

# cao\_\_problem\_\_set\_\_5

Steven Cao

11/27/2019

## Problem 1

```
# load csv
meta <- read.csv("C:/Users/Rosenberg Lab/Documents/Problem-Set-5/platforms.csv")

## Warning in read.table(file = file, header = header, sep = sep, quote =
## quote, : incomplete final line found by readTableHeader on 'C:/Users/
## Rosenberg Lab/Documents/Problem-Set-5/platforms.csv'

# load individual text files into a corpus
texts <- file.path("C:/Users/Rosenberg Lab/Documents/Problem-Set-5", "Party Platforms Data")
docs <- VCorpus(DirSource(texts))
```

## Problem 2

```
# remove punctuation + extras that may have been missed
docs <- tm_map(docs, removePunctuation)
for (j in seq(docs)) {
  docs[[j]] <- gsub("/", " ", docs[[j]])
  docs[[j]] <- gsub("'", "", docs[[j]])
  docs[[j]] <- gsub("-", " ", docs[[j]])
  docs[[j]] <- gsub("\\\\|", "", docs[[j]])
  docs[[j]] <- gsub("@", "", docs[[j]])
  docs[[j]] <- gsub("\\u2028", "", docs[[j]])
  docs[[j]] <- gsub("â\\200", "", docs[[j]])
  docs[[j]] <- gsub("\"", "", docs[[j]])
  docs[[j]] <- gsub(" ", " ", docs[[j]])
}

# remove numbers
docs <- tm_map(docs, removeNumbers)

# replace uppercase with lowercase
docs <- tm_map(docs, tolower)
docs <- tm_map(docs, PlainTextDocument)

# remove stop words and also "will" since that's pretty common
docs <- tm_map(docs,
  removeWords,
  stopwords("english"))
docs <- tm_map(docs, removeWords, c("will"))
docs <- tm_map(docs, PlainTextDocument)
```

```
# remove extra whitespace
docs <- tm_map(docs, stripWhitespace)
docs <- tm_map(docs, PlainTextDocument)
```

### Problem 3

```
## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : strong could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : promote could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : democratic could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : infrastructure could not be fit on page. It will not be
## plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : institutions could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : opportunities could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : citizens could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : community could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : expanding could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : violence could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : businesses could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : goodpaying could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : making could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : ensuring could not be fit on page. It will not be plotted.
```

```

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : practices could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : recognize could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : stronger could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : without could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : countries could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : crisis could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : democracy could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : important could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : native could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : partners could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : poverty could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : protecting could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : young could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : allies could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : lowincome could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[1], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : oppose could not be fit on page. It will not be plotted.

```



```
## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,  
## "Dark2"), : american could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : republican could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,  
## "Dark2"), : americans could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,  
## "Dark2"), : security could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,  
## "Dark2"), : economy could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,  
## "Dark2"), : military could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : administration could not be fit on page. It will not be
## plotted.
```

```
## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,  
## "Dark2"), : private could not be fit on page. It will not be plotted.
```

```

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : country could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : religious could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : amendment could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : policies could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : protect could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : growth could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : need could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : programs could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : republicans could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : ensure could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : especially could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : foreign could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : governments could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : political could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : including could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : make could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : access could not be fit on page. It will not be plotted.

```

```

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : advance could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : constitutional could not be fit on page. It will not be
## plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : healthcare could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : system could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : control could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : jobs could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : many could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : power could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : upon could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : communities could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : legislation could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : trade could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : believe could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : defense could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : encourage could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : now could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : resources could not be fit on page. It will not be plotted.

```

```

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : without could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : administrations could not be fit on page. It will not be
## plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : best could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : business could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : children could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : institutions could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : local could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : service could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : also could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : continue could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : citizens could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : constitution could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : policy could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : property could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : democratic could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : development could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : financial could not be fit on page. It will not be plotted.

```

```

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : international could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : officials could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : assistance could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : abortion could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : millions could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : program could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : regulations could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : among could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : authority could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : based could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : become could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : end could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : family could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : forces could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : individuals could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : liberty could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : provide could not be fit on page. It will not be plotted.

```



```

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : reform could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : respect could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : schools could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : way could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : women could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : workers could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : countrys could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : democrats could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : department could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : efforts could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : internet could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : leadership could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : social could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : society could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : approach could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : environmental could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : innovation could not be fit on page. It will not be plotted.

```

```
## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : investment could not be fit on page. It will not be plotted.

## Warning in wordcloud(docs[2], max.words = 150, colors = brewer.pal(8,
## "Dark2"), : protection could not be fit on page. It will not be plotted.
```



Democrats seem to focus more on “interpersonal relationships” that focus on helping one another, e.g. words like “believe, public, support, care, communities,” and so on. Republicans seem to focus on more “national infrastructure” ideas, e.g. words like “government, federal, state, economic, administration,” and so on.

## Problem 4

```
dem_df <- tibble(dem)
rep_df <- tibble(rep)

dem_ut <- dem_df %>%
  unnest_tokens(word,dem)
rep_ut <- rep_df %>%
  unnest_tokens(word,rep)

dem_sentiment_bing <- dem_ut %>%
  inner_join(get_sentiments("bing")) %>%
  count(sentiment) %>%
```

```

spread(sentiment, n, fill=0) %>%
mutate(sentiment = positive - negative)

## Joining, by = "word"

dem_sentiment_bing

## # A tibble: 1 x 3
##   negative positive sentiment
##   <dbl>      <dbl>      <dbl>
## 1      826      1415        589

rep_sentiment_bing <- rep_ut %>%
  inner_join(get_sentiments("bing")) %>%
  count(sentiment) %>%
  spread(sentiment, n, fill=0) %>%
  mutate(sentiment = positive - negative)

## Joining, by = "word"

rep_sentiment_bing

## # A tibble: 1 x 3
##   negative positive sentiment
##   <dbl>      <dbl>      <dbl>
## 1     1258     1612        354

dem_sentiment_afinn <- dem_ut %>%
  inner_join(get_sentiments("afinn")) %>%
  count(value) %>%
  spread(value, n, fill=0)

## Joining, by = "word"

dem_sentiment_afinnTotal <- sum( as.double(names(dem_sentiment_afinn)) * dem_sentiment_afinn )
dem_sentiment_afinnTotal

## [1] 1326

rep_sentiment_afinn <- rep_ut %>%
  inner_join(get_sentiments("afinn")) %>%
  count(value) %>%
  spread(value, n, fill=0)

## Joining, by = "word"

```

```
rep_sentiment_afinnTotal <- sum( as.double(names(rep_sentiment_afinn)) * rep_sentiment_afinn )
rep_sentiment_afinnTotal
```

```
## [1] 918
```

## Problem 5

The democrats seem to be more optimistic about the future on the whole. Or at least, they have a more positive current mindset. This would make sense given what I saw in the word clouds earlier: Democrats focus more on social relationships (which have a greater affinity to emotionality) while Republicans focus more on impersonal factors like the economy, which would be more emotionally neutral (i.e. “stating the facts for how they are”).

## Problem 6

```
dem_dtm <- DocumentTermMatrix(docs[1])
rep_dtm <- DocumentTermMatrix(docs[2])

dem_ntokens <- dem_dtm$ncol
rep_ntokens <- rep_dtm$ncol

dem_ldaRaw.5 <- LDA(dem_dtm, k=5, method="gibbs")
rep_ldaRaw.5 <- LDA(rep_dtm, k=5, method="gibbs")

tidy(dem_ldaRaw.5)
```

```
## # A tibble: 19,255 x 3
##   topic term      beta
##   <int> <chr>    <dbl>
## 1     1 aapi  0.000453
## 2     2 aapi  0.0000128
## 3     3 aapi  0.0000404
## 4     4 aapi  0.0000424
## 5     5 aapi  0.0000452
## 6     1 abandon 0.0000412
## 7     2 abandon 0.000141
## 8     3 abandon 0.00125
## 9     4 abandon 0.000466
## 10    5 abandon 0.0000452
## # ... with 19,245 more rows
```

```
tidy(rep_ldaRaw.5)
```

```
## # A tibble: 26,230 x 3
##   topic term      beta
##   <int> <chr>    <dbl>
## 1     1 abandon 0.0000321
## 2     2 abandon 0.0000307
## 3     3 abandon 0.0000285
```

```
## 4      4 abandon  0.0000101
## 5      5 abandon  0.000696
## 6      1 abandoned 0.000353
## 7      2 abandoned 0.0000307
## 8      3 abandoned 0.000313
## 9      4 abandoned 0.0000101
## 10     5 abandoned 0.0000331
## # ... with 26,220 more rows
```

## Problem 7

```
terms(dem_ldaRaw.5, 10)
```

```
##      Topic 1      Topic 2      Topic 3      Topic 4
## [1,] "service"    "democrats" "donald"   "build"
## [2,] "well"       "health"    "back"     "nuclear"
## [3,] "better"     "support"   "fair"     "ensuring"
## [4,] "change"     "believe"  "expanding" "infrastructure"
## [5,] "businesses" "people"    "immigration" "protecting"
## [6,] "like"       "americans" "improve"   "country"
## [7,] "military"   "must"      "invest"    "living"
## [8,] "trump"      "american"  "mental"    "civil"
## [9,] "partners"   "also"      "strong"    "good"
## [10,] "years"     "communities" "veterans"  "get"
##      Topic 5
## [1,] "democrats"
## [2,] "americas"
## [3,] "goodpaying"
## [4,] "financial"
## [5,] "full"
## [6,] "lowincome"
## [7,] "growth"
## [8,] "crisis"
## [9,] "recognize"
## [10,] "class"
```

```
terms(rep_ldaRaw.5, 10)
```

```
##      Topic 1      Topic 2      Topic 3      Topic 4      Topic 5
## [1,] "jobs"      "institutions" "make"      "government" "upon"
## [2,] "trade"     "democrats"    "tax"       "federal"    "based"
## [3,] "veterans"  "forces"       "access"    "must"       "democratic"
## [4,] "protecting" "way"          "need"      "states"     "authority"
## [5,] "long"      "safe"         "service"   "american"   "just"
## [6,] "small"     "become"       "policy"    "support"    "international"
## [7,] "greater"   "create"       "among"     "people"     "end"
## [8,] "resources" "china"        "nations"   "national"   "justice"
## [9,] "longer"    "taxpayers"   "millions"  "rights"     "drug"
## [10,] "strong"   "relations"   "encourage" "congress"   "must"
```

There is substantial overlap in what they talk about (e.g. the economy, international relations/security, global economy), but they tend to go about it in different angles (e.g. democrats discuss “communities” in relation

to “support, believe, health”, while republicans discuss “community” in relation to “protecting, trust”) and, beyond that, they talk about different topics (e.g. democrats on education, republicans on economics).

## Problem 8

```
dem_ldaRow.5 <- LDA(dem_dtm, k=5, method="gibbs")
rep_ldaRow.5 <- LDA(rep_dtm, k=5, method="gibbs")
dem_ldaRow.10 <- LDA(dem_dtm, k=10, method="gibbs")
rep_ldaRow.10 <- LDA(rep_dtm, k=10, method="gibbs")
dem_ldaRow.25 <- LDA(dem_dtm, k=25, method="gibbs")
rep_ldaRow.25 <- LDA(rep_dtm, k=25, method="gibbs")

tidy(dem_ldaRow.5)
```

```
## # A tibble: 19,255 x 3
##   topic term      beta
##   <int> <chr>    <dbl>
## 1     1 aapi  0.0000128
## 2     2 aapi  0.000500
## 3     3 aapi  0.0000392
## 4     4 aapi  0.0000444
## 5     5 aapi  0.0000414
## 6     1 abandon 0.0000128
## 7     2 abandon 0.000500
## 8     3 abandon 0.00161
## 9     4 abandon 0.0000444
## 10    5 abandon 0.0000414
## # ... with 19,245 more rows
```

```
tidy(rep_ldaRow.5)
```

```
## # A tibble: 26,230 x 3
##   topic term      beta
##   <int> <chr>    <dbl>
## 1     1 abandon  0.000739
## 2     2 abandon  0.0000321
## 3     3 abandon  0.0000295
## 4     4 abandon  0.00000972
## 5     5 abandon  0.0000317
## 6     1 abandoned 0.000739
## 7     2 abandoned 0.0000321
## 8     3 abandoned 0.0000295
## 9     4 abandoned 0.00000972
## 10    5 abandoned 0.0000317
## # ... with 26,220 more rows
```

```
tidy(dem_ldaRow.10)
```

```
## # A tibble: 38,510 x 3
##   topic term      beta
```

```
##      <int> <chr>      <dbl>
## 1      1 aapi  0.0000842
## 2      2 aapi  0.0000850
## 3      3 aapi  0.000887
## 4      4 aapi  0.0000725
## 5      5 aapi  0.000110
## 6      6 aapi  0.0000118
## 7      7 aapi  0.0000851
## 8      8 aapi  0.0000848
## 9      9 aapi  0.0000824
## 10     10 aapi  0.0000792
## # ... with 38,500 more rows
```

```
tidy(rep_ldaRow.10)
```

```
## # A tibble: 52,460 x 3
##   topic term      beta
##   <int> <chr>      <dbl>
## 1      1 abandon 0.0000613
## 2      2 abandon 0.0000646
## 3      3 abandon 0.0000598
## 4      4 abandon 0.0000602
## 5      5 abandon 0.000671
## 6      6 abandon 0.0000640
## 7      7 abandon 0.0000662
## 8      8 abandon 0.00000942
## 9      9 abandon 0.0000571
## 10     10 abandon 0.000606
## # ... with 52,450 more rows
```

```
tidy(dem_ldaRow.25)
```

```
## # A tibble: 96,275 x 3
##   topic term      beta
##   <int> <chr>      <dbl>
## 1      1 aapi  0.000181
## 2      2 aapi  0.000151
## 3      3 aapi  0.000118
## 4      4 aapi  0.000141
## 5      5 aapi  0.000118
## 6      6 aapi  0.000143
## 7      7 aapi  0.000135
## 8      8 aapi  0.000162
## 9      9 aapi  0.000153
## 10     10 aapi  0.000156
## # ... with 96,265 more rows
```

```
tidy(rep_ldaRow.25)
```

```
## # A tibble: 131,150 x 3
##   topic term      beta
##   <int> <chr>      <dbl>
```

```
## 1      1 abandon 0.000124
## 2      2 abandon 0.000131
## 3      3 abandon 0.000103
## 4      4 abandon 0.0000998
## 5      5 abandon 0.000121
## 6      6 abandon 0.000107
## 7      7 abandon 0.00140
## 8      8 abandon 0.0000978
## 9      9 abandon 0.000108
## 10     10 abandon 0.000108
## # ... with 131,140 more rows
```

## Problem 9

```
dem_perp.5 <- exp(-1 * (dem_ldaRaw.5@loglikelihood / dem_ntokens) )
rep_perp.5 <- exp(-1 * (rep_ldaRaw.5@loglikelihood / rep_ntokens) )
dem_perp.10 <- exp(-1 * (dem_ldaRaw.10@loglikelihood / dem_ntokens) )
rep_perp.10 <- exp(-1 * (rep_ldaRaw.10@loglikelihood / rep_ntokens) )
dem_perp.25 <- exp(-1 * (dem_ldaRaw.25@loglikelihood / dem_ntokens) )
rep_perp.25 <- exp(-1 * (rep_ldaRaw.25@loglikelihood / rep_ntokens) )

dem_perp.5
```

```
## [1] 1.915807e+12
```

```
rep_perp.5
```

```
## [1] 2.924597e+12
```

```
dem_perp.10
```

```
## [1] 1.66351e+12
```

```
rep_perp.10
```

```
## [1] 2.480284e+12
```

```
dem_perp.25
```

```
## [1] 2.106974e+12
```

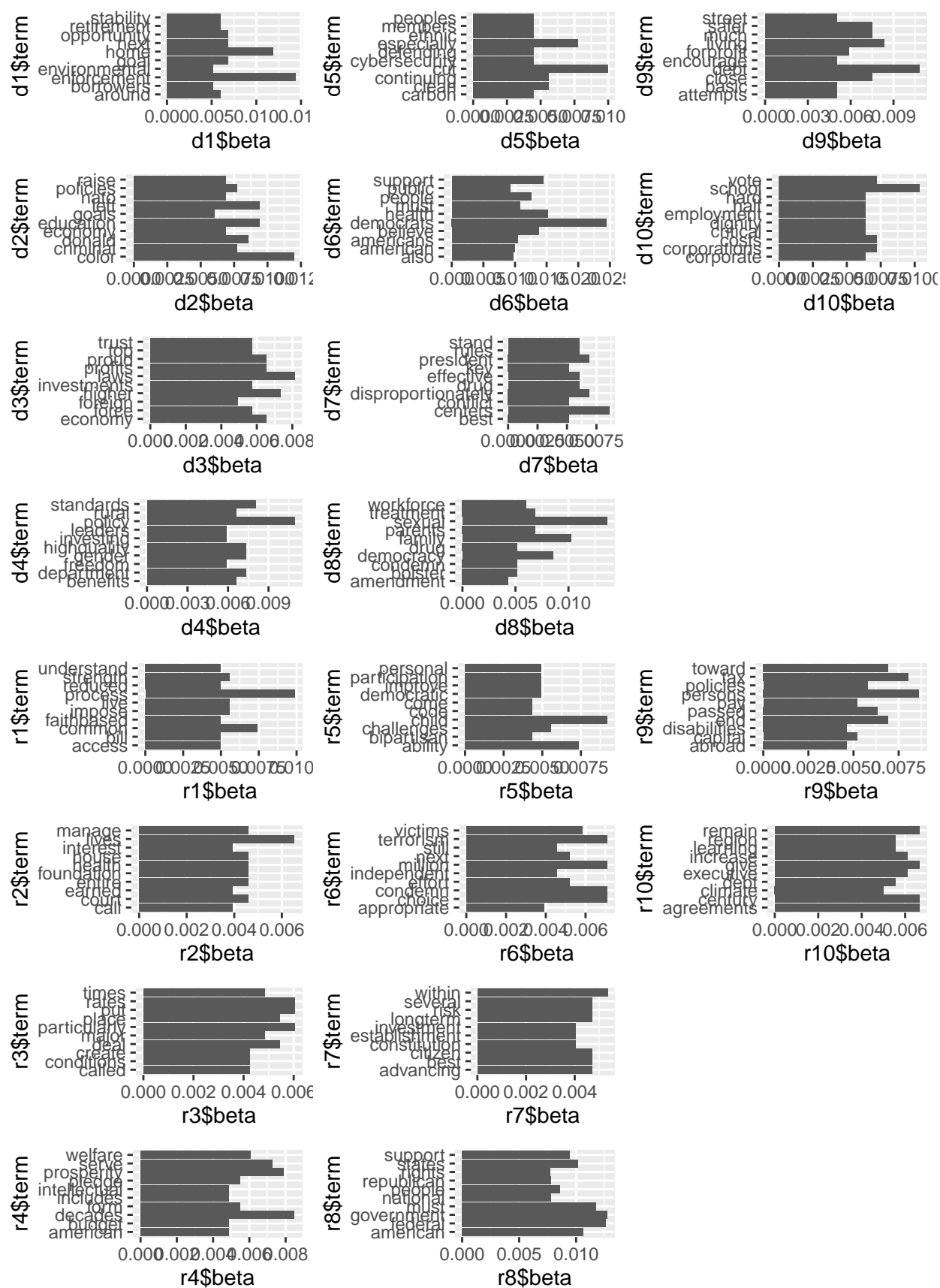
```
rep_perp.25
```

```
## [1] 3.20132e+12
```

When 'k=10', the perplexity is the lowest for both Democrats and Republicans.



## Problem 10



Some themes which the Democrats hone in on: “reform/medical/gun/crisis, promote/industry, safety/minorities, change/historic, support/people/communities, share/income/international/immigrants, stop/gaps/african, promoting/global/earn/private, employment/discrimination, party/isis/pollution”.

Some themes which the Republicans hone in on: “citizen/deserve/sovereignty, agencies/workforce, cooperation/trade, federal/government, liberty, standards/century, global/sector/risk, borders/beliefs, responsibility/climate/workers/lands, important/housing/status/consumer”.

So what they seem to have in common are: the unity/identity of the American people, international issues, and economics.

Despite ‘k=10’ having somewhat lower perplexity scores than ‘k=5’, they are not necessarily more efficient at displaying party differences (i.e. visualising their differences in a neat summary fashion), because it’s easier to keep 5 topics/themes in mind when comparing two platforms than keeping 10 topics/themes in mind. In other words, a fewer-topic model is simpler (and likely, simpler to interpret) than a model with more topics.

## **Problem 11**

Probably the Republicans since their topics appear to be more country-centric than not.