Decoding Presidential Rhetoric

How do the communication strategies and emotional undertones expressed in presidential speeches vary across different administrations?
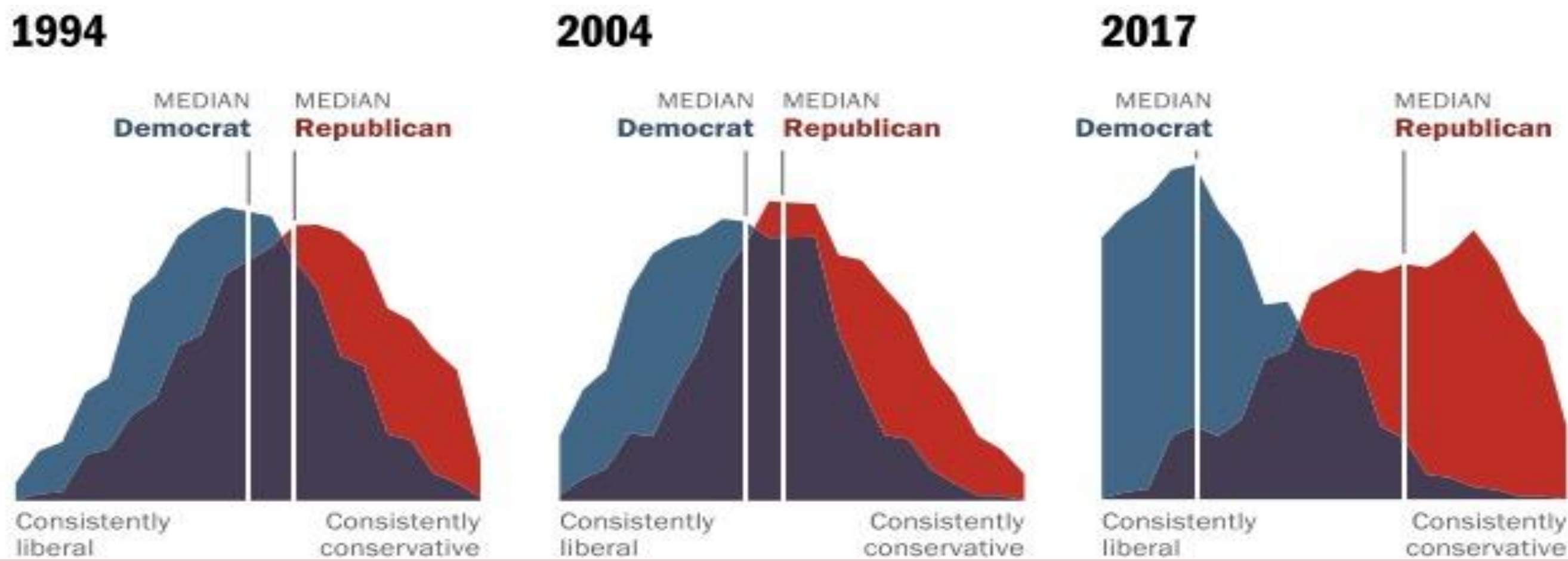
Introduction

Abstract

This project delves into the analysis of presidential speeches to unveil patterns in communication strategies by various administrations throughout history. Utilizing web scraping techniques, we collected data from the University of California, Santa Barbara Presidency Project. Through exploratory analysis, we identified prominent words and calculate pairwise distances among presidents. Sentiment analyses were then conducted which provided insights into the emotional sentiments conveyed in the speeches. Our findings shed light on the nuanced differences in communication styles across presidents and have significant implications for understanding the impact of presidential rhetoric on American politics and society.

Motivation and Insights from Prior Work

The United States is at an alarmingly high level of polarization and is polarizing faster than other democracies (Shapiro et al., 2020). Many people blame former President Trump for being divisive and exacerbating polarization in our country. Similarly, many Republicans argue that President Obama divided the country (Associated Press, 2017). Regardless of party, presidents and their speeches strongly impact the country. Simply put, what presidents say matters. This motivated us to use our technical skills to investigate the speeches of American presidents. Data opens up the possibility for us to compare presidents with each other in terms of how they communicate to the American public, specifically, what words they say.

1994 | **2004** | **2017**

MEDIAN Democrat / MEDIAN Republican

Consistently liberal — Consistently conservative

Data Collection

American Presidency Project: University of California, Santa Barbara

State of the Union Addresses: Annual speeches delivered by the President of the United States to a joint session of the United States Congress, providing an overview of the nation's condition and outlining legislative proposals for the upcoming year.
Inaugural Addresses: Delivered by newly inaugurated presidents at their inauguration ceremonies, setting forth their vision and agenda for their presidency.
Farewell Addresses: Delivered by outgoing presidents upon the conclusion of their terms in office, reflecting on their presidency and offering advice to the nation.
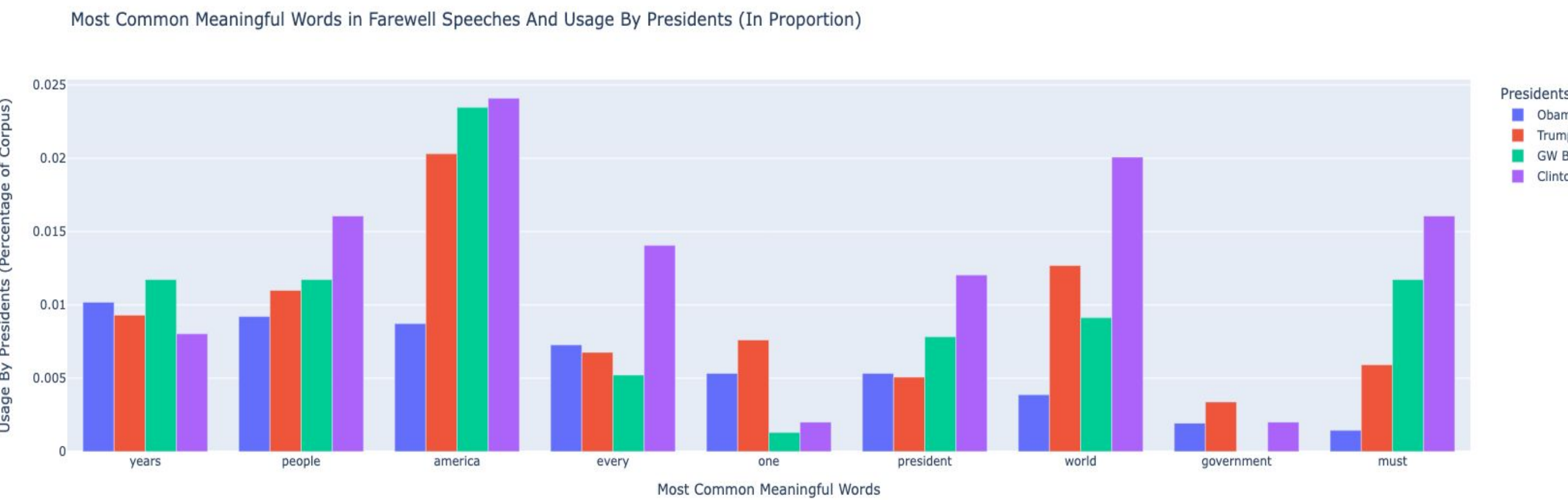
Web Scraping Process (Textual Data):
The web scraping process involved utilizing BeautifulSoup to extract speech data from the HTML structure of the web pages. This process allowed for the retrieval of presidential speeches from the Presidency Project. Through parsing the HTML content of each webpage, the relevant speech data, including the president's name, speech date, and speech content, was extracted and stored in a data frame.

Data Cleaning: Stripping the text of unnecessary characters such as additional spaces and '\n'. Filling in authorship using the previous non-empty President's name.

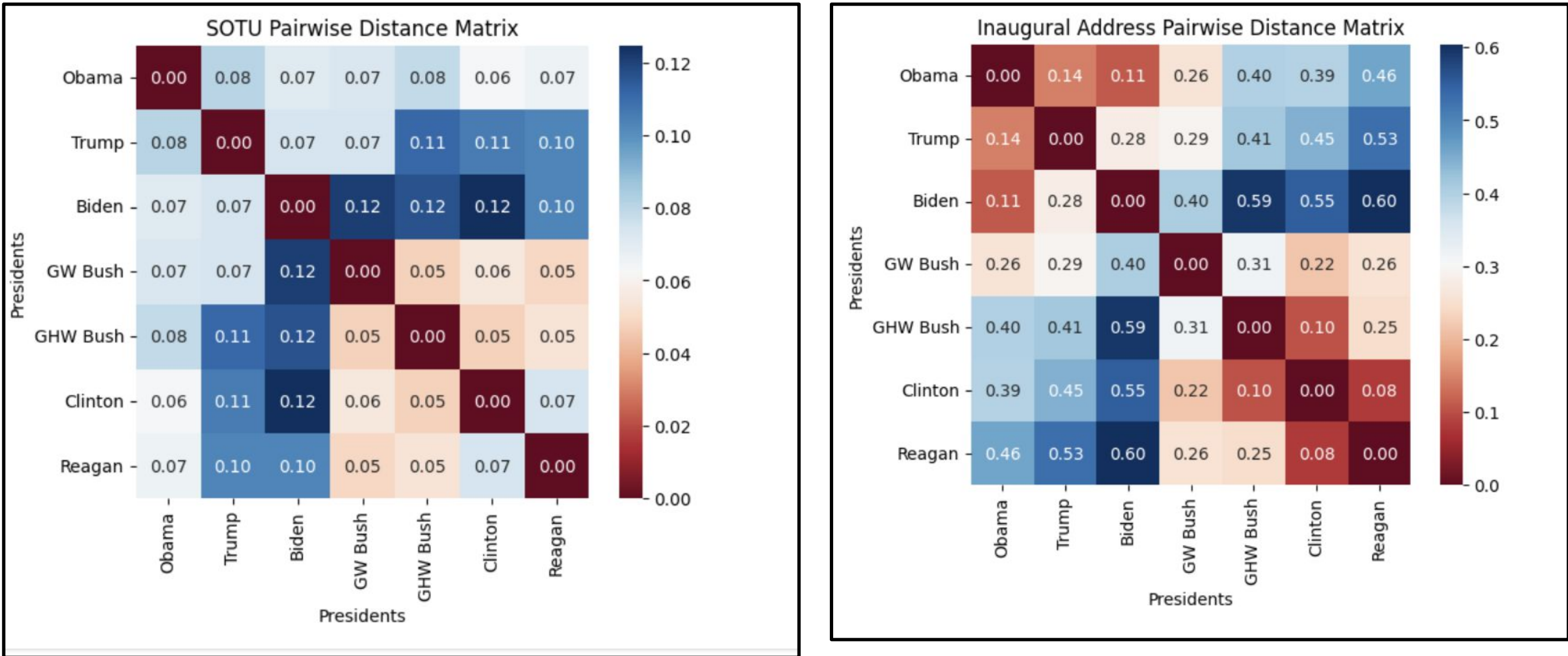## Exploratory Analysis

Term Frequency Matrix

A term frequency matrix is a good way to understand what words Presidents use and how Presidents compare to one another. Through CountVectorizer, we were able to create a TF matrix. An immediate problem we encountered was that the TF matrix was dominated by "non-value" words like pronouns and prepositions. We addressed this by downloading a set of stop words from a Natural Language Toolkit, and filtered those stop words out of the TF matrix. This route was chosen instead of a TF-IDF matrix because we wanted to preserve raw counts of each word used. With the TF matrix, we then able to plot the most common meaningful words used by US presidents and the usage frequencies and calculate pairwise distances. For this step, we filtered it down to the 5-7 most recent presidents for a finer data analysis. We then implemented the above steps for the State of the Union speeches, inaugural addresses, and farewell addresses.

Most Common Meaningful Words in Farewell Speeches And Usage By Presidents (In Proportion)

Clinton used words like America, world, every, must, and president more than other presidents, indicating that his farewell address may have been more energetic and included a stronger call to action.

Pairwise Distances

A major finding was that for State of the Union speeches, Democratic presidents tended to be most similar to Democratic presidents and Republican presidents tended to be most similar to Republican presidents. For Inaugural addresses, similarities were moreso along time period rather than political party.
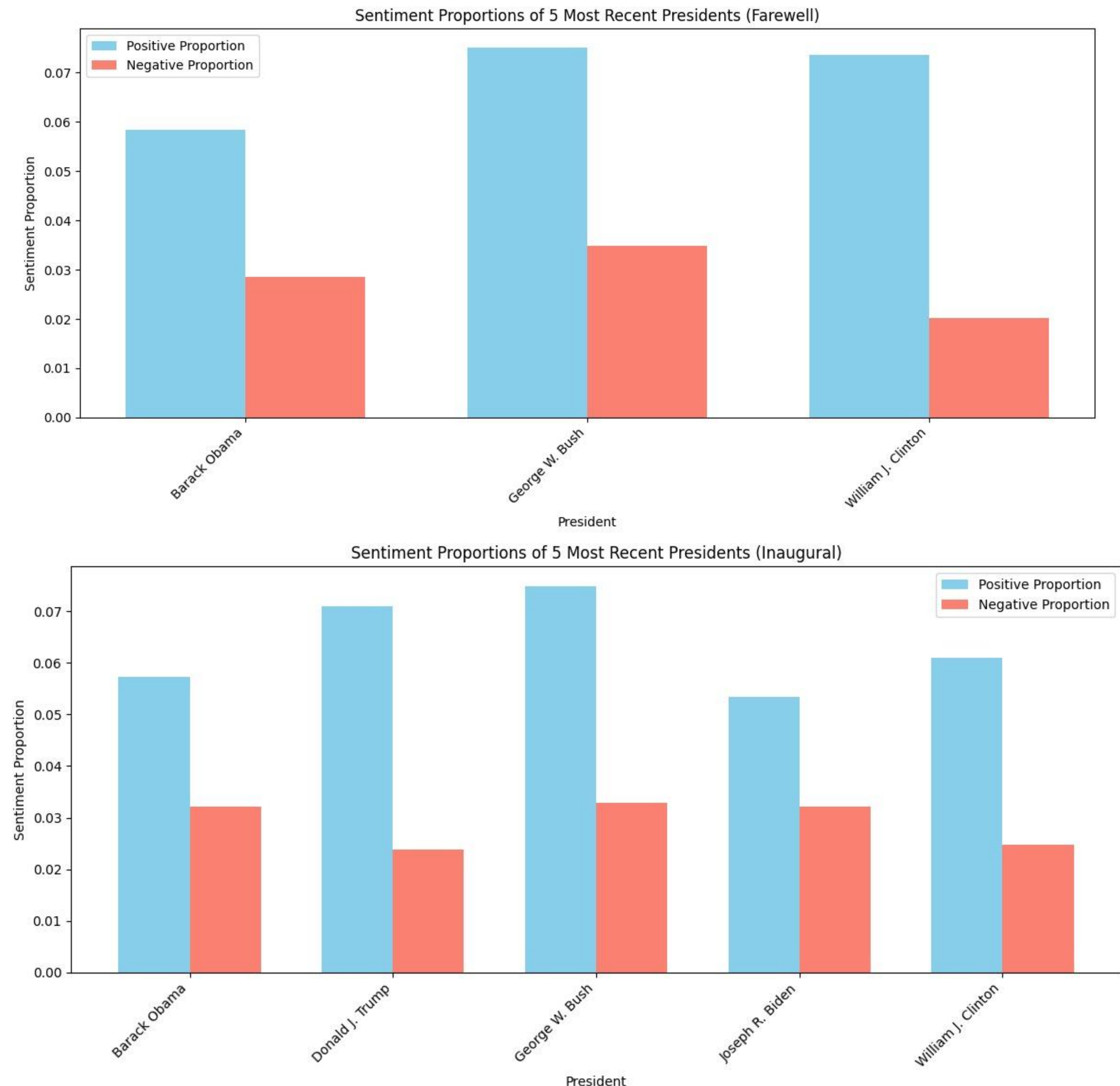
SOTU Pairwise Distance Matrix | Inaugural Address Pairwise Distance Matrix

Heatmap was created using seaborn and matplot

Sentiment Analysis

A sentiment analysis can better grasp the emotional undertones conveyed in presidential speeches, providing a nuanced understanding of the sentiments expressed by various administrations throughout history.

AFFIN Lexicon: A curated collection of words each assigned to a sentiment score.
Sentiment Score: Aggregated sentiment scores of the constituent words within each speech, arriving at a numerical representation encapsulating the overall sentiment expressed in the speech.
Sentiment Proportions: The proportion of positive and negative words.

Sentiment Proportions of 5 Most Recent Presidents (Farewell)

Sentiment Proportions of 5 Most Recent Presidents (Inaugural)

Inaugural vs Farewell: Bush had a slightly higher positive proportion in his inaugural address compared to his farewell address, indicating a consistent tone throughout his presidency. Clinton exhibited a higher positive proportion in his farewell compared to his inaugural, suggesting a shift toward a congratulatory, energetic tone and a more urgent call to action in his farewell.

\Inter-Presidential Sentiment Comparisons: Trump's inaugural address had the highest positive proportion among the selected presidents, indicating a more optimistic tone at the beginning of his presidency. In contrast, Biden's inaugural address exhibited a relatively lower positive proportion, potentially reflecting a more subdued tone as he assumed office amidst challenging circumstances, such as the COVID-19 pandemic and political polarization.

Ethical Considerations

Adherence to Web Scraping Guidelines: Carefully read through the website's robots.txt file and terms of service, ensuring that scraping activities were done responsibly and in accordance with the website.

Server Overload: Implemented time delays between requests, allowing the website to handle our requests without overwhelming the website's resources.

Real World Implications

This study has strong real-world applications for American politics. It is crucial to understand how presidents are communicating and how they differ from one another. Our study provides a concrete, unbiased, evaluation of presidents and the semantics they use. These findings are important, especially in an age of hyper-partisanship in the media. Our study also reveals that political parties may have a strong role in how presidents communicate, as presidents from the same political parties tended to give similar speeches (especially SOTU speeches)