   11.966554954
## 675   11.962035957
## 676   11.929044901
## 677   11.917550141
## 678   11.909278665
## 679   11.896008951
## 680   11.892054050
## 681   11.888669727
## 682   11.879863475
## 683   11.871482600
## 684   11.863676630
## 685   11.863643864
## 686   11.861492376
## 687   11.836456495
## 688   11.822186699
## 689   11.808484705
## 690   11.808091786
## 691   11.791010593
## 692   11.781710344
## 693   11.766812147
## 694   11.763538038
## 695   11.761908896
## 696   11.760365562
## 697   11.759993265
## 698   11.748097266
## 699   11.742854867
## 700   11.721730955
## 701   11.719333834
## 702   11.715715661
## 703   11.709859439
## 704   11.707781653
## 705   11.693996993
## 706   11.693593529
## 707   11.688518229
## 708   11.676940509
## 709   11.667035222
## 710   11.655219855
## 711   11.637520596
## 712   11.636221718
## 713   11.619616416
## 714   11.619071555
## 715   11.613204957
## 716   11.609726354
## 717   11.591677518
## 718   11.583192388
## 719   11.564888927
## 720   11.564026097
## 721   11.556865394
## 722   11.529927116
## 723   11.526934148
## 724   11.525669122
## 725   11.522282841
## 726   11.521480206
## 727   11.517993404
```

```
## 728   11.484683620
## 729   11.470001876
## 730   11.449847658
## 731   11.442225319
## 732   11.435193114
## 733   11.433880344
## 734   11.420995960
## 735   11.419565515
## 736   11.415414226
## 737   11.411721829
## 738   11.405809356
## 739   11.397529454
## 740   11.372129648
## 741   11.367371023
## 742   11.364960172
## 743   11.364241916
## 744   11.356715769
## 745   11.346682381
## 746   11.338380450
## 747   11.333315751
## 748   11.330026238
## 749   11.328545292
## 750   11.319499389
## 751   11.318343915
## 752   11.312059818
## 753   11.305786925
## 754   11.303876329
## 755   11.303257755
## 756   11.290779700
## 757   11.284806341
## 758   11.271165587
## 759   11.262534027
## 760   11.260359457
## 761   11.258237185
## 762   11.202567976
## 763   11.195430285
## 764   11.182719317
## 765   11.181351079
## 766   11.172221101
## 767   11.169897166
## 768   11.152287487
## 769   11.144745975
## 770   11.141669037
## 771   11.132089833
## 772   11.129245440
## 773   11.124563419
## 774   11.123444484
## 775   11.115039337
## 776   11.113398273
## 777   11.096467083
## 778   11.088407031
## 779   11.078503053
## 780   11.066578524
## 781   11.052306829
```

```
## 782   11.033533787
## 783   11.017298456
## 784   11.013076095
## 785   11.006662141
## 786   11.001804468
## 787   10.999265136
## 788   10.994122322
## 789   10.993563911
## 790   10.978196545
## 791   10.969129739
## 792   10.956153792
## 793   10.955213895
## 794   10.952754440
## 795   10.943980605
## 796   10.939155365
## 797   10.939122430
## 798   10.937355228
## 799   10.934272250
## 800   10.909227823
## 801   10.901891921
## 802   10.901596072
## 803   10.863990163
## 804   10.863990163
## 805   10.850271088
## 806   10.843166587
## 807   10.836302273
## 808   10.820662602
## 809   10.815658745
## 810   10.792928056
## 811   10.790431358
## 812   10.777974688
## 813   10.765379037
## 814   10.758671420
## 815   10.753821716
## 816   10.746506798
## 817   10.739014000
## 818   10.735511496
## 819   10.734715137
## 820   10.729169437
## 821   10.727553514
## 822   10.722774696
## 823   10.705201707
## 824   10.703137385
## 825   10.688267499
## 826   10.675411428
## 827   10.657289288
## 828   10.654953700
## 829   10.649903376
## 830   10.647285900
## 831   10.643792689
## 832   10.625796741
## 833   10.619987772
## 834   10.616569444
## 835   10.614063073
```

```
## 836   10.611387983
## 837   10.611049548
## 838   10.610452549
## 839   10.607461267
## 840   10.596467433
## 841   10.584438448
## 842   10.584139628
## 843   10.577118886
## 844   10.570247867
## 845   10.552389522
## 846   10.542730713
## 847   10.533566504
## 848   10.531811522
## 849   10.512311192
## 850   10.509133071
## 851   10.501979717
## 852   10.497878666
## 853   10.491593938
## 854   10.486098008
## 855   10.476774605
## 856   10.460622806
## 857   10.457080374
## 858   10.456616516
## 859   10.452546979
## 860   10.449791938
## 861   10.436149734
## 862   10.425831605
## 863   10.400834819
## 864   10.392488851
## 865   10.387155262
## 866   10.379415525
## 867   10.376531370
## 868   10.372573326
## 869   10.370402356
## 870   10.366634392
## 871   10.355214256
## 872   10.333321186
## 873   10.326575360
## 874   10.320109654
## 875   10.318190709
## 876   10.310019363
## 877   10.308745409
## 878   10.301512752
## 879   10.288469526
## 880   10.284148653
## 881   10.263236976
## 882   10.260061355
## 883   10.236664765
## 884   10.230286702
## 885   10.218827360
## 886   10.210516033
## 887   10.206955101
## 888   10.191434826
## 889   10.182798484
```

```
## 890    10.180503146
## 891    10.171597203
## 892    10.155817600
## 893    10.153464097
## 894    10.149837139
## 895    10.148178709
## 896    10.134536321
## 897    10.133065256
## 898    10.118536101
## 899    10.116737929
## 900    10.112443369
## 901    10.109456711
## 902    10.100760609
## 903    10.098112063
## 904    10.094764976
## 905    10.090370006
## 906    10.077819137
## 907    10.075760797
## 908    10.073029303
## 909    10.071469464
## 910    10.064943368
## 911    10.060302897
## 912    10.060083485
## 913    10.050655435
## 914    10.030427721
## 915    10.030221753
## 916    10.019856292
## 917    10.014811242
## 918    10.000611587
## 919     9.977142048
## 920     9.970059589
## 921     9.964424364
## 922     9.961365250
## 923     9.957957668
## 924     9.948296211
## 925     9.947883744
## 926     9.946496836
## 927     9.936174026
## 928     9.935348937
## 929     9.927603157
## 930     9.927595792
## 931     9.920654417
## 932     9.901512903
## 933     9.885360220
## 934     9.879749774
## 935     9.876646662
## 936     9.867530885
## 937     9.854070160
## 938     9.826388630
## 939     9.817312121
## 940     9.815110090
## 941     9.812665809
## 942     9.812665809
## 943     9.804469826
```

```
## 944    9.801278879
## 945    9.793647077
## 946    9.790969822
## 947    9.768807731
## 948    9.764033833
## 949    9.752082226
## 950    9.742063396
## 951    9.739912747
## 952    9.731401860
## 953    9.730899170
## 954    9.729572354
## 955    9.725022552
## 956    9.713747342
## 957    9.707836695
## 958    9.693289176
## 959    9.684975872
## 960    9.679799628
## 961    9.678044459
## 962    9.667689423
## 963    9.653952425
## 964    9.639436845
## 965    9.623606723
## 966    9.619480401
## 967    9.611377375
## 968    9.597933358
## 969    9.594962251
## 970    9.592935923
## 971    9.578214100
## 972    9.578214100
## 973    9.562852849
## 974    9.561248350
## 975    9.550227098
## 976    9.539081367
## 977    9.530035935
## 978    9.520079721
## 979    9.506678722
## 980    9.499289939
## 981    9.498405620
## 982    9.494025449
## 983    9.484469887
## 984    9.480953407
## 985    9.462693862
## 986    9.458569276
## 987    9.442375463
## 988    9.439279388
## 989    9.437945512
## 990    9.435815979
## 991    9.435815979
## 992    9.434570304
## 993    9.416812445
## 994    9.392046918
## 995    9.376556257
## 996    9.360460519
## 997    9.358634937
```

```
## 998    9.353018021
## 999    9.352212752
## 1000   9.346626965
## 1001   9.344861012
## 1002   9.337828630
## 1003   9.329183575
## 1004   9.324043917
## 1005   9.323628225
## 1006   9.313146448
## 1007   9.311392255
## 1008   9.298098541
## 1009   9.289584508
## 1010   9.285715450
## 1011   9.277033340
## 1012   9.276025546
## 1013   9.271531376
## 1014   9.271516622
## 1015   9.271516622
## 1016   9.262678134
## 1017   9.257306253
## 1018   9.255599613
## 1019   9.252304442
## 1020   9.247956976
## 1021   9.247158067
## 1022   9.229618821
## 1023   9.224733392
## 1024   9.224152462
## 1025   9.219245286
## 1026   9.217990807
## 1027   9.216068469
## 1028   9.201254729
## 1029   9.200634606
## 1030   9.191944766
## 1031   9.191694455
## 1032   9.185311055
## 1033   9.158017330
## 1034   9.157104453
## 1035   9.141572055
## 1036   9.122545641
## 1037   9.122545641
## 1038   9.120275710
## 1039   9.119401649
## 1040   9.111552237
## 1041   9.108064062
## 1042   9.102771011
## 1043   9.102771011
## 1044   9.102079900
## 1045   9.098299947
## 1046   9.096112793
## 1047   9.094651131
## 1048   9.084540002
## 1049   9.082841551
## 1050   9.080517870
## 1051   9.056053565
```

```
## 1052   9.052514416
## 1053   9.050822962
## 1054   9.045648467
## 1055   9.043437742
## 1056   9.033014536
## 1057   9.031976693
## 1058   9.031182955
## 1059   9.029345297
## 1060   9.026849932
## 1061   9.012271991
## 1062   8.998975687
## 1063   8.992068566
## 1064   8.987687456
## 1065   8.981388487
## 1066   8.972094086
## 1067   8.969613696
## 1068   8.969165688
## 1069   8.967373168
## 1070   8.963416303
## 1071   8.950549951
## 1072   8.943088720
## 1073   8.940663811
## 1074   8.940657544
## 1075   8.935220832
## 1076   8.934481586
## 1077   8.911620031
## 1078   8.885133699
## 1079   8.878247358
## 1080   8.876888542
## 1081   8.871095900
## 1082   8.869950060
## 1083   8.869147754
## 1084   8.866637721
## 1085   8.860756408
## 1086   8.860483859
## 1087   8.841843921
## 1088   8.841749443
## 1089   8.822356749
## 1090   8.821711917
## 1091   8.821009931
## 1092   8.816373474
## 1093   8.812334279
## 1094   8.810301944
## 1095   8.806953224
## 1096   8.796726133
## 1097   8.785402869
## 1098   8.782220024
## 1099   8.777043822
## 1100   8.775373282
## 1101   8.766239711
## 1102   8.765586618
## 1103   8.761652521
## 1104   8.758228979
## 1105   8.756152146
```

```
## 1106   8.746501076
## 1107   8.744506198
## 1108   8.742289335
## 1109   8.741797242
## 1110   8.736370885
## 1111   8.712112955
## 1112   8.702258347
## 1113   8.689330228
## 1114   8.685857560
## 1115   8.670786444
## 1116   8.668374503
## 1117   8.662770703
## 1118   8.661972000
## 1119   8.653173268
## 1120   8.649057261
## 1121   8.645940982
## 1122   8.639777619
## 1123   8.634603651
## 1124   8.619238169
## 1125   8.617918217
## 1126   8.617840322
## 1127   8.615177851
## 1128   8.613698813
## 1129   8.613346943
## 1130   8.612821593
## 1131   8.611458025
## 1132   8.610100468
## 1133   8.607232738
## 1134   8.601980205
## 1135   8.601980205
## 1136   8.596389139
## 1137   8.592885366
## 1138   8.577373967
## 1139   8.574699855
## 1140   8.573038094
## 1141   8.568509503
## 1142   8.566649216
## 1143   8.566308929
## 1144   8.560411854
## 1145   8.556629569
## 1146   8.556478971
## 1147   8.541623568
## 1148   8.528595267
## 1149   8.520200365
## 1150   8.518167439
## 1151   8.518167439
## 1152   8.511658539
## 1153   8.510183298
## 1154   8.495762542
## 1155   8.494014851
## 1156   8.481660007
## 1157   8.478815231
## 1158   8.475948790
## 1159   8.469864654
```

```
## 1160   8.462462547
## 1161   8.436039215
## 1162   8.433051902
## 1163   8.433051902
## 1164   8.432172255
## 1165   8.428910865
## 1166   8.427194122
## 1167   8.427093165
## 1168   8.426309683
## 1169   8.423308638
## 1170   8.417287894
## 1171   8.412180790
## 1172   8.410698297
## 1173   8.410410358
## 1174   8.409811411
## 1175   8.400223761
## 1176   8.382447975
## 1177   8.382357554
## 1178   8.380579557
## 1179   8.378496079
## 1180   8.374835949
## 1181   8.371365637
## 1182   8.337019236
## 1183   8.302308580
## 1184   8.301782241
## 1185   8.299087593
## 1186   8.295267601
## 1187   8.273191615
## 1188   8.260077612
## 1189   8.256467525
## 1190   8.255982225
## 1191   8.248414342
## 1192   8.242932692
## 1193   8.238288409
## 1194   8.231980884
## 1195   8.225419156
## 1196   8.216635445
## 1197   8.209583059
## 1198   8.208536470
## 1199   8.204283036
## 1200   8.203651016
## 1201   8.198575597
## 1202   8.190912288
## 1203   8.188730956
## 1204   8.180836444
## 1205   8.180836444
## 1206   8.180120663
## 1207   8.179779792
## 1208   8.172974093
## 1209   8.169144585
## 1210   8.167643116
## 1211   8.149323714
## 1212   8.143279806
## 1213   8.138486579
```

```
## 1214   8.133117450
## 1215   8.123875651
## 1216   8.118036439
## 1217   8.110949087
## 1218   8.104636104
## 1219   8.102856569
## 1220   8.091265415
## 1221   8.086159918
## 1222   8.079299226
## 1223   8.074981306
## 1224   8.073714151
## 1225   8.068158929
## 1226   8.066623804
## 1227   8.062782959
## 1228   8.057601400
## 1229   8.054778009
## 1230   8.042121912
## 1231   8.039497179
## 1232   8.038157799
## 1233   8.029872265
## 1234   8.027098609
## 1235   8.026930350
## 1236   8.014784759
## 1237   8.009300424
## 1238   8.007451552
## 1239   7.998884362
## 1240   7.986615336
## 1241   7.984689735
## 1242   7.964753190
## 1243   7.964605150
## 1244   7.956554585
## 1245   7.956554585
## 1246   7.955548293
## 1247   7.950337207
## 1248   7.949674260
## 1249   7.948006748
## 1250   7.943092156
## 1251   7.938673380
## 1252   7.930780518
## 1253   7.920975430
## 1254   7.917015859
## 1255   7.912167761
## 1256   7.904579288
## 1257   7.899691662
## 1258   7.897706473
## 1259   7.892162846
## 1260   7.884634156
## 1261   7.876782581
## 1262   7.869346307
## 1263   7.858711396
## 1264   7.838650453
## 1265   7.838290158
## 1266   7.834104785
## 1267   7.825354121
```

```
## 1268   7.819147921
## 1269   7.812915678
## 1270   7.807649467
## 1271   7.802497094
## 1272   7.800851245
## 1273   7.799654494
## 1274   7.798457158
## 1275   7.786604077
## 1276   7.780564367
## 1277   7.775486003
## 1278   7.774861967
## 1279   7.769555912
## 1280   7.768104306
## 1281   7.762403830
## 1282   7.755516578
## 1283   7.755260381
## 1284   7.753362223
## 1285   7.751433053
## 1286   7.750364560
## 1287   7.745149685
## 1288   7.738316376
## 1289   7.726187477
## 1290   7.722161899
## 1291   7.718956753
## 1292   7.717426601
## 1293   7.712640226
## 1294   7.696747171
## 1295   7.690920532
## 1296   7.690920532
## 1297   7.688330115
## 1298   7.684380289
## 1299   7.684380289
## 1300   7.681736280
## 1301   7.675690355
## 1302   7.670543383
## 1303   7.669087260
## 1304   7.658395292
## 1305   7.652357496
## 1306   7.648208704
## 1307   7.646373917
## 1308   7.646074041
## 1309   7.646074041
## 1310   7.644855332
## 1311   7.642489391
## 1312   7.618593760
## 1313   7.602513008
## 1314   7.600488681
## 1315   7.600224212
## 1316   7.596585875
## 1317   7.586847176
## 1318   7.572585277
## 1319   7.557851975
## 1320   7.555154467
## 1321   7.553434704
```

```
## 1322   7.548897985
## 1323   7.522272572
## 1324   7.521989928
## 1325   7.518566293
## 1326   7.515724999
## 1327   7.514777329
## 1328   7.514461304
## 1329   7.506998599
## 1330   7.494911219
## 1331   7.494911219
## 1332   7.492067160
## 1333   7.490981738
## 1334   7.487186683
## 1335   7.484679670
## 1336   7.470872574
## 1337   7.466224103
## 1338   7.461197906
## 1339   7.461197906
## 1340   7.440539462
## 1341   7.435355081
## 1342   7.434847375
## 1343   7.427874266
## 1344   7.415222769
## 1345   7.411400423
## 1346   7.402765867
## 1347   7.387837138
## 1348   7.386654899
## 1349   7.386518825
## 1350   7.378453403
## 1351   7.366884665
## 1352   7.366835749
## 1353   7.347360806
## 1354   7.345079439
## 1355   7.343009506
## 1356   7.339305522
## 1357   7.338881566
## 1358   7.335935261
## 1359   7.333545151
## 1360   7.332964801
## 1361   7.324500861
## 1362   7.322743476
## 1363   7.321813136
## 1364   7.316323489
## 1365   7.306069444
## 1366   7.303112632
## 1367   7.297016788
## 1368   7.293494966
## 1369   7.285395984
## 1370   7.274561010
## 1371   7.267541600
## 1372   7.255169642
## 1373   7.254566663
## 1374   7.250165156
## 1375   7.246654708
```

```
## 1376   7.246654708
## 1377   7.244287971
## 1378   7.231831520
## 1379   7.228944319
## 1380   7.203870128
## 1381   7.200879864
## 1382   7.199251048
## 1383   7.186853131
## 1384   7.181897112
## 1385   7.178734595
## 1386   7.177109484
## 1387   7.165101855
## 1388   7.161454570
## 1389   7.158989327
## 1390   7.156729689
## 1391   7.134393931
## 1392   7.126484556
## 1393   7.123886694
## 1394   7.118118626
## 1395   7.118118626
## 1396   7.113097369
## 1397   7.108110689
## 1398   7.106066167
## 1399   7.102743351
## 1400   7.102604274
## 1401   7.098493416
## 1402   7.092061301
## 1403   7.088332763
## 1404   7.083080166
## 1405   7.079769405
## 1406   7.066975930
## 1407   7.061437041
## 1408   7.056747591
## 1409   7.037584312
## 1410   7.031299144
## 1411   7.021640921
## 1412   7.021005373
## 1413   7.019410226
## 1414   7.019379371
## 1415   7.018401174
## 1416   7.018363365
## 1417   7.010247338
## 1418   7.008642961
## 1419   6.999354839
## 1420   6.993851129
## 1421   6.980979975
## 1422   6.968216495
## 1423   6.965924280
## 1424   6.960347502
## 1425   6.936325443
## 1426   6.918748031
## 1427   6.915651535
## 1428   6.914198062
## 1429   6.907434025
```

```
## 1430   6.899355523
## 1431   6.885633805
## 1432   6.874847062
## 1433   6.868743764
## 1434   6.866851137
## 1435   6.864802900
## 1436   6.862290165
## 1437   6.858526150
## 1438   6.853406574
## 1439   6.850412193
## 1440   6.836495401
## 1441   6.834076040
## 1442   6.834076040
## 1443   6.831631886
## 1444   6.820233806
## 1445   6.817300067
## 1446   6.816582073
## 1447   6.815786502
## 1448   6.815279834
## 1449   6.807137997
## 1450   6.804809722
## 1451   6.789726526
## 1452   6.786581020
## 1453   6.785746913
## 1454   6.776784118
## 1455   6.751272786
## 1456   6.751014608
## 1457   6.747567039
## 1458   6.746201134
## 1459   6.725243732
## 1460   6.717066084
## 1461   6.711159074
## 1462   6.702503457
## 1463   6.690945224
## 1464   6.688561599
## 1465   6.687165136
## 1466   6.683628828
## 1467   6.681704707
## 1468   6.679447678
## 1469   6.678222578
## 1470   6.661068843
## 1471   6.647483364
## 1472   6.645919009
## 1473   6.631369440
## 1474   6.627251053
## 1475   6.623429362
## 1476   6.620468213
## 1477   6.619919613
## 1478   6.590318092
## 1479   6.580370512
## 1480   6.578847728
## 1481   6.569072580
## 1482   6.563852276
## 1483   6.559766257
```

```
## 1484   6.558859519
## 1485   6.548484631
## 1486   6.548152052
## 1487   6.546464504
## 1488   6.545928195
## 1489   6.532333913
## 1490   6.511081289
## 1491   6.489988255
## 1492   6.489426729
## 1493   6.476636949
## 1494   6.467932902
## 1495   6.465951045
## 1496   6.463427079
## 1497   6.462443323
## 1498   6.452426461
## 1499   6.449944958
## 1500   6.443477915
## 1501   6.439625499
## 1502   6.429205815
## 1503   6.424816448
## 1504   6.409756557
## 1505   6.408825199
## 1506   6.399347331
## 1507   6.393060468
## 1508   6.386899733
## 1509   6.385274738
## 1510   6.381927761
## 1511   6.364706744
## 1512   6.352463840
## 1513   6.344602701
## 1514   6.340973353
## 1515   6.330730505
## 1516   6.325334760
## 1517   6.325334760
## 1518   6.309326840
## 1519   6.305587651
## 1520   6.290762145
## 1521   6.286384417
## 1522   6.263603608
## 1523   6.261155985
## 1524   6.257532823
## 1525   6.256123294
## 1526   6.246001461
## 1527   6.240117877
## 1528   6.236207533
## 1529   6.229385212
## 1530   6.193348307
## 1531   6.185731429
## 1532   6.183219616
## 1533   6.177745204
## 1534   6.176845647
## 1535   6.164581939
## 1536   6.160795280
## 1537   6.155695144
```

```
## 1538   6.155695144
## 1539   6.129872774
## 1540   6.129753215
## 1541   6.126501874
## 1542   6.107744680
## 1543   6.105982748
## 1544   6.105856092
## 1545   6.101011903
## 1546   6.093202882
## 1547   6.089280230
## 1548   6.073153103
## 1549   6.066009570
## 1550   6.065702753
## 1551   6.052885627
## 1552   6.047183988
## 1553   6.044031722
## 1554   6.044031722
## 1555   6.037163170
## 1556   6.005433077
## 1557   6.001431444
## 1558   5.982616665
## 1559   5.978616622
## 1560   5.973963233
## 1561   5.971865277
## 1562   5.970595837
## 1563   5.964863434
## 1564   5.956084073
## 1565   5.941897251
## 1566   5.940866978
## 1567   5.928276496
## 1568   5.924893051
## 1569   5.922887144
## 1570   5.913155382
## 1571   5.899071493
## 1572   5.891516998
## 1573   5.891427514
## 1574   5.885891962
## 1575   5.884795690
## 1576   5.870824607
## 1577   5.865060784
## 1578   5.860925429
## 1579   5.855463402
## 1580   5.851058597
## 1581   5.843491913
## 1582   5.842401500
## 1583   5.826871816
## 1584   5.820241750
## 1585   5.817443673
## 1586   5.813822101
## 1587   5.811279869
## 1588   5.799841931
## 1589   5.796457078
## 1590   5.795751328
## 1591   5.773158558
```

```
## 1592   5.732823388
## 1593   5.723379322
## 1594   5.716086941
## 1595   5.706396407
## 1596   5.681888946
## 1597   5.678785685
## 1598   5.672004898
## 1599   5.663518867
## 1600   5.655370443
## 1601   5.652887390
## 1602   5.649896574
## 1603   5.649481880
## 1604   5.640570591
## 1605   5.633651082
## 1606   5.619815634
## 1607   5.619240700
## 1608   5.601790570
## 1609   5.593296834
## 1610   5.590791212
## 1611   5.589815533
## 1612   5.589208760
## 1613   5.577520953
## 1614   5.577094362
## 1615   5.575234417
## 1616   5.574843442
## 1617   5.568307985
## 1618   5.552626863
## 1619   5.551925963
## 1620   5.535101323
## 1621   5.523646534
## 1622   5.523172752
## 1623   5.515686119
## 1624   5.504624050
## 1625   5.497663248
## 1626   5.491352788
## 1627   5.487937528
## 1628   5.487721435
## 1629   5.481815524
## 1630   5.476795258
## 1631   5.466187989
## 1632   5.457304992
## 1633   5.450602633
## 1634   5.448733136
## 1635   5.439201606
## 1636   5.437427893
## 1637   5.435844848
## 1638   5.425300793
## 1639   5.425300793
## 1640   5.422680780
## 1641   5.407157356
## 1642   5.405338302
## 1643   5.402384521
## 1644   5.402257393
## 1645   5.388980185
```

```
## 1646   5.385114803
## 1647   5.380631346
## 1648   5.363398969
## 1649   5.363236133
## 1650   5.352864628
## 1651   5.343934792
## 1652   5.339108016
## 1653   5.327538617
## 1654   5.323312947
## 1655   5.321823187
## 1656   5.307933918
## 1657   5.296193047
## 1658   5.283700691
## 1659   5.281214487
## 1660   5.279849753
## 1661   5.274244671
## 1662   5.272891190
## 1663   5.266769847
## 1664   5.255858483
## 1665   5.255611388
## 1666   5.251719733
## 1667   5.251717363
## 1668   5.233589242
## 1669   5.229014953
## 1670   5.202569979
## 1671   5.196619307
## 1672   5.189229336
## 1673   5.184206337
## 1674   5.183888104
## 1675   5.177994344
## 1676   5.163809276
## 1677   5.162878616
## 1678   5.158343378
## 1679   5.153366687
## 1680   5.148109367
## 1681   5.145577160
## 1682   5.140687975
## 1683   5.136159172
## 1684   5.135458362
## 1685   5.133380402
## 1686   5.130362709
## 1687   5.122419772
## 1688   5.117778903
## 1689   5.116669332
## 1690   5.106624646
## 1691   5.105014542
## 1692   5.102597523
## 1693   5.075945430
## 1694   5.064588392
## 1695   5.054683297
## 1696   5.052542394
## 1697   5.046862654
## 1698   5.042644726
## 1699   5.040129467
```

```
## 1700   5.036304218
## 1701   5.034599983
## 1702   5.031139454
## 1703   5.028425375
## 1704   5.026635090
## 1705   5.018573778
## 1706   5.018573778
## 1707   5.018573160
## 1708   5.012201182
## 1709   5.010360986
## 1710   4.997567666
## 1711   4.995259528
## 1712   4.984336528
## 1713   4.972316717
## 1714   4.935750774
## 1715   4.930545126
## 1716   4.927096242
## 1717   4.909563571
## 1718   4.908805850
## 1719   4.898073211
## 1720   4.895664772
## 1721   4.891661466
## 1722   4.883138951
## 1723   4.861011696
## 1724   4.855228748
## 1725   4.854177606
## 1726   4.854177606
## 1727   4.849480740
## 1728   4.845791734
## 1729   4.838379609
## 1730   4.836025589
## 1731   4.829533103
## 1732   4.818360641
## 1733   4.816923415
## 1734   4.789866136
## 1735   4.789262136
## 1736   4.777864790
## 1737   4.775661672
## 1738   4.766526554
## 1739   4.739238944
## 1740   4.717609001
## 1741   4.713014522
## 1742   4.710580790
## 1743   4.709260528
## 1744   4.708460810
## 1745   4.689414558
## 1746   4.673260735
## 1747   4.666757249
## 1748   4.648730869
## 1749   4.647362522
## 1750   4.640423043
## 1751   4.637853294
## 1752   4.630261252
## 1753   4.623001238
```

```
## 1754   4.617197822
## 1755   4.614643171
## 1756   4.608334086
## 1757   4.602799523
## 1758   4.585445313
## 1759   4.584738714
## 1760   4.582372481
## 1761   4.582001106
## 1762   4.572483904
## 1763   4.571350412
## 1764   4.567109794
## 1765   4.565848564
## 1766   4.554488111
## 1767   4.553394095
## 1768   4.551559307
## 1769   4.547224758
## 1770   4.530501954
## 1771   4.528298513
## 1772   4.523176932
## 1773   4.520350927
## 1774   4.517127981
## 1775   4.496059975
## 1776   4.486900716
## 1777   4.483901983
## 1778   4.479988038
## 1779   4.477618882
## 1780   4.475949191
## 1781   4.465952113
## 1782   4.465704610
## 1783   4.465704610
## 1784   4.465162007
## 1785   4.446952164
## 1786   4.446259997
## 1787   4.446159710
## 1788   4.445434314
## 1789   4.442625566
## 1790   4.439437175
## 1791   4.434940976
## 1792   4.433677830
## 1793   4.428329518
## 1794   4.426157347
## 1795   4.423568031
## 1796   4.421213032
## 1797   4.414161556
## 1798   4.402300461
## 1799   4.394716766
## 1800   4.374325854
## 1801   4.366849891
## 1802   4.349355919
## 1803   4.323303627
## 1804   4.311284949
## 1805   4.303672221
## 1806   4.289453831
## 1807   4.288454352
```

```
## 1808   4.285480089
## 1809   4.277083671
## 1810   4.277006668
## 1811   4.276843645
## 1812   4.267849729
## 1813   4.267257443
## 1814   4.265914325
## 1815   4.261757375
## 1816   4.236547562
## 1817   4.226360514
## 1818   4.222119218
## 1819   4.221865628
## 1820   4.196313059
## 1821   4.194163389
## 1822   4.186827144
## 1823   4.183894772
## 1824   4.165302745
## 1825   4.154936810
## 1826   4.154802697
## 1827   4.150622468
## 1828   4.142878737
## 1829   4.133828976
## 1830   4.129716879
## 1831   4.129108290
## 1832   4.121383422
## 1833   4.109778325
## 1834   4.108037486
## 1835   4.076844873
## 1836   4.069913879
## 1837   4.065423429
## 1838   4.064452294
## 1839   4.061883783
## 1840   4.061692292
## 1841   4.056631430
## 1842   4.052870126
## 1843   4.052684692
## 1844   4.043972451
## 1845   4.038178651
## 1846   4.016317996
## 1847   4.008160078
## 1848   4.003745927
## 1849   4.001506354
## 1850   4.000359757
## 1851   3.993217461
## 1852   3.980953296
## 1853   3.976186985
## 1854   3.963726843
## 1855   3.954893543
## 1856   3.952227065
## 1857   3.947100324
## 1858   3.937103952
## 1859   3.934282452
## 1860   3.930113076
## 1861   3.927742991
```

```
## 1862   3.925172922
## 1863   3.909025925
## 1864   3.909025925
## 1865   3.897949009
## 1866   3.895803078
## 1867   3.888885835
## 1868   3.876863570
## 1869   3.875961620
## 1870   3.873893358
## 1871   3.854230067
## 1872   3.844590031
## 1873   3.843478622
## 1874   3.838772790
## 1875   3.832856599
## 1876   3.831259410
## 1877   3.827823453
## 1878   3.823555315
## 1879   3.821642998
## 1880   3.807872388
## 1881   3.792487158
## 1882   3.784214069
## 1883   3.781599934
## 1884   3.767257942
## 1885   3.764376261
## 1886   3.761178790
## 1887   3.756829651
## 1888   3.746805997
## 1889   3.737817444
## 1890   3.730635626
## 1891   3.715984713
## 1892   3.710450546
## 1893   3.705952344
## 1894   3.700929919
## 1895   3.700358042
## 1896   3.699749335
## 1897   3.694820577
## 1898   3.692763948
## 1899   3.687946157
## 1900   3.681847958
## 1901   3.681021995
## 1902   3.666275956
## 1903   3.664171749
## 1904   3.662516851
## 1905   3.657512627
## 1906   3.655401048
## 1907   3.645210322
## 1908   3.637276262
## 1909   3.610723186
## 1910   3.592475130
## 1911   3.585367635
## 1912   3.553437919
## 1913   3.548455024
## 1914   3.534690211
## 1915   3.521473063
```

```
## 1916   3.520781838
## 1917   3.499515315
## 1918   3.490852476
## 1919   3.488106729
## 1920   3.484643413
## 1921   3.484037663
## 1922   3.480830438
## 1923   3.474633749
## 1924   3.467289280
## 1925   3.455409345
## 1926   3.448052727
## 1927   3.445373591
## 1928   3.427512858
## 1929   3.403068014
## 1930   3.403068014
## 1931   3.398535572
## 1932   3.397830849
## 1933   3.390733268
## 1934   3.377855605
## 1935   3.372914170
## 1936   3.369325480
## 1937   3.364942821
## 1938   3.359381725
## 1939   3.355023043
## 1940   3.334166426
## 1941   3.333220975
## 1942   3.328686224
## 1943   3.319118348
## 1944   3.309093433
## 1945   3.308713136
## 1946   3.284174089
## 1947   3.266209630
## 1948   3.252500773
## 1949   3.239289766
## 1950   3.237341034
## 1951   3.236466934
## 1952   3.229934676
## 1953   3.217782223
## 1954   3.206627599
## 1955   3.205782779
## 1956   3.202742564
## 1957   3.186972450
## 1958   3.184984953
## 1959   3.171624859
## 1960   3.149956178
## 1961   3.144256133
## 1962   3.143125757
## 1963   3.141392231
## 1964   3.133966025
## 1965   3.122068004
## 1966   3.114249010
## 1967   3.092010228
## 1968   3.065301497
## 1969   3.063846418
```

```
## 1970   3.050597842
## 1971   3.045824537
## 1972   3.044639566
## 1973   3.044088619
## 1974   3.041947657
## 1975   3.041947657
## 1976   3.039829719
## 1977   3.036820323
## 1978   3.030400456
## 1979   3.010653673
## 1980   2.997200208
## 1981   2.993906144
## 1982   2.978326462
## 1983   2.977695408
## 1984   2.958107995
## 1985   2.953546234
## 1986   2.945124477
## 1987   2.924766180
## 1988   2.919743842
## 1989   2.901289648
## 1990   2.899204170
## 1991   2.898404140
## 1992   2.894098422
## 1993   2.886746935
## 1994   2.853452414
## 1995   2.853038488
## 1996   2.841561360
## 1997   2.822283399
## 1998   2.814984471
## 1999   2.813899956
## 2000   2.807814632
## 2001   2.807529716
## 2002   2.802606395
## 2003   2.794379271
## 2004   2.784263605
## 2005   2.783272034
## 2006   2.775922516
## 2007   2.764739438
## 2008   2.761704193
## 2009   2.747409427
## 2010   2.745414438
## 2011   2.742523482
## 2012   2.731974866
## 2013   2.729161029
## 2014   2.720976653
## 2015   2.716979469
## 2016   2.686469845
## 2017   2.672113991
## 2018   2.668301407
## 2019   2.664737208
## 2020   2.662592789
## 2021   2.656859682
## 2022   2.637712968
## 2023   2.625024739
```

```
## 2024   2.621472605
## 2025   2.620838222
## 2026   2.611206413
## 2027   2.602214989
## 2028   2.593232183
## 2029   2.592599111
## 2030   2.577437752
## 2031   2.576589206
## 2032   2.575893785
## 2033   2.541273138
## 2034   2.533268592
## 2035   2.531903775
## 2036   2.502784782
## 2037   2.500403807
## 2038   2.497428887
## 2039   2.473538784
## 2040   2.468743462
## 2041   2.452114037
## 2042   2.439808725
## 2043   2.434643804
## 2044   2.433833610
## 2045   2.427973836
## 2046   2.426612926
## 2047   2.414653057
## 2048   2.412399763
## 2049   2.408821177
## 2050   2.396924848
## 2051   2.393754239
## 2052   2.392982336
## 2053   2.392135685
## 2054   2.391118339
## 2055   2.369861873
## 2056   2.361308496
## 2057   2.360996620
## 2058   2.344854835
## 2059   2.337473802
## 2060   2.324561059
## 2061   2.308191612
## 2062   2.306672429
## 2063   2.297480053
## 2064   2.294864718
## 2065   2.292584297
## 2066   2.290762510
## 2067   2.269724288
## 2068   2.255519314
## 2069   2.253796352
## 2070   2.244033135
## 2071   2.221496293
## 2072   2.217595371
## 2073   2.204227400
## 2074   2.194002640
## 2075   2.192687672
## 2076   2.186955216
## 2077   2.181446965
```

```
## 2078   2.175485691
## 2079   2.152665634
## 2080   2.149462673
## 2081   2.138273788
## 2082   2.136727862
## 2083   2.136546569
## 2084   2.124869290
## 2085   2.120527092
## 2086   2.115350916
## 2087   2.110143033
## 2088   2.105767397
## 2089   2.085735621
## 2090   2.085194413
## 2091   2.059898557
## 2092   2.054249614
## 2093   2.044935939
## 2094   2.044520330
## 2095   2.022152564
## 2096   2.021946450
## 2097   2.007579070
## 2098   2.005747359
## 2099   1.995484096
## 2100   1.959386913
## 2101   1.949109609
## 2102   1.936197824
## 2103   1.933642791
## 2104   1.932410776
## 2105   1.927284781
## 2106   1.926216511
## 2107   1.922123627
## 2108   1.909999726
## 2109   1.881500677
## 2110   1.873772306
## 2111   1.868000873
## 2112   1.856368836
## 2113   1.847577466
## 2114   1.845113972
## 2115   1.840210866
## 2116   1.836414653
## 2117   1.834582722
## 2118   1.828526525
## 2119   1.800715871
## 2120   1.798360801
## 2121   1.796854683
## 2122   1.783664447
## 2123   1.780609703
## 2124   1.772886848
## 2125   1.772858237
## 2126   1.764897640
## 2127   1.750092891
## 2128   1.746185699
## 2129   1.742220596
## 2130   1.739073251
## 2131   1.729399614
```

```
## 2132   1.726562020
## 2133   1.716332492
## 2134   1.707850894
## 2135   1.695975328
## 2136   1.682320122
## 2137   1.681621943
## 2138   1.681130000
## 2139   1.677690119
## 2140   1.657693168
## 2141   1.652449951
## 2142   1.648306287
## 2143   1.620322168
## 2144   1.616179691
## 2145   1.605426357
## 2146   1.601919591
## 2147   1.586589227
## 2148   1.574069253
## 2149   1.540463399
## 2150   1.536228400
## 2151   1.533127379
## 2152   1.530031379
## 2153   1.525997201
## 2154   1.506628372
## 2155   1.498069793
## 2156   1.492923143
## 2157   1.491643436
## 2158   1.481772259
## 2159   1.477203653
## 2160   1.470727449
## 2161   1.462721802
## 2162   1.462657256
## 2163   1.462363986
## 2164   1.435567583
## 2165   1.435011418
## 2166   1.415218865
## 2167   1.414562584
## 2168   1.414264293
## 2169   1.396223924
## 2170   1.389174741
## 2171   1.386721113
## 2172   1.386577492
## 2173   1.376094088
## 2174   1.348269344
## 2175   1.345524264
## 2176   1.339039750
## 2177   1.334734657
## 2178   1.332445152
## 2179   1.330073917
## 2180   1.319061348
## 2181   1.313858827
## 2182   1.309430613
## 2183   1.301080083
## 2184   1.293658311
## 2185   1.287603957
```

```
## 2186  1.277319387
## 2187  1.273642912
## 2188  1.260144989
## 2189  1.257382637
## 2190  1.250219123
## 2191  1.246458860
## 2192  1.242432954
## 2193  1.237760351
## 2194  1.232053633
## 2195  1.226115256
## 2196  1.204553027
## 2197  1.185354387
## 2198  1.176810513
## 2199  1.174954576
## 2200  1.173609995
## 2201  1.168946802
## 2202  1.168008866
## 2203  1.151977869
## 2204  1.139491989
## 2205  1.135816916
## 2206  1.130479288
## 2207  1.126867254
## 2208  1.122393158
## 2209  1.121317809
## 2210  1.118368663
## 2211  1.107851780
## 2212  1.083311996
## 2213  1.071585055
## 2214  1.063449651
## 2215  1.061961431
## 2216  1.050066581
## 2217  1.028397028
## 2218  1.006354060
## 2219  0.982134036
## 2220  0.980842371
## 2221  0.975694690
## 2222  0.968963948
## 2223  0.962920448
## 2224  0.960514335
## 2225  0.919450337
## 2226  0.904609466
## 2227  0.783391876
## 2228  0.757195977
## 2229  0.753459976
## 2230  0.752947502
## 2231  0.744100316
## 2232  0.741257870
## 2233  0.706395316
## 2234  0.705183562
## 2235  0.700464047
## 2236  0.693428154
## 2237  0.683115477
## 2238  0.669294154
## 2239  0.667947018
```

```
## 2240   0.664343674
## 2241   0.661182940
## 2242   0.654857785
## 2243   0.636993553
## 2244   0.634422955
## 2245   0.626098298
## 2246   0.625868772
## 2247   0.622228987
## 2248   0.607345026
## 2249   0.605158348
## 2250   0.593396244
## 2251   0.591755924
## 2252   0.589839029
## 2253   0.583925997
## 2254   0.581361592
## 2255   0.569785269
## 2256   0.547627888
## 2257   0.507122915
## 2258   0.494716816
## 2259   0.469197122
## 2260   0.467265899
## 2261   0.459600104
## 2262   0.441712235
## 2263   0.439767211
## 2264   0.437589324
## 2265   0.423433819
## 2266   0.396744238
## 2267   0.378168788
## 2268   0.375293796
## 2269   0.370607500
## 2270   0.366952498
## 2271   0.341449997
## 2272   0.335231351
## 2273   0.333962021
## 2274   0.329942579
## 2275   0.310117233
## 2276   0.296927473
## 2277   0.296390431
## 2278   0.295763387
## 2279   0.273110368
## 2280   0.233794234
## 2281   0.223469984
## 2282   0.216770647
## 2283   0.167731991
## 2284   0.161554709
## 2285   0.145905580
## 2286   0.126241541
## 2287   0.100668542
## 2288   0.097292188
## 2289   0.082267606
## 2290   0.079048461
## 2291   0.067462869
## 2292   0.060788194
## 2293   0.049138436
```

```
## 2294  0.038124868
## 2295  0.036251595
## 2296  0.027145441
## 2297  0.022517945
## 2298  0.019232166
## 2299  0.011862863
## 2300  0.006639727
## 2301 -0.015910851
## 2302 -0.027851313
## 2303 -0.029508488
## 2304 -0.039134282
## 2305 -0.062152412
## 2306 -0.072723000
## 2307 -0.074782133
## 2308 -0.079000911
## 2309 -0.079629903
## 2310 -0.083629606
## 2311 -0.102704414
## 2312 -0.115123891
## 2313 -0.115704661
## 2314 -0.131223288
## 2315 -0.141868895
## 2316 -0.154083668
## 2317 -0.161701561
## 2318 -0.173863086
## 2319 -0.185856038
## 2320 -0.197750416
## 2321 -0.203433026
## 2322 -0.211747477
## 2323 -0.216133551
## 2324 -0.235517934
## 2325 -0.251594987
## 2326 -0.255840484
## 2327 -0.272597687
## 2328 -0.299194517
## 2329 -0.299848323
## 2330 -0.308713648
## 2331 -0.311959198
## 2332 -0.319427205
## 2333 -0.320468365
## 2334 -0.322377271
## 2335 -0.343534131
## 2336 -0.347751641
## 2337 -0.354688371
## 2338 -0.375393745
## 2339 -0.377562476
## 2340 -0.377584672
## 2341 -0.388971533
## 2342 -0.393981691
## 2343 -0.406449739
## 2344 -0.410632097
## 2345 -0.435967136
## 2346 -0.489145902
## 2347 -0.495890340
```

```
## 2348 -0.505441227
## 2349 -0.508126576
## 2350 -0.513139881
## 2351 -0.528099371
## 2352 -0.537915900
## 2353 -0.564317152
## 2354 -0.571337700
## 2355 -0.575246897
## 2356 -0.584556001
## 2357 -0.598685461
## 2358 -0.638030400
## 2359 -0.653005998
## 2360 -0.661714128
## 2361 -0.675340741
## 2362 -0.704614975
## 2363 -0.760497476
## 2364 -0.768040989
## 2365 -0.770512851
## 2366 -0.782542759
## 2367 -0.807932179
## 2368 -0.819407750
## 2369 -0.855679636
## 2370 -0.866989986
## 2371 -0.868155728
## 2372 -0.906103620
## 2373 -0.906799487
## 2374 -0.921620640
## 2375 -0.926299785
## 2376 -0.932176301
## 2377 -0.939430420
## 2378 -0.953034556
## 2379 -0.956434248
## 2380 -0.968753440
## 2381 -0.972484015
## 2382 -0.994841328
## 2383 -1.006287389
## 2384 -1.008904323
## 2385 -1.010788611
## 2386 -1.039721184
## 2387 -1.040562768
## 2388 -1.097819535
## 2389 -1.104694503
## 2390 -1.117886416
## 2391 -1.119069764
## 2392 -1.120576552
## 2393 -1.148461299
## 2394 -1.156038846
## 2395 -1.162331298
## 2396 -1.163016375
## 2397 -1.168362704
## 2398 -1.200695870
## 2399 -1.206475922
## 2400 -1.207278674
## 2401 -1.236081273
```

```
## 2402 -1.263020225
## 2403 -1.297537224
## 2404 -1.304724687
## 2405 -1.337558779
## 2406 -1.341415601
## 2407 -1.348249296
## 2408 -1.351949026
## 2409 -1.370206951
## 2410 -1.386237639
## 2411 -1.397498106
## 2412 -1.399568629
## 2413 -1.445809614
## 2414 -1.508961864
## 2415 -1.511175024
## 2416 -1.527123972
## 2417 -1.560521293
## 2418 -1.563077625
## 2419 -1.646957864
## 2420 -1.683320181
## 2421 -1.700283521
## 2422 -1.706357171
## 2423 -1.737017134
## 2424 -1.753407099
## 2425 -1.792330129
## 2426 -1.795517129
## 2427 -1.825835217
## 2428 -1.900039985
## 2429 -1.911814454
## 2430 -1.920160192
## 2431 -1.925639490
## 2432 -1.935199994
## 2433 -1.951730737
## 2434 -1.976358153
## 2435 -1.980532588
## 2436 -1.988700943
## 2437 -1.993378836
## 2438 -2.017385129
## 2439 -2.019142590
## 2440 -2.190002729
## 2441 -2.228620095
## 2442 -2.318858358
## 2443 -2.326755650
## 2444 -2.352510784
## 2445 -2.355499757
## 2446 -2.447336043
## 2447 -2.503811645
## 2448 -2.511275513
## 2449 -2.530328811
## 2450 -2.557347580
## 2451 -2.562821779
## 2452 -2.597689556
## 2453 -2.630552650
## 2454 -2.651329000
## 2455 -2.670234730
```

```
## 2456 -2.676684601
## 2457 -2.696995853
## 2458 -2.731064412
## 2459 -2.788534821
## 2460 -2.790946477
## 2461 -2.822469676
## 2462 -2.986136406
## 2463 -3.065570081
## 2464 -3.093291176
## 2465 -3.120300623
## 2466 -3.164825278
## 2467 -3.186598129
## 2468 -3.301963864
## 2469 -3.344908536
## 2470 -3.379617694
## 2471 -3.513291873
## 2472 -3.531282781
## 2473 -3.545513811
## 2474 -3.586349871
## 2475 -3.677696554
## 2476 -3.708520123
## 2477 -3.949689268
## 2478 -4.227259546
## 2479 -4.348797110
## 2480 -4.503035389
## 2481 -4.695026475
## 2482 -4.927142839
## 2483 -5.458832821
## 2484 -5.490063677
## 2485 -5.554520565
## 2486 -5.694985028
## 2487 -5.706543277
## 2488 -5.739542651
## 2489 -6.055721873
## 2490 -6.284795236
## 2491 -7.100593190
## 2492 -7.179143000
## 2493 -7.535656938
## 2494 -9.729884003
```

Not quite what we want, right? Now again with our pre-processed tokens where we removed stopwords.

```
news_corp_coll_no_stop <- textstat_collocations(toks_news_subs, size = 2:3, min_count = 10)
news_corp_coll_no_stop
```

```
##                  collocation count count_nested length    lambda           z
## 1                illeg immigr   212          127      2 5.7824640 52.36049163
## 2               climat chang   129           77      2 7.5216777 49.18609439
## 3                  unit state   271          122      2 7.1768200 48.13933718
## 4               death penalti   253          131      2 8.5472499 44.16924071
## 5                 global warm   119           76      2 9.2488545 42.94497411
## 6             same-sex marriag    95           68      2 7.0723026 42.49855264
## 7                 gay marriag    87           62      2 6.4906110 42.29083411
## 8                 gun control   109           77      2 6.4649117 41.80255062
## 9                   last week    85           79      2 5.9129219 41.43818474
```

```
## 10            last year   107      88      2  4.9870936 41.40871926
## 11       tobacco compani    71      42      2  5.6377909 37.29871084
## 12       attorney general    57      47      2  6.8903487 36.89494769
## 13          capit punish    58      42      2  8.8830770 35.44383007
## 14           high school    52      42      2  6.6795450 35.08282241
## 15         same-sex coupl    48      29      2  6.7053637 34.72280305
## 16          second amend    45      27      2  7.3844664 34.29732468
## 17            white hous    53      27      2  7.0740004 34.07608164
## 18           civil union    43      27      2  7.7939803 33.74706874
## 19          greenhous gas    42      42      2  8.6096184 32.34147660
## 20       tobacco industri    52      30      2  5.6756855 32.32963794
## 21              year ago    71      59      2  6.4524827 32.19556217
## 22            last month    51      49      2  5.3796060 32.04238043
## 23           natur servic    37      27      2  6.9737735 31.43299839
## 24             smoke ban    59      40      2  4.7416357 31.38136345
## 25          death sentenc    57      22      2  4.9091486 31.26613675
## 26          presid obama    45      39      2  5.7264418 31.07780811
## 27      background check    75      59      2 10.2756762 30.56137905
## 28        conceal weapon    35      28      2  8.1595001 29.98529908
## 29            rifl associ    39      39      2  8.6572153 29.74032418
## 30             gas emiss    33      32      2  7.2068423 29.40623130
## 31        presid clinton    39      33      2  5.8975821 29.23635555
## 32              gun show    50      28      2  4.8029265 29.01742765
## 33            power plant    29      21      2  7.2184839 28.87336842
## 34        greenhous gase    48      31      2  9.7966712 28.84928482
## 35          custom enforc    29      15      2  7.3072065 28.13601207
## 36             court rule    51      43      2  4.4522941 28.11886534
## 37            lung cancer    26      11      2  8.4342815 27.89555446
## 38           justic depart    30      18      2  6.0991164 27.72547769
## 39             law enforc    41      38      2  5.1031998 27.26736757
## 40            even though    27      19      2  6.6185678 26.34913210
## 41              next week    30      25      2  5.5849231 26.12758018
## 42              u. suprem    35      35      2  5.0254418 25.91911041
## 43          north carolina    32      25      2  9.8410654 25.83258144
## 44             nation rifl    41      41      2  7.2065379 25.80405071
## 45              gay coupl    28      16      2  5.6384593 25.75045291
## 46           immigr reform    48      36      2  5.5774404 25.23650641
## 47             green card    20      10      2  8.7635208 25.21553578
## 48             presid bush    30      22      2  5.3344883 25.11635985
## 49            news confer    22       9      2  7.6782077 24.74215408
## 50            district judg    29      29      2  5.2222150 24.71049802
## 51            border patrol    40      30      2  8.4347375 24.65466727
## 52            feder govern    32       8      2  4.8090140 24.36180251
## 53           carri conceal    20      20      2  6.9149459 24.35932301
## 54             health care    22      19      2  6.4226950 24.28443069
## 55              gun owner    32      25      2  5.5635921 24.25470929
## 56          assault weapon    21      12      2  7.8670171 24.21450119
## 57              mental ill    17       9      2  7.9509023 23.97879780
## 58            public health    27      23      2  5.0983932 23.65766825
## 59             associ press    18      12      2  7.1397939 23.57615390
## 60            undocu immigr    40      27      2  5.4071994 23.50399766
## 61              row inmat    18      18      2  6.7255664 23.49893336
## 62             gun violenc    30      16      2  5.8872960 23.22136715
## 63             civil right    25      18      2  5.3881123 23.22046072
```

100

```
## 64               bear arm    16          8   2  8.5851366 23.21257522
## 65          homeland secur    40         27   2  9.4557322 23.08440655
## 66               death row   102         73   2  8.7631782 23.01117623
## 67                 st paul    16         13   2  7.9105873 22.91592191
## 68             appeal court    33         29   2  4.4225013 22.83124604
## 69              u. district    28         28   2  4.8225642 22.80822237
## 70            driver licens    19         12   2  7.6355533 22.67056762
## 71           constitut amend    19        10   2  5.9674320 22.63003327
## 72                next year    32         20   2  4.5793512 22.58164190
## 73             public safeti    20        13   2  6.4535721 22.19376897
## 74         district attorney    22        19   2  5.2527273 22.11871927
## 75              offici said    50         45   2  3.4678290 22.07028879
## 76              carbon emiss    19        14   2  5.7218800 21.99165853
## 77              dioxid emiss    17        17   2  6.5702973 21.95061906
## 78                 park lot    14          7   2  7.2262107 21.90436231
## 79            bush administr    19        10   2  5.6438700 21.84277083
## 80              capitol hill    14         8   2  9.3625065 21.78105769
## 81             without parol    20         20   2  8.0148278 21.70222420
## 82           tobacco product    22        17   2  5.6852688 21.44633385
## 83              citi council    21         17   2  5.4507046 21.42707863
## 84             decid whether    17         6   2  5.9109577 21.40221308
## 85           defens attorney    18         13   2  5.7609659 21.30661154
## 86            domest partner    12         11   2  8.2637277 21.22128015
## 87              electr chair    13         9   2  8.9818526 21.11541632
## 88               illeg alien    21         10   2  6.6037568 21.07354209
## 89           general assembl    15         6   2  7.0298978 21.06244631
## 90              long island    14          7   2  6.7819738 21.04891188
## 91         presidenti candid    13        12   2  7.0332489 21.03424046
## 92                york citi    24         24   2  4.6628146 20.89101468
## 93            detent center    14         12   2  7.1571878 20.88076000
## 94                 san jose    18         14   2  8.6056790 20.81790141
## 95               first time    27         10   2  4.2933707 20.67581817
## 96       environment protect    14        14   2  6.5371859 20.65431433
## 97               next month    19         16   2  5.2141797 20.58499195
## 98               polic offic    18        16   2  5.3032974 20.41880206
## 99                  gun law    41         37   2  3.3850502 20.39995772
## 100            social secur    14         12   2  6.3995468 20.36691145
## 101                al gore    17         14   2  9.9087421 20.15227952
## 102               five year    24         11   2  4.8489707 20.14545173
## 103             found guilti    14         3   2  6.5894546 20.12290376
## 104            obama administr  17        12   2  5.3569785 20.06308867
## 105             famili member    18        12   2  5.1693756 20.03056219
## 106             patrol agent    12         11   2  6.8878597 19.98438064
## 107              young peopl    20         18   2  5.2676247 19.97174872
## 108              polic chief    14         13   2  6.1772924 19.94218218
## 109           cigarett maker    21         13   2  7.8852981 19.92472768
## 110             undocu worker    14         7   2  6.0438293 19.87625673
## 111            lesbian coupl    14         7   2  7.0756173 19.86960241
## 112              presid trump    21        19   2  4.7264309 19.82595970
## 113        domest partnership    11         9   2  9.0854222 19.81277474
## 114                sea level    11         9   2  7.6710817 19.80656887
## 115            execut director    18        16   2  5.2182166 19.73567944
## 116             perman resid    12         11   2  7.2316312 19.70354867
## 117               three year    27         21   2  4.1670109 19.68560806
```

```
## 118             u. citizen    21       6      2  4.7175475 19.65600841
## 119             new jersey    60      45      2  8.1438093 19.64177330
## 120               mr trump    19      16      2  4.9056094 19.62429362
## 121          district court    25      25      2  4.2896320 19.61044014
## 122           civil liberti    15      13      2  8.3193208 19.60087399
## 123              feder judg    25      11      2  4.1760030 19.52440871
## 124            donald trump    29      27      2  9.2306724 19.52237493
## 125             execut order    20      17      2  4.7435940 19.51302499
## 126              vice presid    36      32      2  8.3353731 19.41557288
## 127           mexican border    15       7      2  5.6174738 19.31514592
## 128                 st loui    39      32      2 11.1467133 19.21986204
## 129       tobacco settlement    18      15      2  5.2316351 19.14384338
## 130             plead guilti    12       7      2  9.3988760 19.05920015
## 131                ban smoke    28      12      2  3.8389705 19.05896458
## 132               innoc peopl    18      15      2  5.3228019 19.04436498
## 133              unusu punish    11       3      2  7.6015878 19.03964139
## 134                  georg w    10      10      2  8.9818774 19.00879829
## 135           reduc greenhous    12      12      2  6.0634864 18.81825280
## 136               high court    22      17      2  4.4009628 18.77287221
## 137             protect agenc    13      13      2  5.6015958 18.44003067
## 138             circuit court    18      17      2  5.8171078 18.37782324
## 139              prime minist    12      12      2 10.4172943 18.34587007
## 140                 bay area    11      11      2  7.5614368 18.32102522
## 141            marriag licens    18      16      2  4.6727059 18.31911472
## 142               take effect    16       9      2  4.9196441 18.23757504
## 143                 two year    30      17      2  3.5076326 18.12458860
## 144                immigr law    46      31      2  2.7999927 18.03022146
## 145             florida suprem    15      15      2  4.9971175 17.99858899
## 146              state suprem    28      28      2  3.6954054 17.92438150
## 147           central america    10       3      2  6.6138336 17.92078363
## 148              agent arrest    11      11      2  5.9564243 17.91882892
## 149         judiciari committe    10       9      2  8.3023413 17.88272326
## 150             kyoto protocol    13       5      2 10.5805074 17.87527221
## 151              capit murder    14       6      2  5.1728127 17.82934205
## 152               circuit judg    13      13      2  6.0153498 17.82498598
## 153                month ago    14      12      2  5.1385540 17.73453274
## 154             estim million    13       9      2  5.4997160 17.70419919
## 155          brown williamson    11      11      2  9.7194928 17.69915872
## 156                six month    13       6      2  5.4010192 17.69423050
## 157            prison without    13       9      2  5.2956077 17.67785099
## 158             european union    10       5      2  7.7388132 17.63557841
## 159               told report    16      15      2  4.7148879 17.60246030
## 160             council member    12      11      2  5.5606862 17.60057032
## 161             tuesday night    11       7      2  5.8805639 17.59260467
## 162              major leader    13      10      2  5.2551948 17.57502521
## 163             defens lawyer    12       9      2  5.5084882 17.48801325
## 164                day labor    12       9      2  5.7826752 17.42735448
## 165              legal status    13       6      2  5.4618543 17.39853176
## 166             mental retard    14       5      2  9.6476278 17.29994199
## 167           determin whether    10       5      2  6.2571639 17.26846215
## 168              lethal inject    31      12      2 12.7932718 17.25990501
## 169                natur gas    10       4      2  5.9837316 17.22748687
## 170            cigarett smoke    22      17      2  3.8885184 17.19202612
## 171                make sure    12       5      2  5.9026707 17.14457590
```

```
## 172          court decis   21    18   2  3.9790890 17.07820055
## 173            sen john   10    10   2  5.9047825 17.07562591
## 174            web site   11     7   2  9.9535081 16.99330044
## 175            law firm   15    15   2  5.2877669 16.97770890
## 176              w bush   10    10   2  7.5510879 16.95596434
## 177          marri coupl   11     8   2  5.5677132 16.94279197
## 178       clinton administr   12     9   2  5.2421471 16.92684428
## 179          feder court   26    11   2  3.4748071 16.87588014
## 180         drug administr   11     9   2  5.4991047 16.85004543
## 181            two week   19    16   2  4.0704142 16.84787247
## 182       american citizen   15     9   2  4.6353354 16.80916793
## 183            los angel   28    19   2 13.1285311 16.71495553
## 184         oklahoma citi   13    10   2  6.6187257 16.70737985
## 185             two men   14    10   2  4.8193340 16.68906901
## 186         san francisco   41    32   2 11.0577361 16.67303431
## 187          suprem court  179   153   2 10.6039028 16.60323799
## 188             mr bush   14    13   2  4.6973282 16.51284741
## 189      first-degre murder   20     9   2  8.6144201 16.39743587
## 190            week ago   13    12   2  4.8638016 16.38385639
## 191         border secur   13    10   2  4.7784175 16.24653519
## 192          life sentenc   14     8   2  4.5819538 16.23277863
## 193           three day   14    10   2  4.5493624 16.12477332
## 194         medic school   10    10   2  5.8156297 16.11144003
## 195            mr obama   14    13   2  4.5402032 16.07782548
## 196            citi hall   11     8   2  5.7796829 15.99174932
## 197          big tobacco   13     8   2  4.8411836 15.83768090
## 198          human servic   10     9   2  5.3422967 15.80857870
## 199        nation largest   12    12   2  5.0860876 15.74012009
## 200        republican parti   12     8   2  4.8447318 15.70991928
## 201            four year   17    14   2  4.1266860 15.67463822
## 202           court judg   22    22   2  3.4885379 15.64910373
## 203         million illeg   17    15   2  3.9218815 15.48259551
## 204           u. circuit   11    11   2  5.4076381 15.47001045
## 205            feder law   23    13   2  3.3475815 15.40998983
## 206        trump administr   11     8   2  4.8606059 15.32667564
## 207          human right   12    10   2  4.7205348 15.27303306
## 208          check system   10    10   2  5.0946869 15.26933551
## 209           teen smoke   12     9   2  5.0970775 15.25327265
## 210          cigarett tax   12     8   2  4.6058573 15.20474711
## 211    republican presidenti   10    10   2  5.2533017 15.19642911
## 212            everi day   11    10   2  4.8076483 15.07087842
## 213          feder appeal   15    14   2  4.0505858 15.01914981
## 214            gun lobbi   12     6   2  4.9368691 14.84625274
## 215        reynold tobacco   10     9   2  5.7900923 14.72420540
## 216          school shoot   10    10   2  4.8815702 14.62223777
## 217         gun ownership   15     9   2  6.8260303 14.55438331
## 218         said yesterday   22    22   2  3.3725323 14.54982230
## 219           court order   16     9   2  3.8219384 14.53801581
## 220          public place   11     9   2  4.5402879 14.39352456
## 221            look like   11     7   2  4.5590664 14.34967438
## 222        state attorney   22    22   2  3.2219000 14.33587576
## 223           senat vote   14     9   2  3.9351462 14.21789652
## 224         attorney offic   11    11   2  4.4293569 14.14873896
## 225          feder district   13    13   2  4.0822446 14.14372679
```

```
## 226             least one   13     11      2  4.1378701 14.10026287
## 227          carbon dioxid   54     42      2 11.7948451 14.09696234
## 228           philip morri   37     32      2 14.7022303 14.06067306
## 229         advocaci group   14     13      2  8.3415872 13.92472399
## 230              new trial   16      4      2  3.6433775 13.89723757
## 231          global climat   11     10      2  4.3356556 13.88249895
## 232             take place   10      5      2  4.5472425 13.82750270
## 233            recent year   15      5      2  3.7718387 13.82317965
## 234              gun buyer   10      5      2  5.6262369 13.81774514
## 235           enforc offici   11     11      2  4.2802461 13.69224376
## 236              less like   10      4      2  4.4867971 13.56205694
## 237            legal resid   10      7      2  4.4606932 13.53114034
## 238             unit nation   16     13      2  3.4703800 13.47520578
## 239              u. border   14      9      2  3.7190566 13.47246044
## 240          counti district  11     11      2  4.1817498 13.42719139
## 241            legal immigr   24     16      2  2.8475064 13.39439934
## 242               new york  190    153      2 10.9786020 13.37435748
## 243              state law   31     20      2  2.4748988 13.34603064
## 244           execut action   10      9      2  4.3815790 13.31521054
## 245            amend right   11     11      2  4.1616546 13.31154404
## 246             year later   12     11      2  4.0988911 13.23132901
## 247              right now   14     12      2  3.6251559 13.19039247
## 248             three time   12      8      2  3.9313722 13.17584686
## 249       republican leader   10      8      2  4.3022420 13.16136084
## 250             marco rubio   12     11      2 11.7791182 13.07833060
## 251             feder build   10      6      2  4.2856944 13.02844629
## 252              new studi   15      7      2  3.4898521 12.99994799
## 253            penalti case   14     14      2  3.5501392 12.93251904
## 254            countri illeg   14      5      2  3.4997694 12.76841750
## 255             sandi hook   12     12      2 13.6037211 12.70393737
## 256               ban gay   11     11      2  3.9298277 12.70135569
## 257             fossil fuel   14      3      2 11.0987549 12.69939178
## 258        elementari school   13     12      2  8.6449484 12.49854009
## 259            marriag act   10      6      2  4.0280845 12.38680551
## 260              bill pass   10      8      2  4.0156056 12.32676145
## 261              first day   11      2      2  3.7632851 12.22039657
## 262            feder author   10      4      2  3.8615732 11.93415286
## 263              u. immigr   24     10      2  2.5044860 11.91749328
## 264              past year   10      1      2  3.9216107 11.74965063
## 265                new law   21     14      2  2.5482801 11.43454813
## 266              gay right   10     10      2  3.6400255 11.32620981
## 267        state legislatur   10      3      2  3.8542503 11.25941097
## 268              york time   10     10      2  3.6145484 11.23883685
## 269            state univers   13     11      2  3.2546597 11.23412705
## 270            immigr polici   15     14      2  2.9994335 11.17733756
## 271            immigr agent   12      7      2  3.3766940 11.07704298
## 272              law requir   11     10      2  3.3974071 11.02427309
## 273            million peopl   12      3      2  3.2278703 10.99950331
## 274              everi year   10     10      2  3.5850269 10.95127463
## 275             state senat   17     15      2  2.7157361 10.90134723
## 276            legal right   11      2      2  3.3238024 10.88204326
## 277           immigr status   10      6      2  3.6916987 10.84653118
## 278            state depart   14      7      2  2.9750806 10.78000458
## 279               carri gun   10      3      2  3.4916639 10.77881656
```

```
## 280             u. attorney  10        10       2  3.4057854 10.63558008
## 281        secondhand smoke  25        14       2  8.9400504 10.62681410
## 282           clara counti   14        14       2  9.0707546 10.54919438
## 283               judg rule  10         8       2  3.3311650 10.43662725
## 284               gun right  14        12       2  2.8066126 10.35049908
## 285            florida death  10        10       2  3.2874994 10.23100162
## 286            immigr servic  13        11       2  2.8799634 10.07540091
## 287              said tuesday  14        14       2  2.7833049 10.02524777
## 288                 u. senat  10         7       2  3.1878025 10.00141116
## 289              clinton said  13        12       2  2.8698698  9.92261154
## 290               law school  11         9       2  3.0139657  9.88510375
## 291              allow smoke  10         7       2  3.1391855  9.82515332
## 292                mani peopl  11         6       2  2.9541723  9.71991875
## 293            american peopl  11         7       2  2.9408000  9.67825714
## 294              state offici  15        13       2  2.5076778  9.53932225
## 295                 just one  10         9       2  3.0345183  9.52587687
## 296                offici say  10         9       2  2.9793272  9.37421055
## 297              tobacco use  10         8       2  2.9456239  9.27660315
## 298                 las vega  12         6       2 15.2131672  9.17944311
## 299            immigr offici  15        12       2  2.4059563  9.15803332
## 300             santa clara   15        15       2 13.3913431  9.12968528
## 301            immigr system  11         7       2  2.7887623  9.04383040
## 302              said monday  12        12       2  2.6796599  9.00939162
## 303               r.j reynold  14        11       2 13.0928419  8.95440346
## 304               york state  11        11       2  2.6208166  8.57343140
## 305               palm beach  18        17       2 12.3091870  8.53745226
## 306            prosecutor said  10         7       2  2.7612876  8.47290826
## 307               polic said  12        10       2  2.4988750  8.45754816
## 308             immigr author  10         8       2  2.5992947  8.11148430
## 309               state illeg  14        14       2  2.1806790  8.08715415
## 310              aldap guerra  10        10       2 16.1374385  7.97434099
## 311                san diego  15        10       2 11.3391301  7.87042406
## 312              barack obama  18        16       2 11.0648694  7.70706096
## 313           immigr children  10         7       2  2.4174027  7.58475938
## 314           surgeon general  14        13       2 10.8799216  7.55019279
## 315                 gun bill  10         6       2  2.3679711  7.50765954
## 316                 jeb bush  11        11       2 10.6944974  7.38894237
## 317               immigr judg  12         9       2  2.1135861  7.28238556
## 318              feder immigr  13        13       2  2.0173161  7.23572644
## 319               immigr bill  13        11       2  1.9431298  6.97738277
## 320             superior court  22        22       2  9.8263796  6.86949856
## 321                 also said  14         6       2  1.8304914  6.79693127
## 322             new hampshir   11         7       2  8.9762908  6.21143481
## 323                group said  10         8       2  1.8351035  5.80449019
## 324                report said  10         4       2  1.7447507  5.52851218
## 325          u. district judg  19         0       3  1.5070590  1.84718192
## 326         nation rifl associ  38         0       3  2.3873439  1.82309762
## 327 environment protect agenc  12         0       3  2.6485016  1.48331829
## 328          feder appeal court  12         0       3  0.9215788  1.26221019
## 329            unit state illeg  13         0       3  1.6293495  0.97967263
## 330          greenhous gas emiss  32         0       3  1.4635987  0.86725050
## 331         state attorney general  10         0       3  0.7620860  0.51203311
## 332              georg w bush  10         0       3  0.9520101  0.44226468
## 333             death row inmat  18         0       3 -0.1884630 -0.09175967
```

```
## 334                new york citi    24            0        3 -0.4244976 -0.19542314
## 335            death penalti case    14            0        3 -0.4565222 -0.29498220
## 336               ban gay marriag    10            0        3 -0.4518635 -0.45797218
## 337                new york time    10            0        3 -1.2363369 -0.56595757
## 338          million illeg immigr    15            0        3 -0.4525303 -0.57711656
## 339               new york state    11            0        3 -1.1491213 -0.62058135
## 340           border patrol agent    11            0        3 -1.1516453 -0.67489874
## 341           superior court judg    14            0        3 -2.0026996 -0.96471566
## 342            santa clara counti    14            0        3 -2.5944974 -1.10616734
## 343              u. circuit court    10            0        3 -1.1202972 -1.14751203
## 344            state suprem court    28            0        3 -1.9571417 -1.23165429
## 345           reduc greenhous gase    10            0        3 -3.0706107 -1.89361472
## 346           florida suprem court    15            0        3 -3.3468296 -2.01259588
## 347             carbon dioxid emiss    16            0        3 -4.4404477 -2.03202821
## 348               u. suprem court    35            0        3 -3.2499897 -2.05894509
## 349              suprem court rule    17            0        3 -4.9681113 -3.14166142
## 350           carri conceal weapon    12            0        3 -3.8560733 -4.57681400
```

Finally, we can add collocations to the original tokens object (single n-grams), and remove unnecessary padding we preserved in the initial core tokens object.

```
# Add to tokens document
news_toks_all <- tokens_compound(toks_news_subs, news_corp_coll_no_stop, concatenator = " ") %>%
    tokens_select(padding=FALSE) %>% # remove padding
    tokens_wordstem # stem words after adding collocations
tail(news_toks_all)
```

```
## Tokens consisting of 6 documents and 9 docvars.
## text995 :
##  [1] "game"      "nra"       "parallel" "imag"      "told"      "stori"
##  [7] "power"     "word"      "newtown"  "conn"      "monday"    "mother"
## [ ... and 126 more ]
##
## text996 :
##  [1] "assault"           "rifl"              "get"               "support"
##  [5] "g.o.p"             "hous"              "move"              "democrat"
##  [9] "denounc"           "bow"               "gun lobbi"         "republican lead"
## [ ... and 117 more ]
##
## text997 :
##  [1] "advi"              "urg"                "new jersey"
##  [4] "high court"        "chang"              "appeal"
##  [7] "death sentenc"     "special"            "judg"
## [10] "appoint"           "state suprem court" "studi"
## [ ... and 130 more ]
##
## text998 :
##  [1] "gun"       "money"     "trump"    "public"    "grim"      "afterglow"
##  [7] "littleton" "kill"      "senat"    "stumbler"  "spent"     "much"
## [ ... and 149 more ]
##
## text999 :
##  [1] "state"     "weigh"     "make"     "beach"     "cleaner"   "smoke-fr"
##  [7] "zone"      "roust"     "restaur"  "bounc"     "bar"       "smoker"
## [ ... and 140 more ]
```

```
## 
## text1000 :
##  [1] "wal-mart"      "death penalti" "one"           "disput"
##  [5] "circumst"      "peopl"         "fundament"     "right"
##  [9] "assert"        "moral"         "religi"        "object"
## [ ... and 144 more ]
```

Regular Expressions

Oftentimes, we identify a pattern in a body of text from the DGP that could lead to false conclusions. Say that we want to create a network linking members of Congress to other speakers they address in their legislative speeches to uncover informal coalitions among members. Creating this network would involve using computational text methods to identify members of Congress that come up in the text. But members of Congress often yield their time to colleagues at the end of their speech – a procedure that may have little to do with coalition building. Leaving the statement where time is yielded would introduce considerable noise and error into the analysis.

One way we can handle such problematic patterns is editing the text before turning it into a corpus. We can use basic functions in R (like `gsub` and `grep`) and packages like `stringi` and `stringr` to edit text, relying heavily on regular expressions (RegEx) to edit out problematic text. Regular expressions are useful because they can identify a sequence of characters that specifies a search pattern in text. Once that search pattern has been identified, we can use commands in R to make changes to the text. Like many programming opportunities, RegEx can be frustrating to work with but has incredible potential. Websites like *https://regex101.com/*, online R resources for data science – Robert Peng's explanation of regular expressions in R, for example or Dave Child's cheat sheet, and Stack Overflow are your friends when you need to edit text.

It's important to note that in many situations, patterns that stem from the DGP will *not* cause systematic bias in a text analysis. In the example below, we demonstrate one way to use RegEx to remove the text that comes before the phrase "legislative counsel's digestab" in the bills data, including the title of the bill and other preamble. Whether these terms and preamble are problematic depends heavily on the research question. Whereas the phrase would likely cause few inferential challenges in a topic model used to measure issue emphases in legislative bills, it might be problematic in a word frequency count used to measure the distribution of venues legislators pay attention to in policymaking.

```r
# Using regular expressions to drop common phrase and preamble in bill text
bills_data <- readRDS("sample_bills_docid.rds")
bills_data$text[2]
```

```
## [1] "\n                    california legislature-\n                    2007-2008 regular session\n
```

```r
bills_data$text_processed <- gsub(".*legislative counsel's digest", "", bills_data$text)
  # Create new column for edited text to preserve original data

# Check a sample of the processed text
bills_data$text_processed[2]
```

```
## [1] "sb 267, as introduced, romero.\n                    department of corrections and rehabili
```

Homework

Discussion Question

1. When might collocation analysis be problematic on a tokens object? Explain in a few sentences.

Project Questions (show your work)

1. Pick a corpus. You could pick your own corpus or, if you want a suggestion, you could make a corpus for the twitter immigration data ("sample_immigration_tweets_2013-2017.csv") and save the corpus as an RDS object.

2. Identify an analysis goal. For example, using the twitter immigration data, you could subset the corpus to years 2016-2017 and aim to identify the most frequently used terms/concepts in immigration twitter data during the change of presidential administrations.

3. Tokenize and pre-process your corpus with your goal in mind, explaining these decisions.

4. Repeat Step 4 but making different (yet still rational) pre-processing decisions, again explaining your decisions.

5. Finally, discuss: what worked and what didn't, and why? Explain the similarities and differences between your two approaches and what you learned.