SHORT PAPER #4: "Trump vs. Obama"

✓ Published   ⋮

What's exciting about this challenge: In 2017, Donald Trump broke the wordcloud by using much the same language as Obama had, but for very different purposes. The result was that the wordclouds for each politician looked nearly the same. Newspapers that published wordclouds about Trump felt the need to editorialize extensively to explain to readers why the wordclouds looked the way they did. Newspapers across America have been reluctant to use text mining as a tool for talking about politics. You will help them find a way out of this dilemma.

The challenge in a nutshell: a simple word cloud of Trump's official speeches and Obama's official speeches makes it look like the presidents share most of the same values. Clearly, the wordcloud fails the purpose of political analysis here. You will use other strategies to compare datasets of the presidents' use of text to compare the norms, beliefs, and ideals of the two

What's challenging about this challenge: The most interesting results are probably hidden at the level of grammar, that is, the presidents' statements about who did what to whom. What does Trump say the American people are, and what should they do? Ask the same for Obama. This kind of analysis is most sure-footed when using parts-of-speech analysis to compare subject-verb-object structures. It requires careful attention and manipulation of the code with attention to linguistic forms, as well as careful reading.

Where to get the data: https://www.kaggle.com/datasets/littleotter/united-states-presidential-speeches ⤷ (https://www.kaggle.com/datasets/littleotter/united-states-presidential-speeches)

Tools to use: ngrams, word embeddings, parts-of-speech analysis. Previous weeks' notebooks give you the tools to do this kind of work. Making KWIC tables for each keyword-president or noun-verb-president or keyword-collocate-president combination.

What to turn in: 1-2 visualizations plus an essay of 2-3 pages with footnotes using quotations to support your interpretation of the visualizations and showing the source of each quotation.

Length and format: 4-6 pages, double-spaced, excluding visualizations. Embed visualizations in word document and label their axes, providing a title like "Figure 1." In the paper's introduction, detailed prose should explain what Figure 1 is and describe *all* of the data in Figure 1. In the

paper's body, detailed descriptions should both quote *and* summarize each appearance of relevant keywords in all of the presidents' speeches. All quotations should be cited in Turabian format, giving the location in the original primary sources of the quotation. Analysis of the primary sources should comprise a series of well-structured paragraphs whose introduction and conclusion sentences draw insight from the speeches. The paper's conclusion should include observations about the newspaper's contribution to American politics, including how the newspaper's orientation changed over time. Pages should be numbered.

Points	10
Submitting	a file upload
File Types	doc and pdf

Due	For	Available from	Until
Apr 29	Everyone	-	-