

2016 Presidential Elections Based on Party Platforms

Chuck Cartledge

August 30, 2016 at 10:32pm

Table of contents I

1 Explorations

2 Analysis and predictions

3 Conclusion

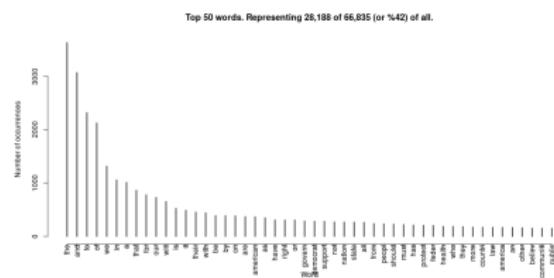
4 Files

Textual mining

Raw text/tokens from political platforms

Perform elementary “data wrangling” on the text:

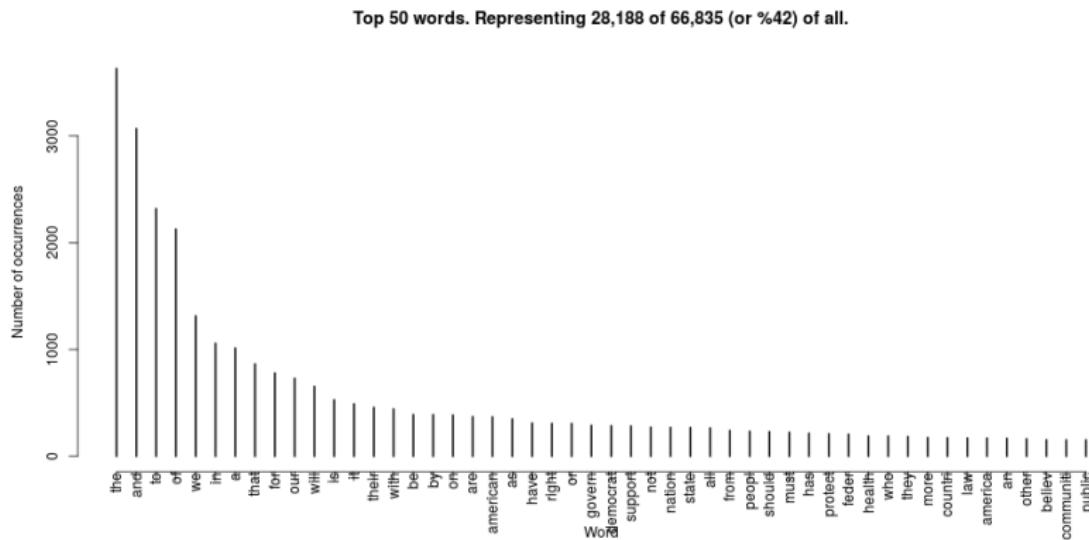
- Splitting the pdf into individual words,
- Converting the words to lower case,
- Removing punctuation marks,
- Changing a word to its stem where possible, and
- Removing numbers



See the impact of “stop words.”

Textual mining

Same image.

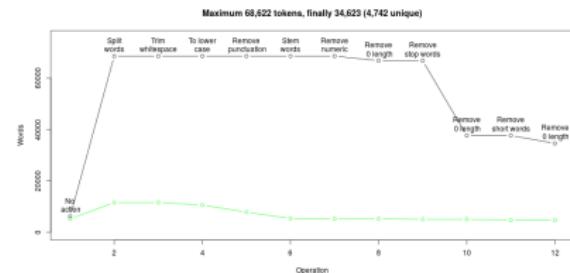


See the impact of “stop words.”

Textual mining

Document the exploration

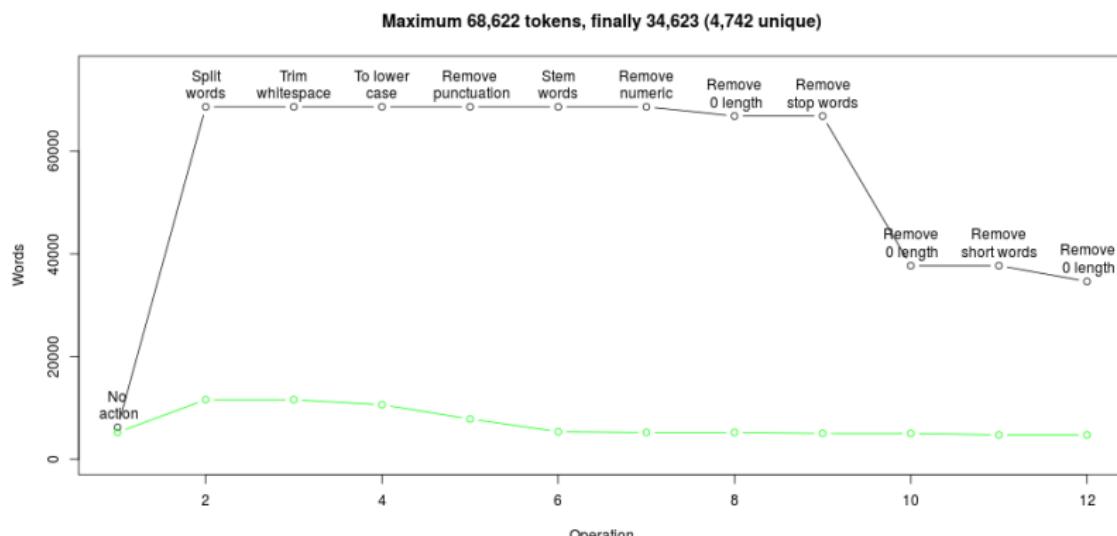
Write a program to show the impact of different text applications on the data.
Structure the program to make it easy to change the order.



Exploration program is attached.

Textual mining

Same image.



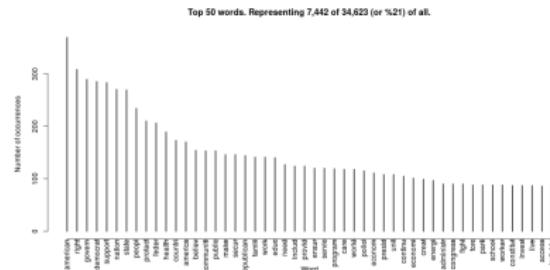
Exploration program is attached.

A code snippet

```
funcs <- c(noAction,  
           splitOnSpaces,  
           trimWhitespace,  
           toLowerCase,  
           removePunctuations,  
           stemify,  
           removeNumeric, removeZeroLength,  
           removeStopWords, removeZeroLength,  
           removeShortWords, removeZeroLength  
           )  
  
for (f in funcs)  
{  
  temp <- f(words, stopWords=stopWords)  
  
  words <- temp$words  
  pointLabel <- c(pointLabel, temp$operation)  
  pointCount <- c(pointCount, length(words))  
  uniqueCount <- c(uniqueCount, length(unique(words)))  
}  
}
```

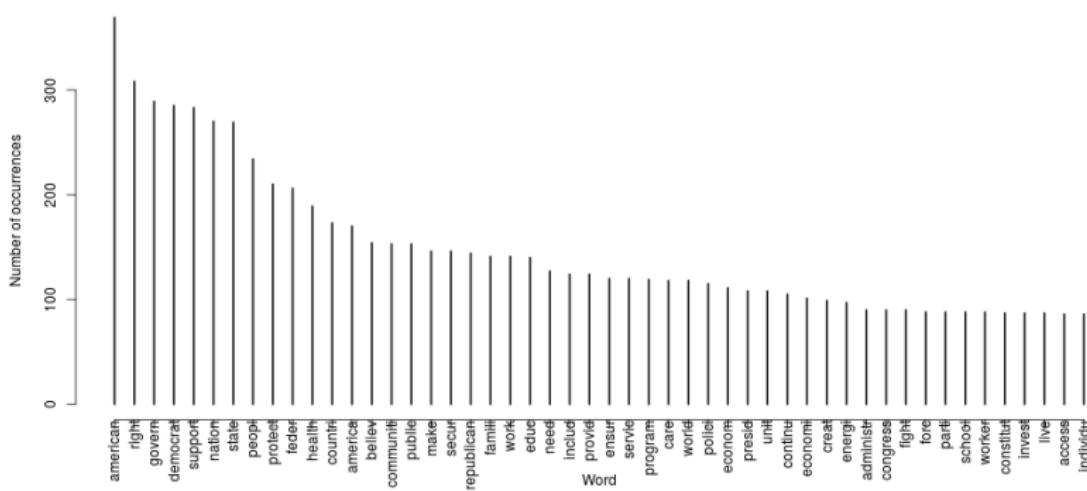
“Processed” text/tokens from political platforms

All the “stop words” have been removed. The mechanical processing of text has been done. Remaining text/tokens are “important.”



Textual mining

Same image.



Assumptions

A set of assumptions

- Each party has a dedicated core (65%)
- Each platform has positive and negative words
- Words can sway undecided voters
- Not all states require party registration
- Percentage of eligible voters is unknown

Every model has assumptions.



Assumptions

Not all political parties are considered

The three largest parties were used in the model:

- Democratic
- Libertarian
- Republican

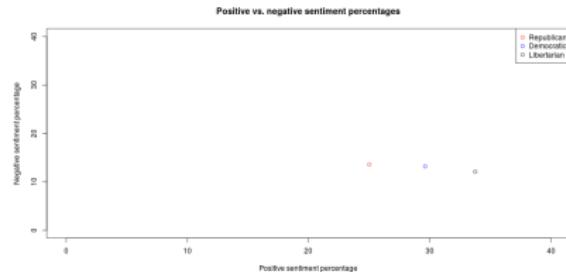
Other parties were considered, but believed to be very much niche organizations and would not have a significant on the outcome.

Prediction

Sentiment analysis of the party platforms.

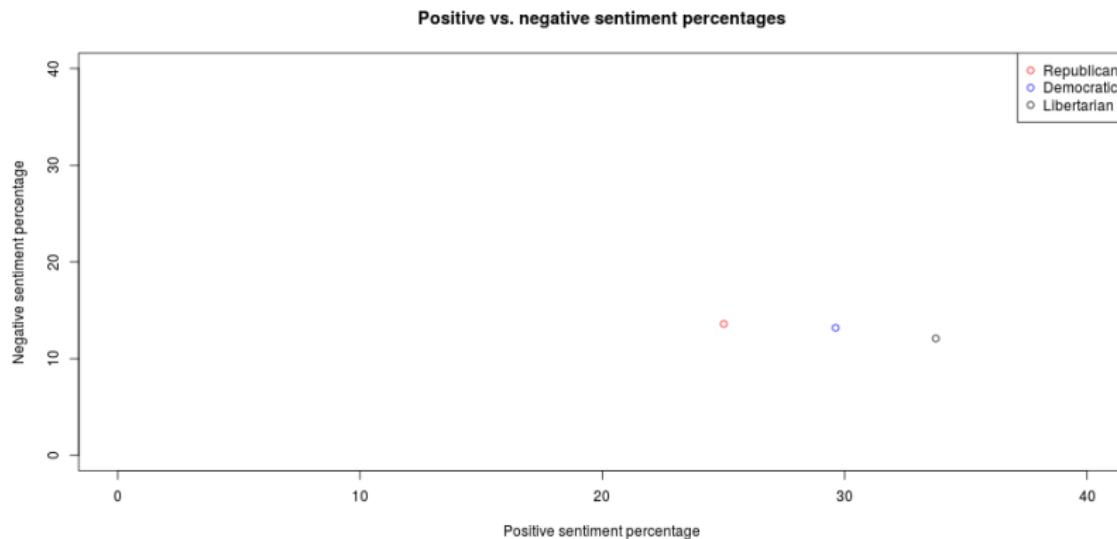
Sentiment analysis was applied to three party platforms.

Each had approximately the level of negativity, but varied on positivity.



Prediction

Same image.



Democratic Party Platform word cloud (100 words)

A word cloud visualization representing the 100 words of the Democratic Party Platform. The words are arranged in a shape where the letters 'democrats' form the main body, and 'will support' forms the outline. The size of each word indicates its frequency or importance within the platform's text. The color of the words varies, with many appearing in shades of green, blue, and black.

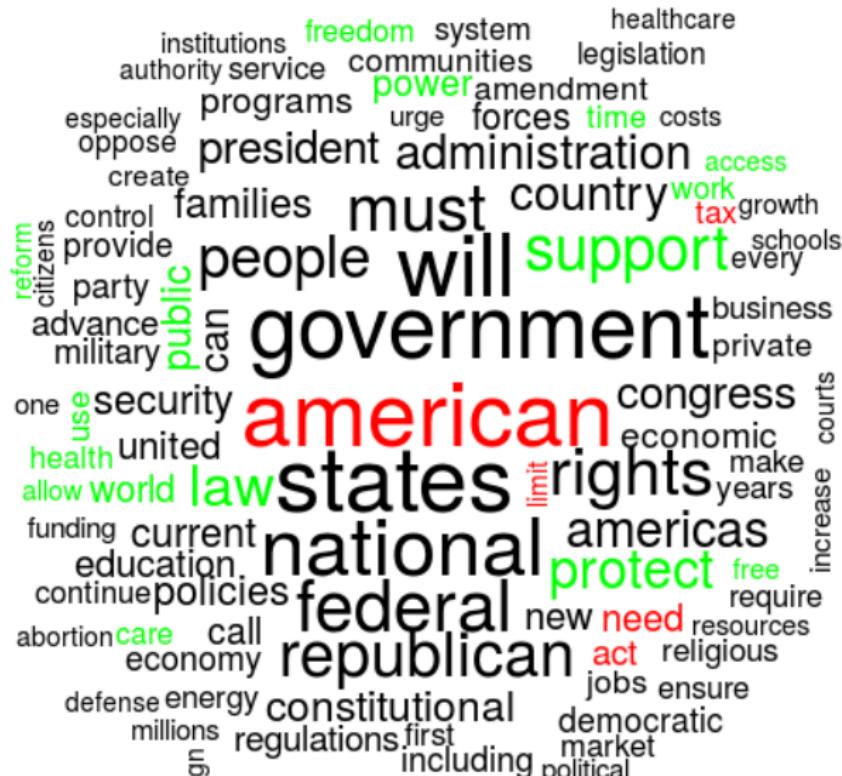
president nuclear
strong strengthen
development corporations
immigration economy help energy
trump living workers investments
global education public continue
children institutions
ensure communities access
address years
voting families united
military end work
students health people
laws believe
jobs rights
protect world
need care make
expand americans services
funding provide must also fight
affordable must also fight
every country states federal
across security new system many
economic including america future
efforts committed increase
opportunities donald climate prevent
enforcement housing disabilities

Prediction

Libertarian Party Platform word cloud (100 words)

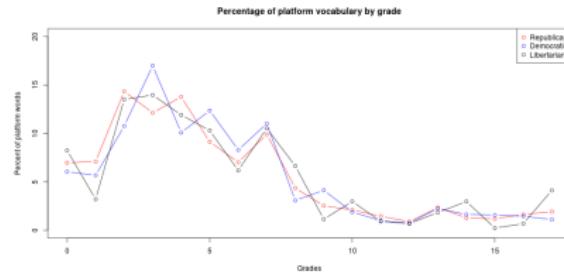


Republican Party Platform word cloud (100 words)



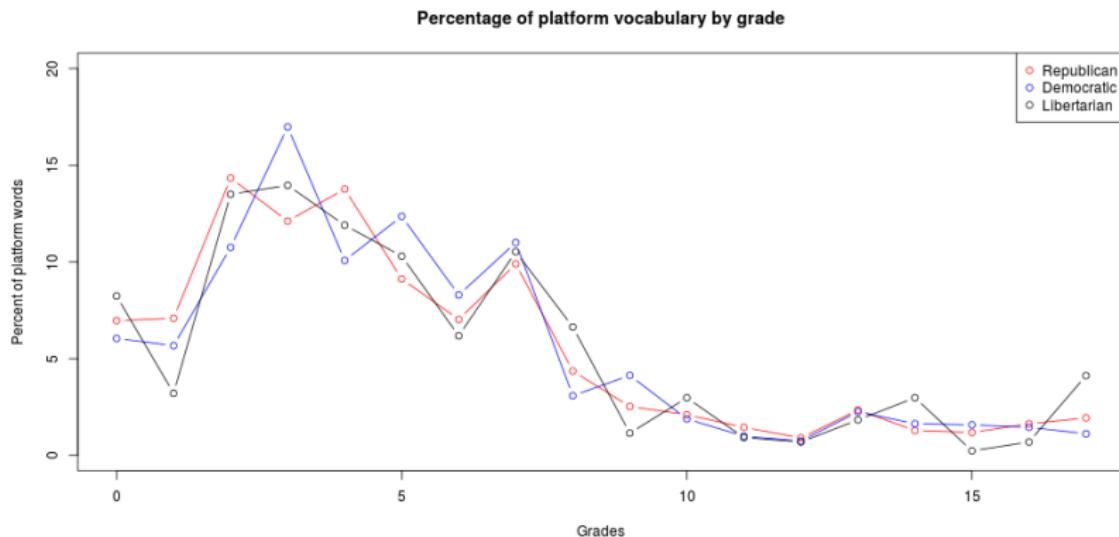
Vocabulary of each platform by grade level

Pro party commentators have made much about how their opponent's platform is “dumbed down” to appeal to that party's core. Looks like they are all about the same.



Prediction

Same image.



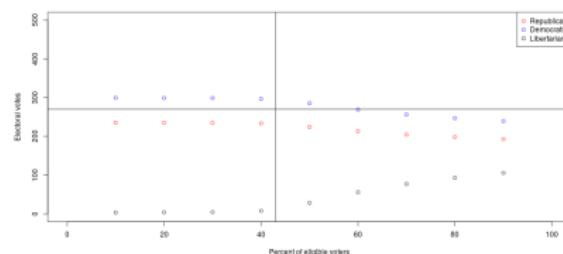
Prediction

How will the electoral college go?

If:

- The core is loyal, and
- The remainder are apportioned IAW the positive sentiment, and
- The electoral college is apportioned IAW the party votes.

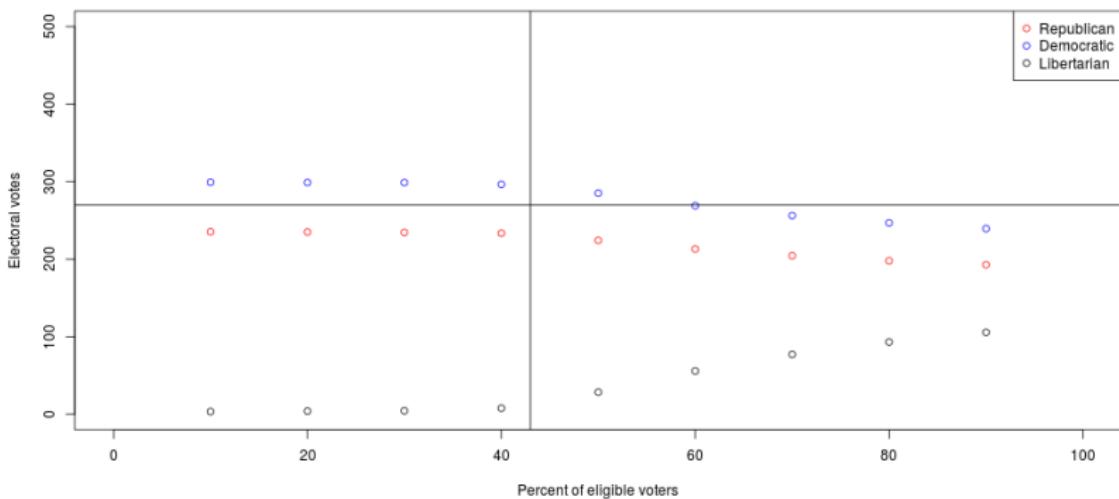
We can vary the percentage of eligible voters and see what happens.



Horizontal to win. Vertical to vote.

Prediction

Same image.



Horizontal to win. Vertical to vote.

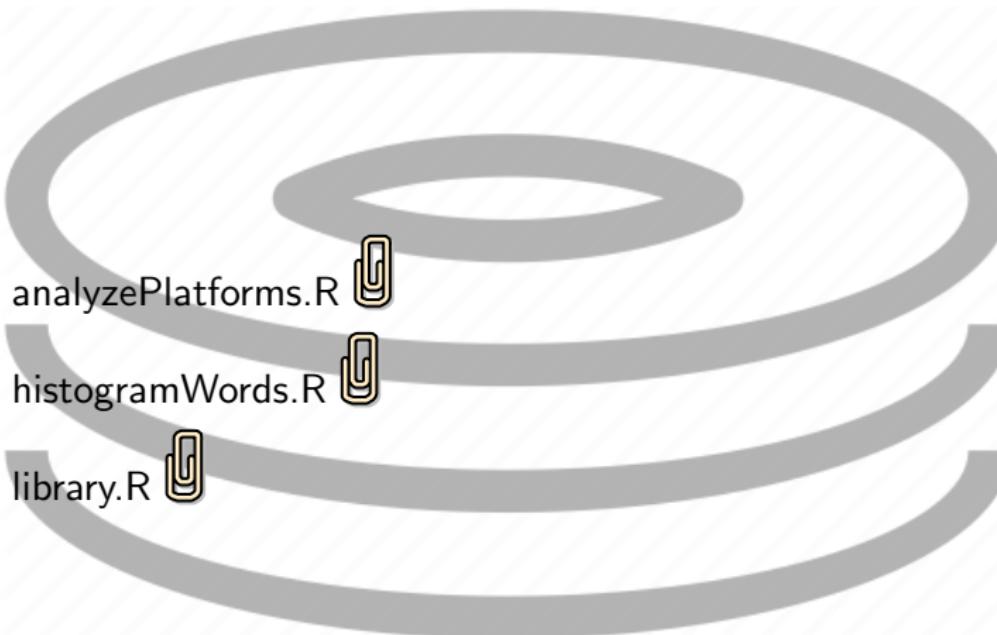
What have we covered?

- Made a simple model to predict Presidential election.
- Model based on readily available and invariant data.
- Model **does not** include any real-time events.
- Model predicts that the Democratic party will win.



Next time: population density over time.

Files of interest

- 
- ① analyzePlatforms.R 
 - ② histogramWords.R 
 - ③ library.R 