



Presidential Debates

NLP Analysis Topic Modelling



Finding the Differences in Debates Styles Democrats vs Republicans

NLP Analysis Topic Modelling

The Data

The Presidential Debate Transcripts 1960-2016

- Every 4 years three debates with 1 Democrat and 1 Republican presidential candidate
- Nixon refused to debate 1964 and 1968
- Also one vice presidential debate each election

The Debates



Natural Language Processing toolkits

- NLTK - word tokenizer, stop words.
 - SpaCy - tagger, parser, lemmatizer.
 - Gensim - LDA Topic Modelling.
 - SKLearn - Vectorizer, TFIDF
 - VADER - Sentiment Analysis
-
- Pretrained Models - ULMFiT, Transformer, ELMo

Stopwords and Lemmatizer in SpaCy

- Remove speech tics and short interjections

spacy

- Anonymize the speakers
 - 'Nixon' filter
 - 'Secretary' filter

Is there a Difference Between Democrats & Republicans?

- Logistic Regression
- KNN
- Random Forest

Accuracy =

0.6805

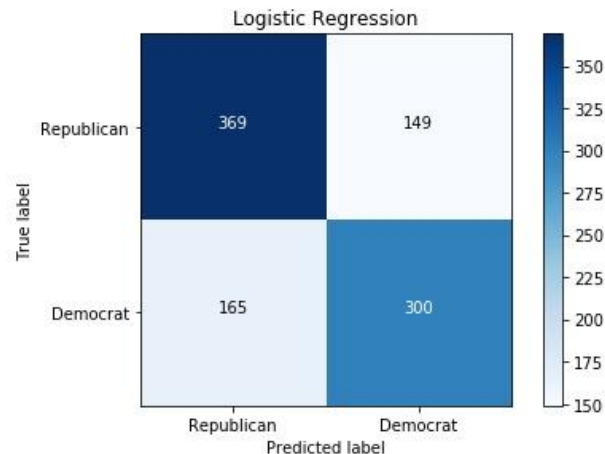
Baseline Accuracy score: 0.6805696846388606

Baseline Scores:

	precision	recall	f1-score	support
0	0.69	0.71	0.70	518
1	0.67	0.65	0.66	465
micro avg	0.68	0.68	0.68	983
macro avg	0.68	0.68	0.68	983
weighted avg	0.68	0.68	0.68	983

Confusion Matrix, without normalization

```
[[369 149]
 [165 300]]
```



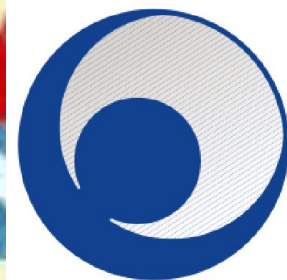
Is there a Difference Between Democrats & Republicans?

- The model shows that there is distinguishing information between the parties even with anonymized speakers

To actually find a difference we'll need

TOPIC MODELLING

Topic Modelling with Gensim

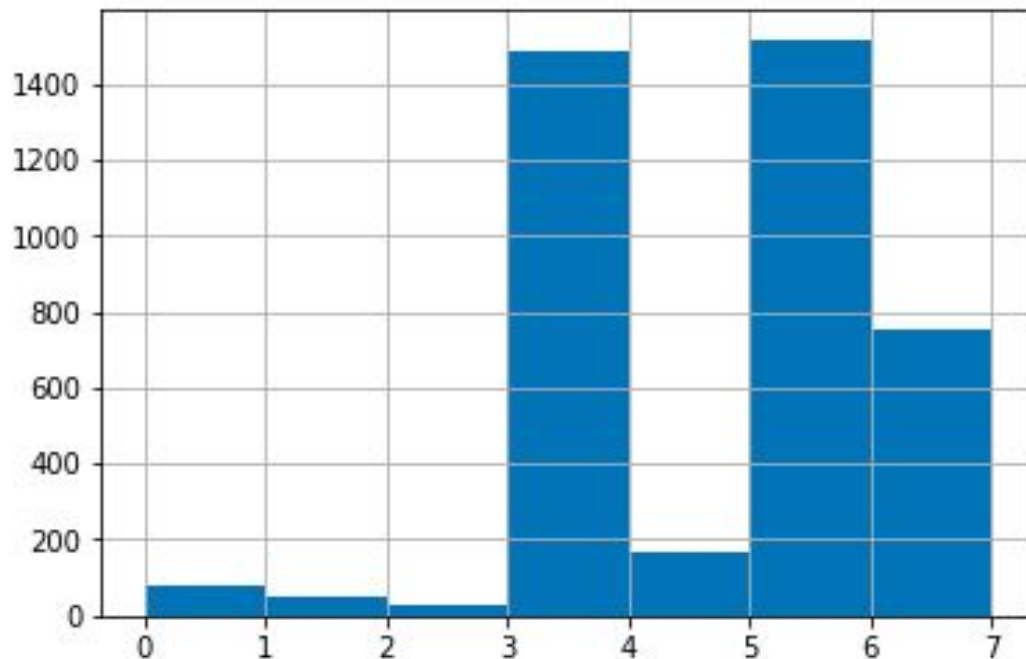


Distribution of Topics

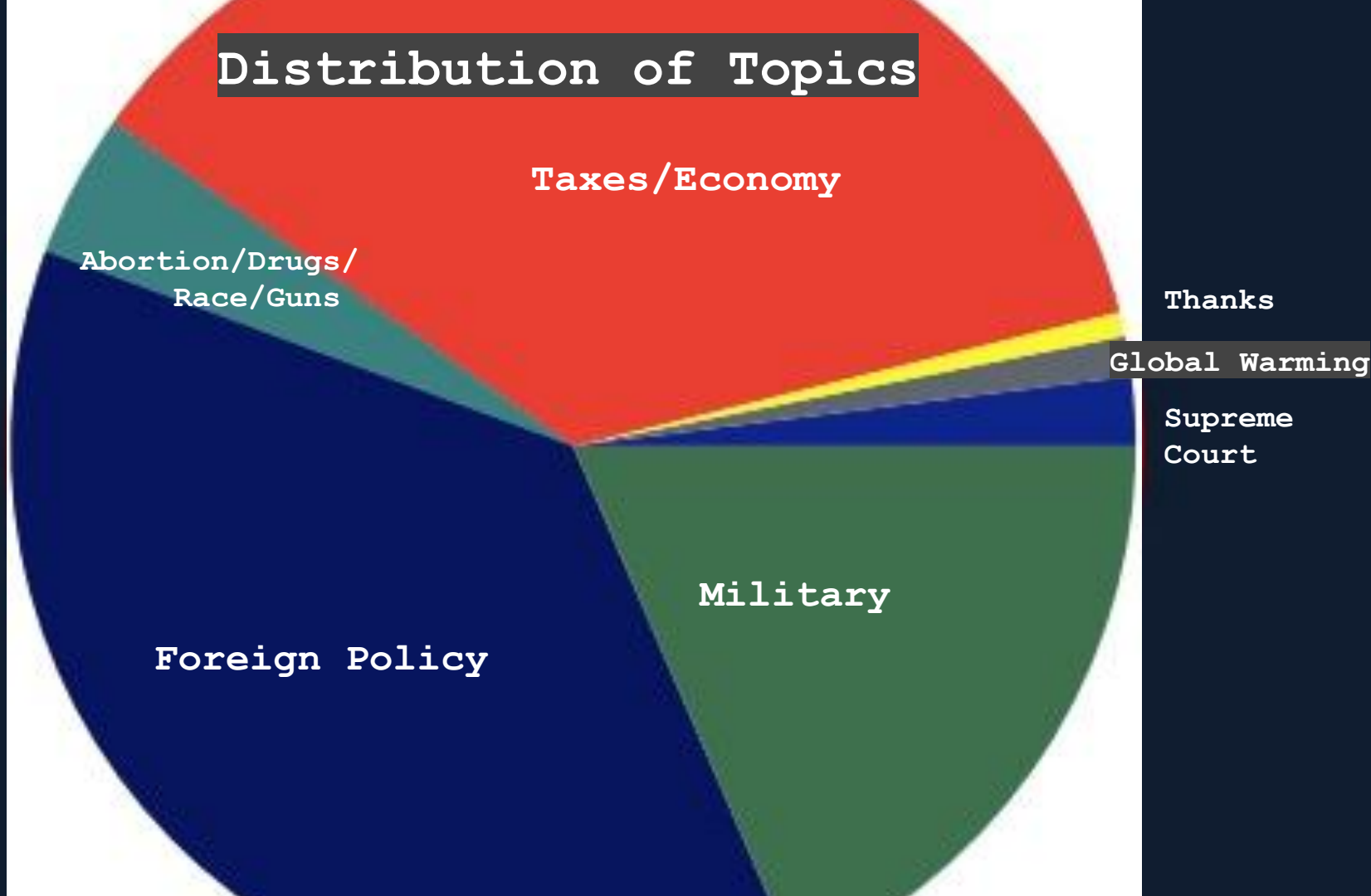
- Added Bigrams

12 Subjects

7 Topics



Distribution of Topics



Taxes/Economy

Abortion/Drugs/
Race/Guns

Thanks

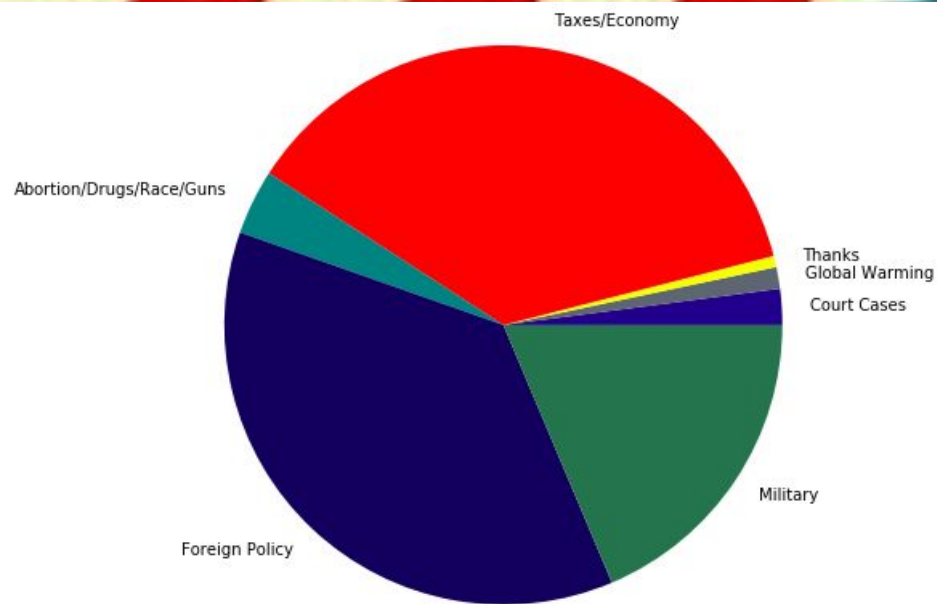
Global Warming

Supreme
Court

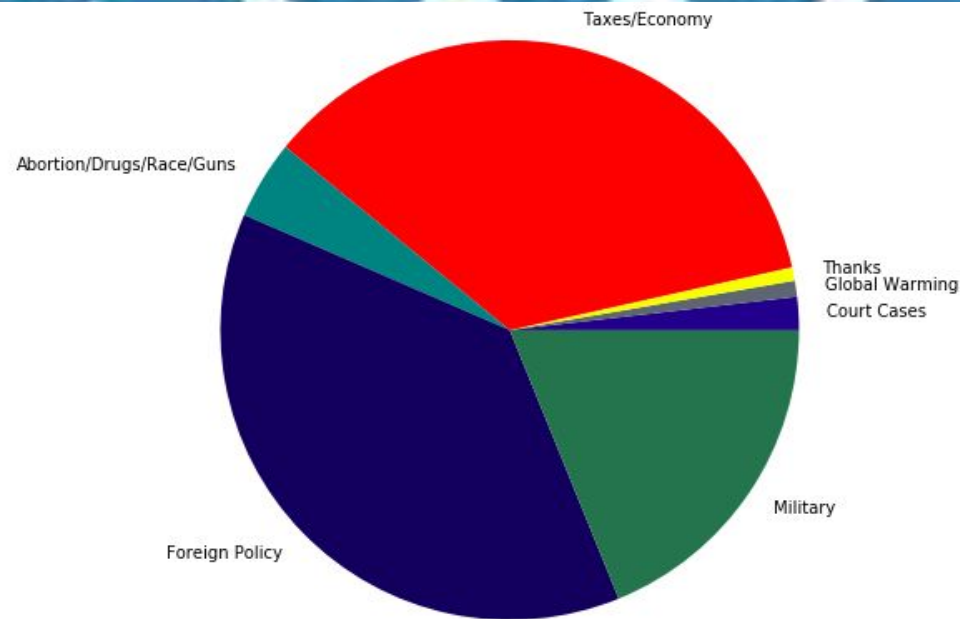
Military

Foreign Policy

Topics by Party



Democrats



Republicans

Democrats had fewer lines across all debates

Percent Difference Democrats/Republicans

-13.525%

['Court Cases: -2.6%',
'Global Warming: 21.1%',
'Thanks: -25.0%',
'Taxes/Economy: -10.4%',
'Abortion/Drugs/Race/Guns: -25.8%',
'Foreign Policy: -15.6%',
'Military: -14.5%']

“Malarkey”

- There are 3 times used “malarkey”
- FOREIGN POLICY
- 2012-10-11-10: “With all due respect, that’s a bunch of malarkey”
- 2012-10-11-46: “It – it is. We Irish call it malarkey.”
- 2012-10-11-48: “We Irish call it malarkey. But last thing. The secretary of defense has made it absolutely clear, we didn’t walk anything back. We will not allow the Iranians to get a nuclear weapon...”

Eric Ma



Data Scientist/Analyst

-Flatiron School-

Get in touch with me

Email: eric.ma206@gmail.com

Github: <https://github.com/EricMa206/DebatesTopicModeling>