

SPRINGBOARD DATA SCIENCE INTENSIVE

Capstone Project: Analysis of the 2016 US Presidential Debates and
Associated Social Media

The Questions

- What features of political language differ from normal language?
- How did social media influence the election?
- What aspects of Donald Trump's language and strategy allowed him to win?
- Was aspects of of the language used by all candidates is interesting or surprising?

Most Useful Python Libraries

- NLTK
- Textstat
- TextBlob

Grade Comprehension Level

- The overall grade comprehension level of political speech is much lower than would be expected from a national leader, assuming that leaders are selected for demonstrated raw intelligence.
- It appears that political speech fulfills a completely separate role from science and engineering in terms of the demonstration of high value.
- Clinton's speech comprehension level is much higher than that of Trump in relative terms, but in absolute terms both are very low.
- Democrats scored higher than Republicans, but, of the Republicans analysed, Ben Carson scored much higher than his other rival Republicans.

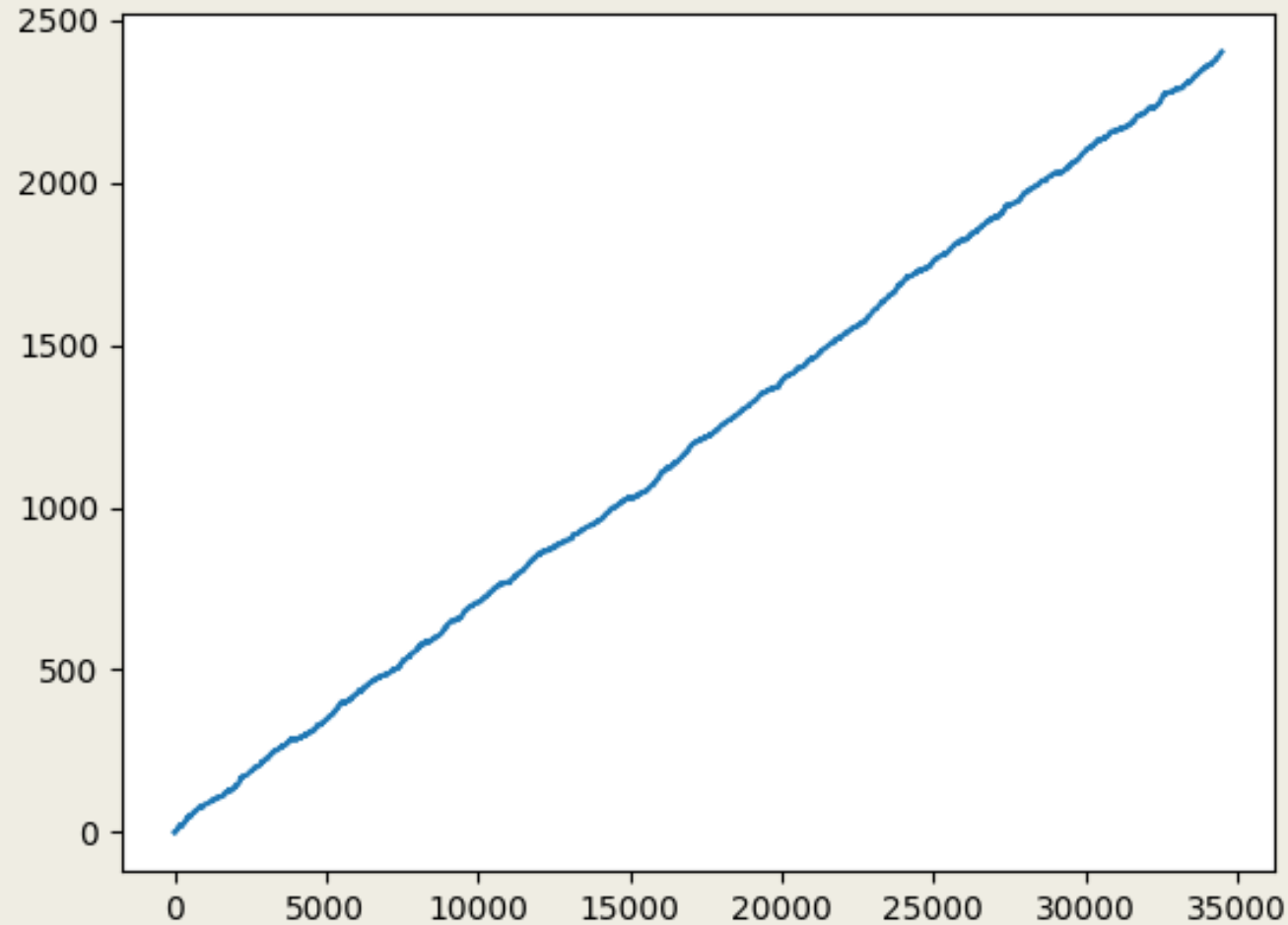
Sentiment

- The Cumulative sentiment and distribution of sentiment, per sentence, was extremely uniform.
- This demonstrates, perhaps unsurprisingly, that politically speech is characterised by extreme preselection selection bias.
- Despite Donald Trump's reputation for causing antagonism, outrage and negativity, Trump's sentiment profile is more 'normal' than many of his rivals.
- Trump's overall sentiment levels were measurably and appreciably lower than those of Former Secretary Clinton, but not by much in on the absolute scale (1.5%).

PRESIDENTIAL PRIMARY SENTIMENT

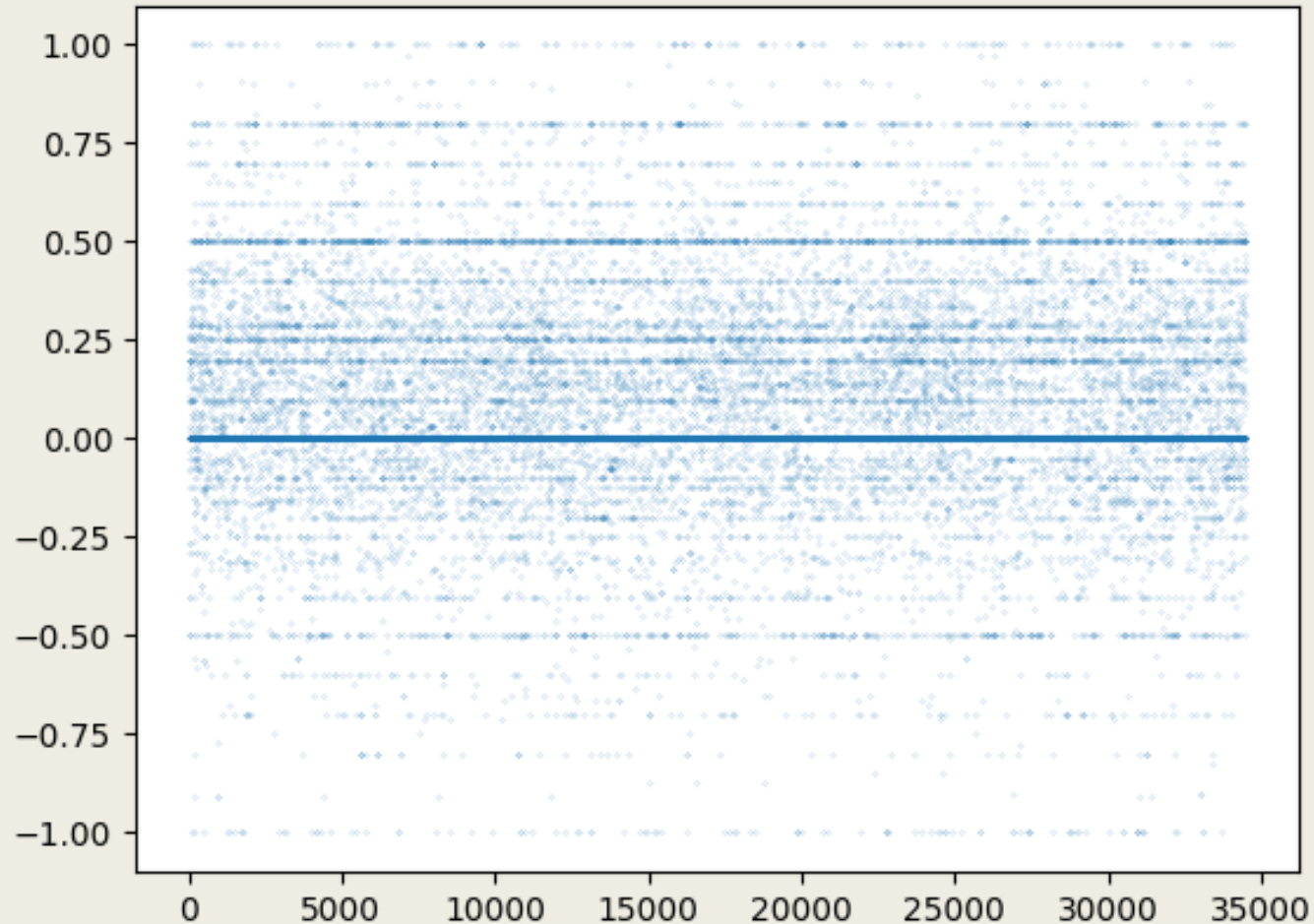
High Preselection

CUMULATIVE SENTIMENT



- $R^2 = 0.99+$
- Slope = 0.694
- Almost a perfect straight line!

CUMULATIVE SENTIMENT



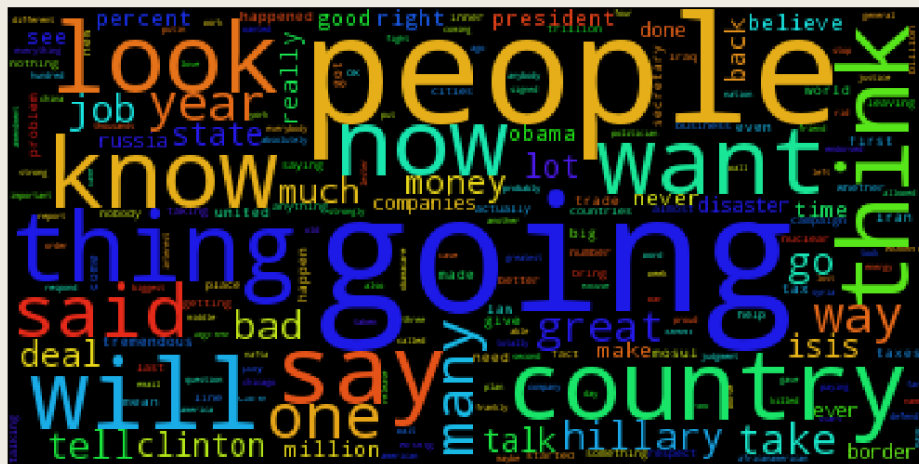
- StdDev = 0.24
- Four standard deviations either side of the mean to capture 99.9% almost perfectly

Trump vs Clinton

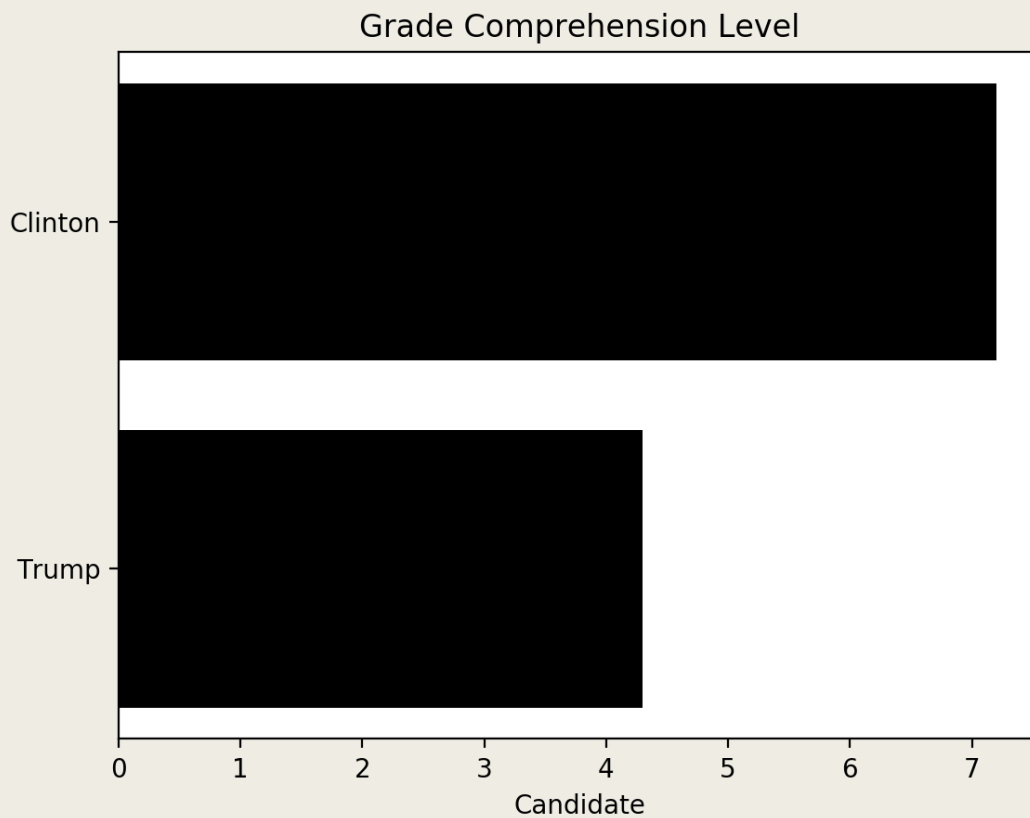
- Trump was able to use Twitter much more effectively than Clinton was and was able to win valuable contentious ground there without going to the risk of inflaming sentiment or alienating the audience on live television.
- Trump was able to infer meaning and leave more of his speech to the interpretation of the audience, which seems to be a major feature of selection in political speech.
- Trump was better able to create two angles of attack between social media and the debates and was therefore able to take advantage of more dynamic posture with more potential lines of attack, similar to a chess player, boxer or martial artist.
- Clinton, by comparison, appears to have been caught by surprise and played too rigid a game.
- Trump appeared crazy, but within the context of the rules of the game he was crazy like a fox.

THE HEAD TO HEAD TELEVISED DEBATES

Trump vs Clinton



TRUMP



FLESCH-
KINCAID

The image features the word "twitter" in its signature blue, lowercase, rounded font. The text is centered within a white, stylized frame that resembles a corner bracket or a partial border. The frame consists of a vertical line on the left, a horizontal line at the top, and a horizontal line at the bottom, with the right side being an open L-shape. The entire composition is set against a solid black background.

twitter

@realDonaldTrump vs @HillaryClinton

- Text analysis between the debates and Twitter would suggest that Twitter played in major role in the election outcome.
- Trump was able to dissociate his social media persona from his professional persona and able to go use Twitter to go outwith the bounds of the highly constrained 'normal' language of the political debates.
- Speech in debates was not predictive of speech on Twitter. Both candidates appear to have adopted alternate social media personas.

@REALD
ONALTR
UMP

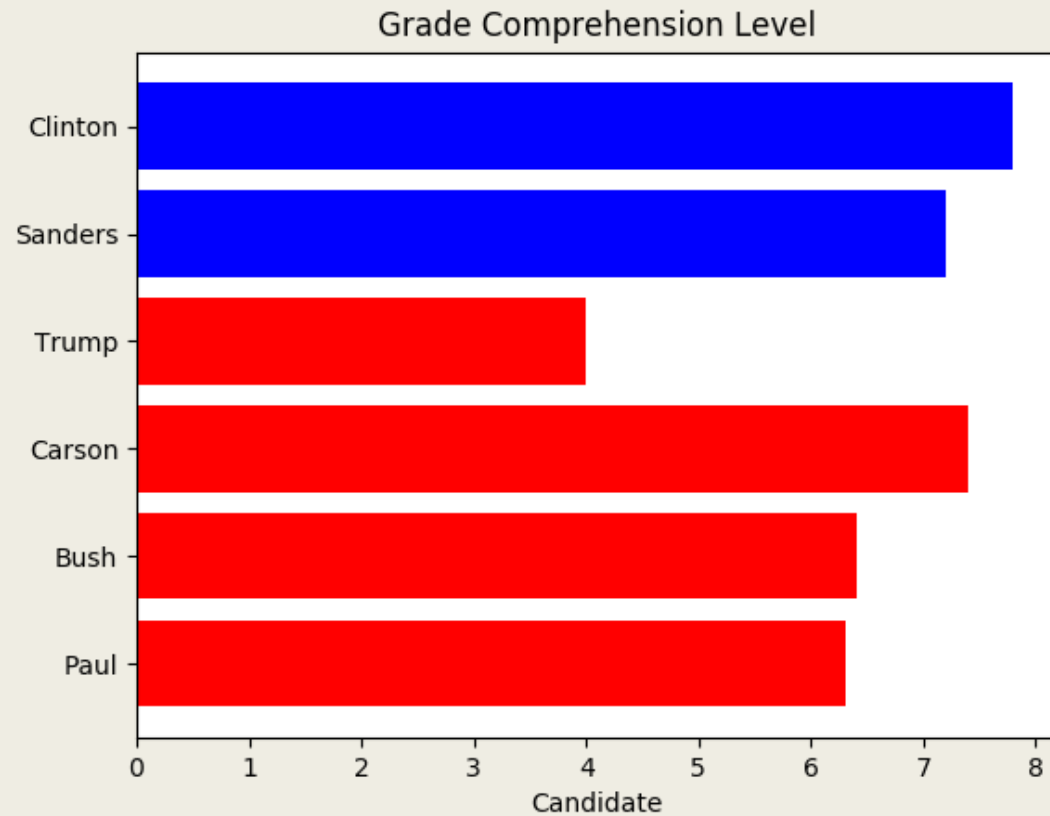


@HILLARY
CLINTON

PRIMARY DEBATES

Candidate Word Clouds

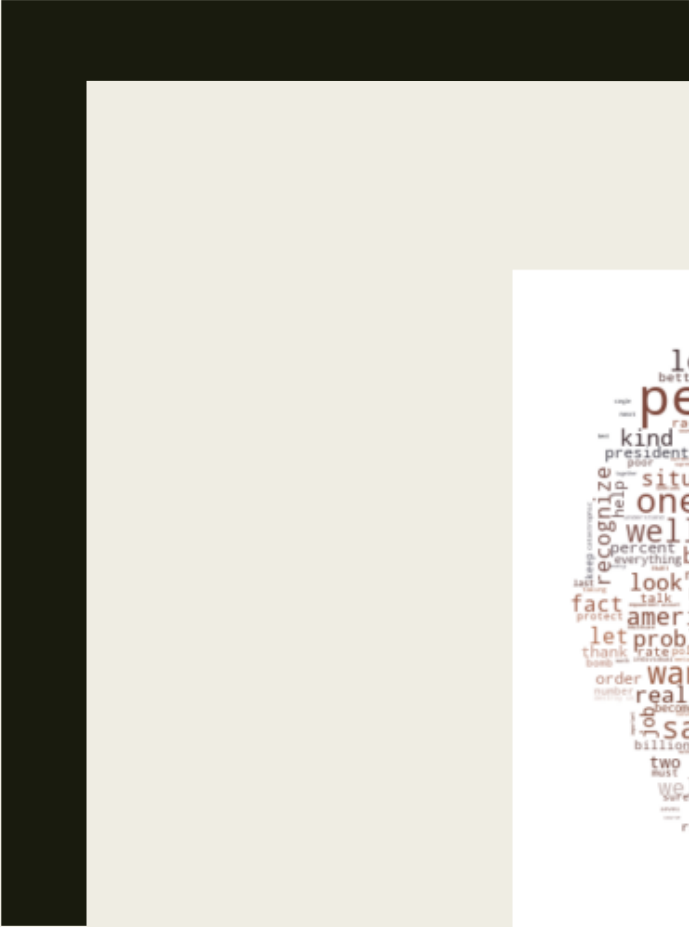




FLESCH-
KINCAID

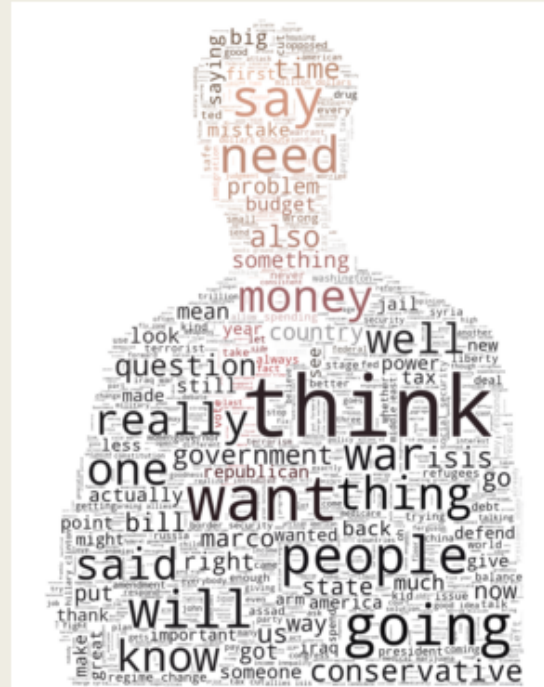
HILLARY CLINTON

DONALD TRUMP



coming
 give
 tax
 now
 big
 right
 now
 way
 us
 anybody
 care
 place
 cost
 even
 got
 made
 come
 system
 fair
 go
 year
 pay
 good
 able
 start
 money
 make
 put
 think
 fact
 matter
 everybody
 great
 done
 need
 thing
 baby
 concerned
 can
 little
 away
 also
 over
 see
 year
 pay
 good
 able
 start
 money
 make
 put
 think
 fact
 matter
 everybody
 great
 done
 need
 thing
 baby

[illegible]



RAND
PAUL

Recommendations and Further Work

- I would recommend to all readers not to get hung up on political speech. Political speech appears to be a last-resort means of creating scapegoats for our individual frustrations And inability to produce enough value to solve our own problems. Political speech is very uniformly low comprehension, much like a weapon and performs a war-like function.
- Social media will continue to play an increasingly important part in the political process, due to its ability to create more dynamic means of political attack and defence.
- Part of speech processing
- Many aspects require too much contact to gain insight from NLP and we should only use these tools as a means to shed greater light on events and test hypotheses. NLP is not a replacement for human intuition and contextual understanding, it rather adds a great deal of value to pre-existing intuition and contextual understanding and is a great way to gain further insight through dialogue. NLP is not a Philosopher's Stone and should not be used as such.
- Caveat Emptor! When making decisions with NLP, look at the data yourself and start a dialogue. Treat the recommendations like you're buying gold and give them a good scratch!