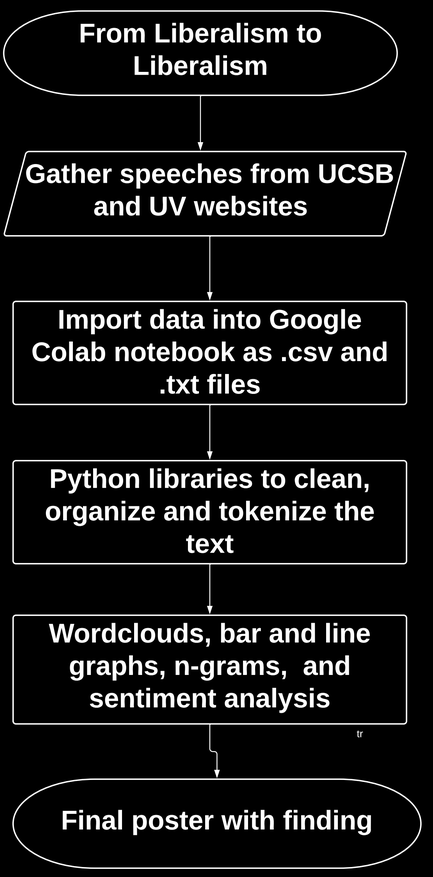
A Textual Analysis of the Great Depression:

DH100: Theory and Method in the Digital Humanities | Dr. Adam Anderson | Juliana Torres | Summer 2021



From Liberalism to the New Deal in words:

When FDR became president in 1933, he encountered a nation that had nosedived into despair. The Great Depression was entering its fourth year, and the unemployment rate was at 24.9%, the highest in the country's history. Around 26,000 businesses failed. Stock prices plummeted. It transformed American life.

His predecessor Herbert Hoover was a classic liberal and believed that economic downturns were part of a capitalist economy and were to be expected. As he said in a speech in 1931, government assistance weakened the American character. Hoover hoped that the solution to the Depression would come from voluntary aid and "associational action" by community members and business people and by maintaining a positive mind frame, often making public statements that "the tide had turned." In the eyes of many Americans, President Hoover's response seemed inadequate and uncaring. His nineteen-century approach to Liberalism no longer served the nation's needs, and FDR's New Deal marked the beginning of a new relationship between Americans and their government.

FDR openly opposed that sentiment and invited the country to establish a new relationship with its federal government. During Franklin Delano Roosevelt's 12-year-long era as president, he gave 27 fireside chats to address a nation recovering from an depression and walking into a world war. Contrary to Hoover, FDR hoped to more government intervention and a change in expectations from the people and the presidency.

My project seeks to find linguisttic evidence of this sentimental shift between Hoover's Liberalism and FDR's New Deal. By analyzing the most frequently used words, sentiment, and topics of the speeches, I hope to find what made one so comforting while the other seemed so out of touch.

Wordcloud for the combined speeches

What is my data?

My data consist of the 27 full-text speeches given by FDR from 1933 until 1939, the end of the New Deal.

Additionally, I will analyze 16 speeches giving by Hoover from 1929, the first year of the Depression, and 1933, his last year as president. These speeches have been made available online by UCSB's American Presidency Project and by the University of Virginia.

The data will be organized by date in order to provide a chronological analyzes of tone and themes.

What is my question?

Can we find evidence of the shift from classical Liberalism to the New Deal in language?

To answer these questions, I will analyze:

1) What were the most used words on their speeches?
2) How did they created sentences and who was the focus of their messages? What are the most repeated expressions?
3) What is the general sentiment of their language?

I. tools and methods

A Textual Analysis of the great depression:

Data Preparation:

My first steo was to gather the data from the UC Santa Barbara and University of Virginia websites. Since I was dealing with a reasonable amount of data, I manually copied and pasted the speeches for both presidents. I created a .csv file for each president containing all of their speeches organized by data.

I also created a master .txt file with all of the speeches for FDR and another one with all of the speeches for Hoover.

I then used Python libraries and functions to clean, organize and tokenize and standardize the speeches:

```
##basic libraries
import numpy as np
import pandas as pd
from os import path
import re
from collections import Counter

##libraries to clean and organize my text data
import unicodedata
from string import punctuation
import nltk
from nltk.tokenize import word_tokenize
from nltk.tokenize import sent_tokenize
from nltk.corpus import stopwords
from nltk.corpus import words
from nltk.stem.snowball import SnowballStemmer
nltk.download("stopwords")
nltk.download("averaged_perceptron_tagger")
```

Data Analysis and Visualization:

I plan to present the data in steps. First the data in a Pandas dataset to give an idea of basic data organization. Then I will use the following libraries to analyze and present the data:

```
##visualization library
import matplotlib.pyplot as plt
%matplotlib inline
plt.style.use('seaborn-bright')
from PIL import Image
from wordcloud import WordCloud, STOPWORDS
from nltk.util import ngrams, bigrams, trigrams
import spacy
import seaborn as sns
```

1. Text Frequency Analysis:

I will use Matplot Lib and Wordcloud to create visualization of the most frequently used words on their speeches.

2. N-Grams:

Next I will do a closer analysis of word choice by using n-grams to identify patterns in phrase and expression formation. For example: maybe above the wordcloud shows that both presidents used the word "Government" several times. I want to analyze what was the content of those uses. What words were associated with the word "government" by Hoover and which were associated by FDR.

3. Topic Modelling

I will use Gensim pre-treined model to identify themes within FDR and Hoover's speeches.

4. Sentiment Analysis:

Finally I will perform sentiment analysis to try to identify the overall feeling of those statements and present it using matplot lib.

II. Data Visualization

A Textual Analysis of the great depression:

Anticipation:

Because of the nature of the text, I anticipated that the results of the wordcloud would be similar, with FDR's most frequent words being related with the Federal Government, and Hoover's speeches focusing on banks and . I also anticipated that Hoover's rhetoric and vocabulary would be more technical and directed at institutions, rather than the American people, while FDR would have a simpler and more accessible vocabulary and would often aim his message directly to the people.

Results:

The texts proved to be more similar than I predicted, which might have several different reasons unrelated with their political discrepancies and more caused by their similar educational background, social class, and occupation.

Both presidents turned their focus towards the state. The most used word by both of them was "Government" with "Country" and "People" also appearing upon the most used words.

This similarity might show more than it appears. We see in this preliminary textual analysis that the focus of both men will be in acknowledging the role of the state in managing the crisis. Even for politicians like Hoover, who believed in a small government, the role of the state was undeniable.

With these results in mind, I moved towards N-Grams to identify the most common expressions used by the presidents. This way I hope to take a closer look on the types of repetitions and frequent phrases the presidents used to make their point.

.

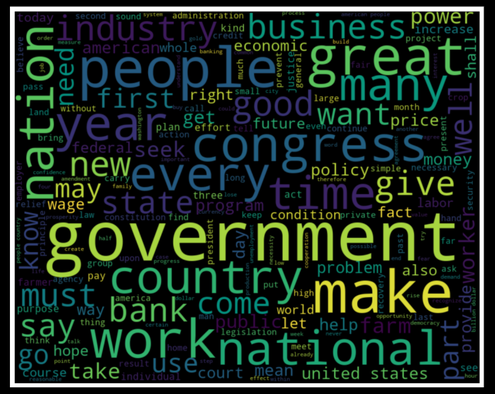


Fig. 1. FDR full text wordcloud.

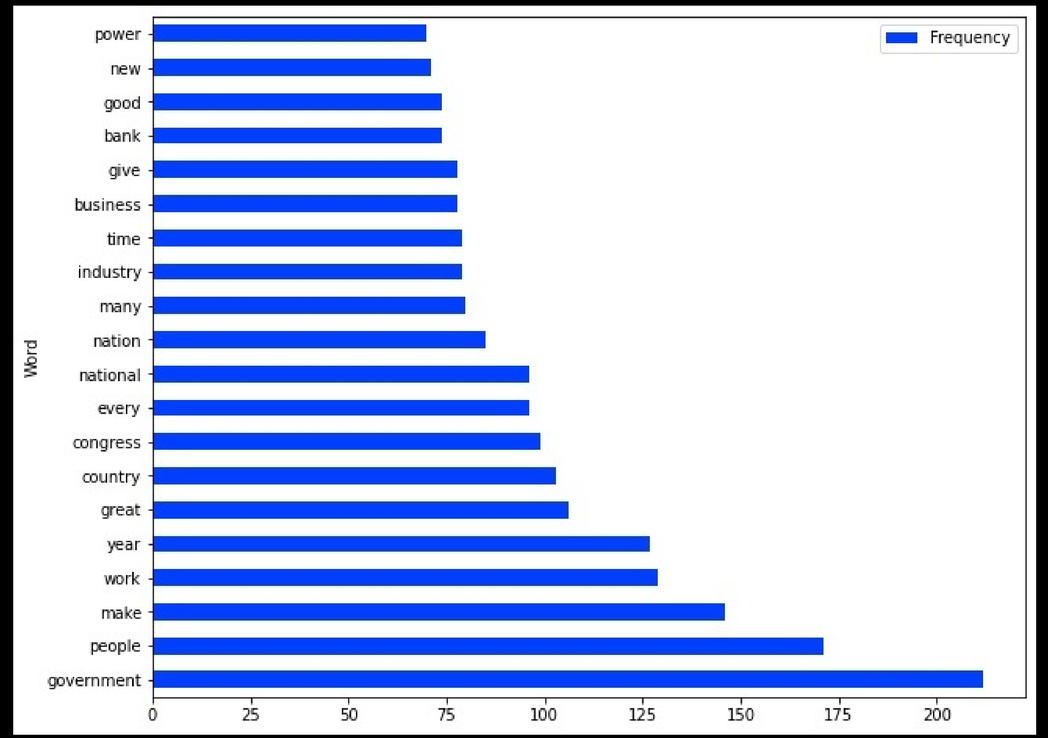


Fig. 2. FDR word frequency barplot

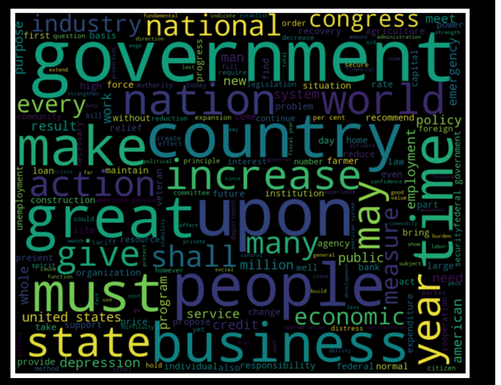


Fig. 3. Hoover full text wordcloud

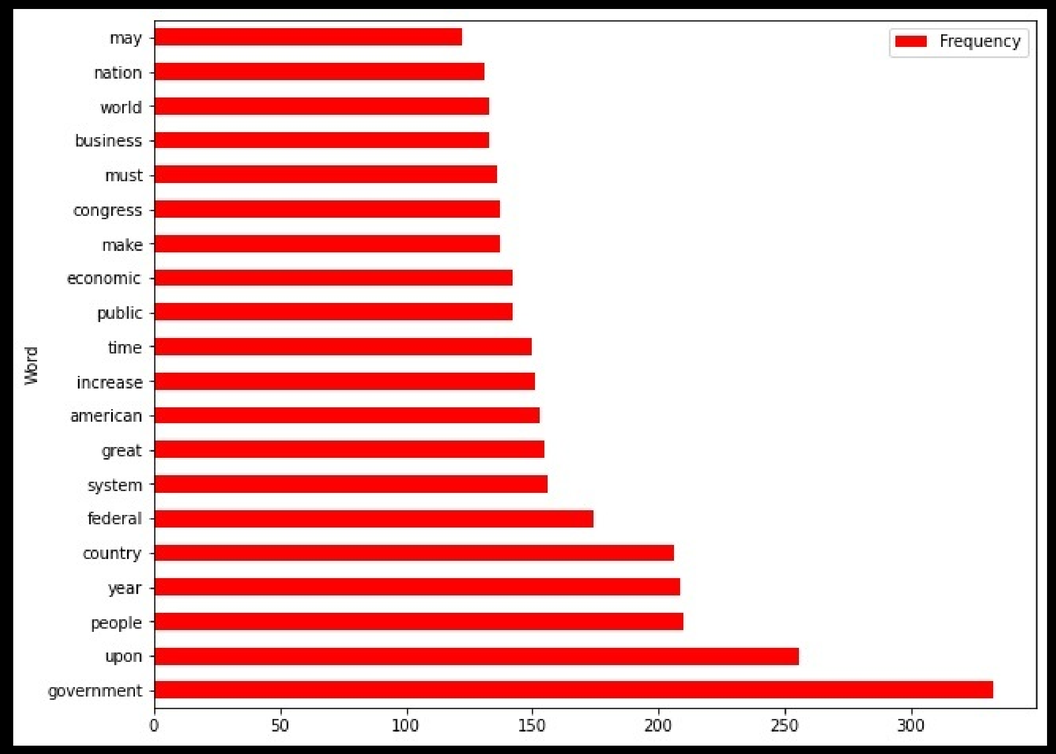


Fig. 4. Hoover frequency barplot

II. Data Visualization

A Textual Analysis of the Great Depression:

Bigrams

Figures 5 and 6 show the bigrams for FDR and Hoover. The results matched my preduction that the vocabulary would be very similar, with more personal enfasis for FDRs texts.

The bigrams for FDR show a more direct communication with offers of solution to the problems. We can see that in the constant mention of New Deal projects like the "Relief Roll" and "industrial recovery."

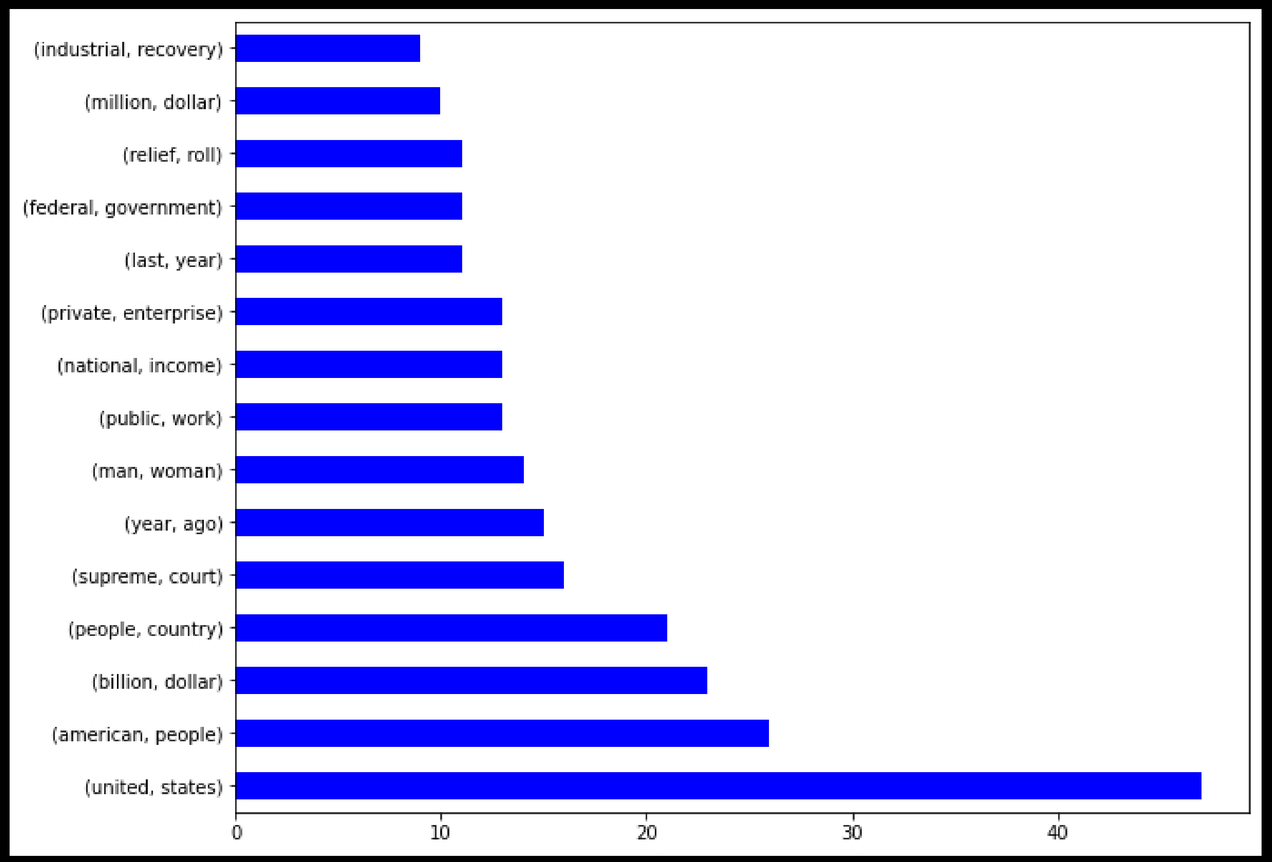


Fig. 5. FDR Bigrams

Hoover's speeches also have warm and direct language like "man and woman" and "american people," the speech seems to focus on technical and specific details that impact the lives of big business and those living in rural areas. Mentions like "fiscal year," "foreign country," and "tariff commission" show a president speaking a language that the masses could not relate to.

This language suggests that part of what made FDR speeches so effective was his ability to go straight to the point that affected the American people. Further analysis into trigrams might cristalize this idea.

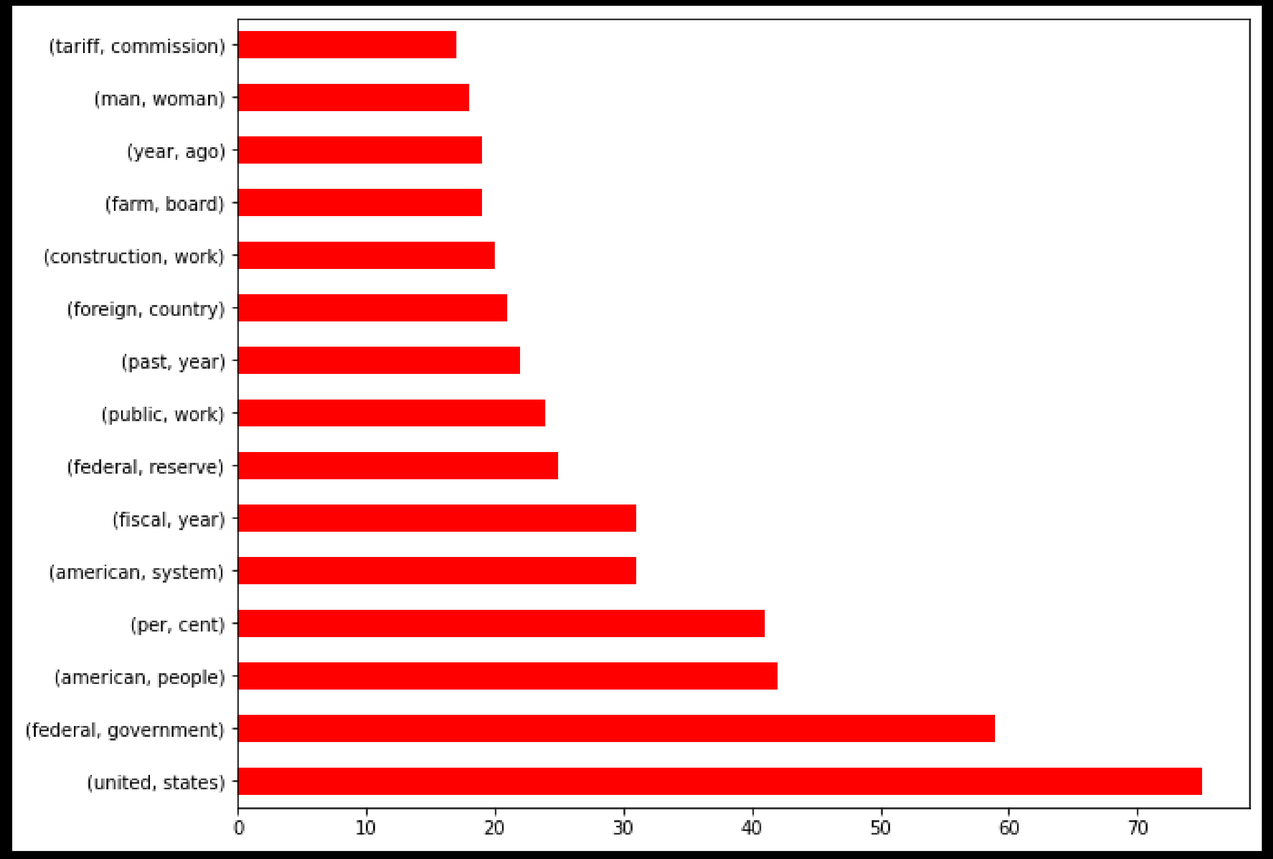


Fig. 6. Hoover Bigrams

A Textual Analysis of the great depression:

Trigrams

The same can be seen on these Trigrams.

FDR employs a language that speaks directly to the public and continues to mention govermental programs that are addressing the Depression. In fact, he mentions himself (or his position) three times.

Another interesting find is FDR's four mentions of "demand restoration of confidence" which illustrates his campaign to recreate the American financial system in a way that common people could trust and depend on. FDR also has other interesting solutions like "anti trust law:" and the interesting appearance of the Social Security creation with the mentioning of "reach age seventy."

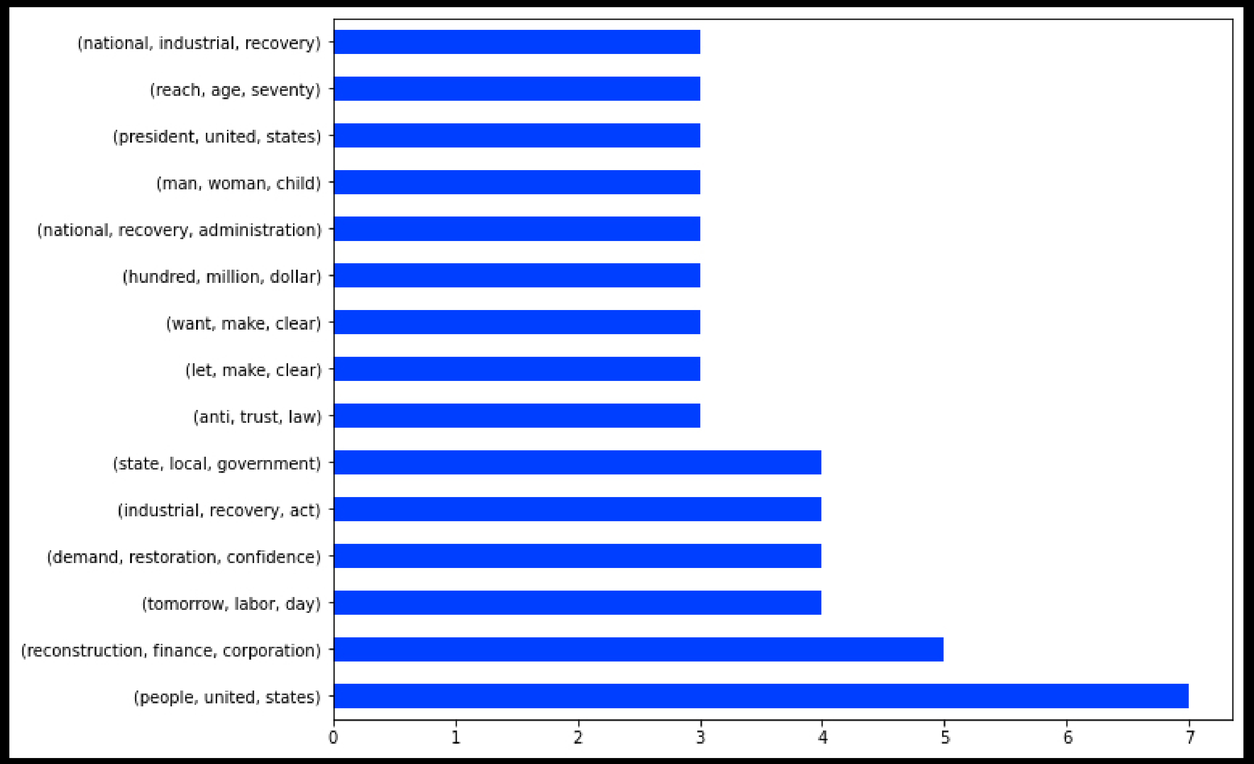


Fig. 6. FDR Trigrams

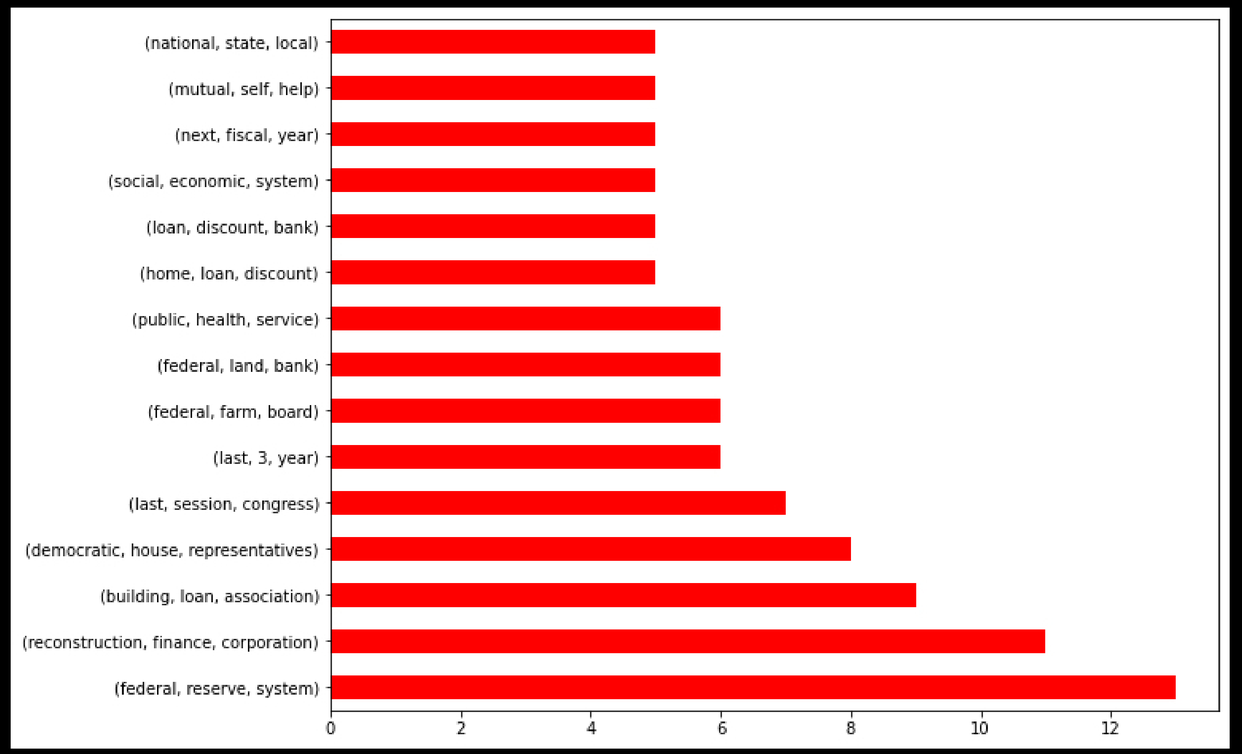


Fig. 7. Hoover Trigrams

II. Data Visualization

A Textual Analysis of the Great Depression:

Sentiment Analysis

I used NLTKs VADER to perform a simple sentiment analysis for the complete body of their speeches. I expected Hoover to have more negative to neutral tone and FDR to have positive tone across the board, as many of his Fireside Chats had an optimistic tinge.

The results showed yet another similarity among the presidents. In line with my expectations, FDR scored positively on all of his speeches, while Hoover had only one negative tone in a speech less than a month after the crash of the market.

Again the results might be misleading and a deeper analysis of the text, with close reading methods, might distill the nuance between these texts.

NOTE:The model is fitted differently for each set of speeches so the Y axis have different lower limits. The V-shaped drop in Hoover's graph represents a negative score.

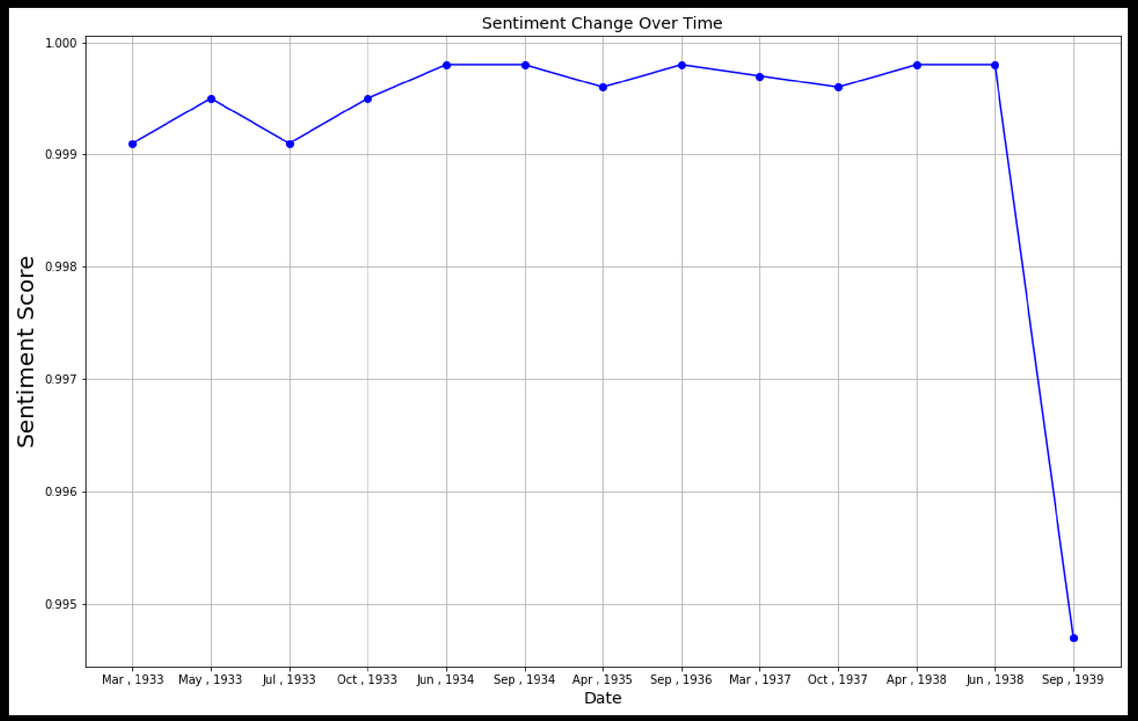


Fig. 8. FDR Sentiment Analysis

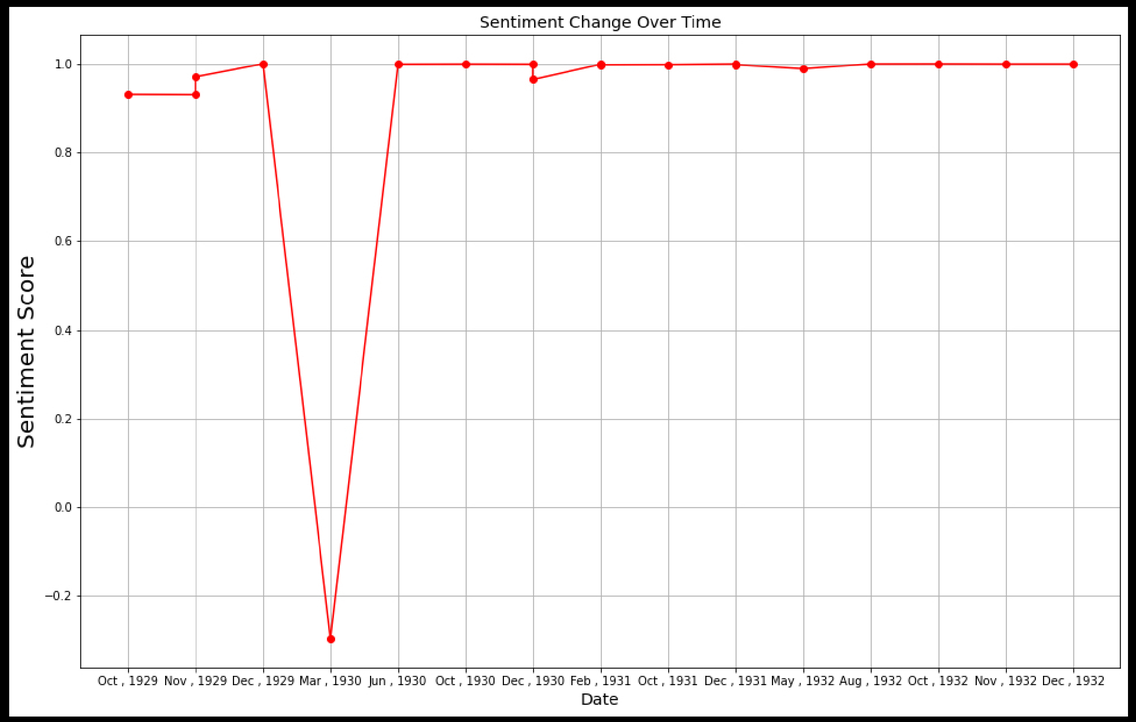


Fig. 9. Hoover Sentiment Analysis

A Textual Analysis of the Great Depression:

DH100: Theory and Method in the Digital Humanities | Dr. Adam Anderson | Juliana Torres | Summer 2021

What is my takeaway?

FDR and Hoover had a very similar tone and vocabulary, but what was undeniable was the frequent use of the word "Government" even for classical liberal like Hoover. There is not much we can take from an exploratory data analysis that could be taken as an analysis, but the results do point us towards the next steps.

Where to go from here?

How can I take this research forward?

The dataset:

Ideally, I'd like to gather every speech given by FDR and Hoover during the Great Depression. This might require further digitalization of the records, and maybe some data scrapping.
The Methods:

The next steps would be approaching these texts from a computational hermeneutics method.

Using Kenneth Burke's five steps to analyze a text, I will use Named-Entity Recognition to identify the following:
what was done (act)

when or where it was done (scene)
who did it (agent)
how he did it (agency)
and why (purpose)

This way I hope to have a better overall understanding of what changed in the rhetoric around the Federal Government and the American people.

Who helped me?

Data Source:

UCSB American Presidency Project
https://www.presidency.ucsb.edu/

University of Virginia:
https://millercenter.org/

Technical Tools:

Towards Data Science:
https://towardsdatascience.com/from-dataframe-to-n-grams-e34e29df3460

https://towardsdatascience.com/sentimental-analysis-using-vader-a3415fef7664

Monkey Learn:
https://monkeylearn.com/blog/sentiment-classification/

Data Camp
https://www.datacamp.com/community/tutorials/wordcloud-python

DLab Python Notebooks

Cecilia Li, Summer 2020
shorturl.at/kANOP

DH100: Theory and Method in the Digital Humanities | Dr. Adam Anderson | Juliana Torres | Summer 2021