Let us not seek the Republican answer or the Democratic answer, but the right answer. Let us not seek to fix the blame for the past. Let us accept our own responsibility for the future.

Senator John F. Kennedy

--Speech at Loyola College Alumni Banquet, Baltimore, Maryland, February 18, 1958.

WebScraping. Staggering Acts of Genius were performed, leveraging BeautifulSoup and MongoDB, to effectively scrape and store speeches/debates the Miller Center at the University of Virginia and the Commission on Presidential Debates.

IndicoIO. An awesome text analyzer available for use through an API. Using our Mongo database and some easy Python code, we fed speeches and debates to add a variety of indicators to our database including:

* **Sentiment:** Rates text as positive or negative (0-1 scale)
* **Personality:**  Predicts a personality for the speech writer on four criteria (extraversion, openness, agreeable, conscientiousness)
* **Emotion:** rates the text for 5 emotions (anger, fear, joy, sadness, surprise) to the probability that the author is expressing the respective emotion.
* **Personas:** Predicts the Myers Briggs persona of an author based on a sample of text.
* **Political:** Gauges the political leanings of a phrase or document (conservative, liberal, green, libertarian) For our purposes, we grouped conservative and libertarian as “Conservative” and liberal and green as “Liberal”

**Word Cloud** Word clouds are effective messages to showcase key topics or words in a body of text.  We wanted to show what were the most discussed topics throughout American history. We analyzed presidential speeches and election debates with a word vectorizer to weight how often certain key words were said per decade.

**Stop Words** We leveraged the SKLearn Count Vectorizer to generate the word-cloud. However, the stop-words included in this package are well-known to be rather slim. We augmented the stop-words list with additional words to avoid skewing our word cloud with conjunctions or unnecessary adverbs.

Stacked Area Chart:  
Indico predicts the emotion expressed by an author in a sample of text. The expressed emotions registered by the library are Anger, Fear, Joy, Sadness, Surprise. The stacked area chart shows an aggregate of speeches delivered over a decade. Scale has been increased by a factor of 100.

Stack Bar Chart:

[9:06 PM](https://gtatl201902data3.slack.com/archives/GLEFRMQJV/p1565139998004400)

Around 1000 speeches and debates were rated from 1 - 0 using IndicoIo based on its political leaning. This stacked bar chart illustrates the average per decade of both sides.