

NLP & Inaugural Addresses

Andrew Smith

Inaugural Address: Background

- Used to set forth vision and goals of the nation
- Rare that policy is discussed (unlike the State of the Union), focus is on continuity and bipartisanship
- Seek to assuage fears or concerns of the time

Inaugural Address: Background

- Used to set forth vision of America and set forth their goals of the nation
- Rare that policy is discussed (unlike the State of the Union), focus is on continuity and bipartisanship
- Seek to assuage fears or concerns of the time

Inaugural Address: Background

- Used to set forth vision of America and set forth their goals of the nation
- Rare that policy is discussed (unlike the State of the Union), focus is on continuity and bipartisanship
- Seek to assuage fears or concerns of the time

The only thing we have to fear
is fear itself.

Franklin D. Roosevelt



www.thequotes.in

Data Sources & Data Description

- Variables:
 - Corpus - Text files of all inaugural Addresses (58 total)
 - Added political party for each president

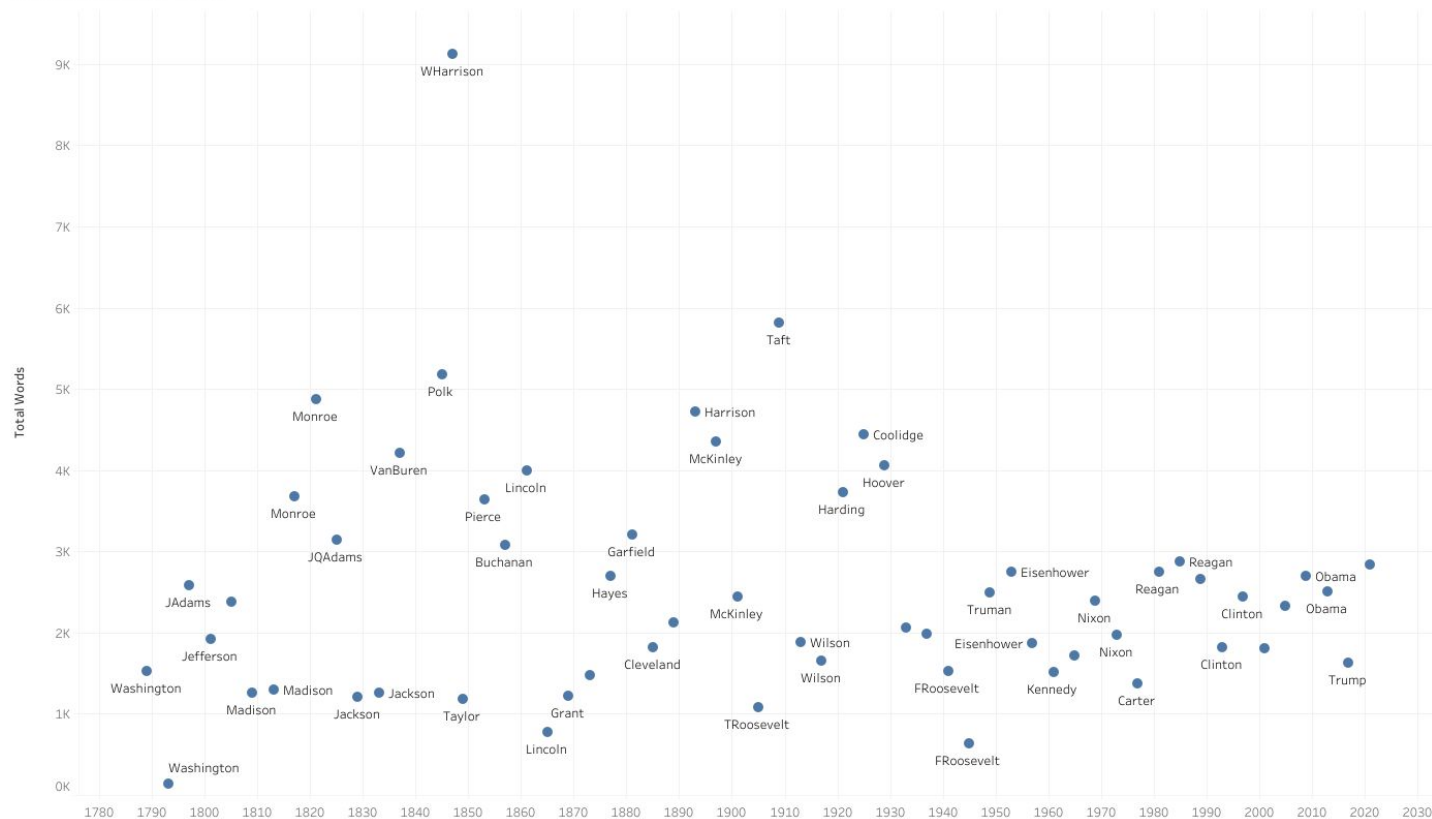
Workflow Process

- Exploratory Data Analysis
 - Address Word Length
 - Words / Sentence
 - “I” vs. “We”
- First Round of Topic Modelling
 - Removed default stop words
 - Added additional stop words
 - Lemmatized
 - Removed words that were not nouns, adjective, verbs, adverbs
 - Modelled with LDA and NMF
 - Optimized for min_df / max_df, ngram_range
 - Tested various number of topics
- Second Round of Topic Modelling
 - Tried CorEx
 - Continue with NMF
 - Tested from the sentence and paragraph level

Exploratory Data Analysis: *Total Words*

Minor decline in speech word length

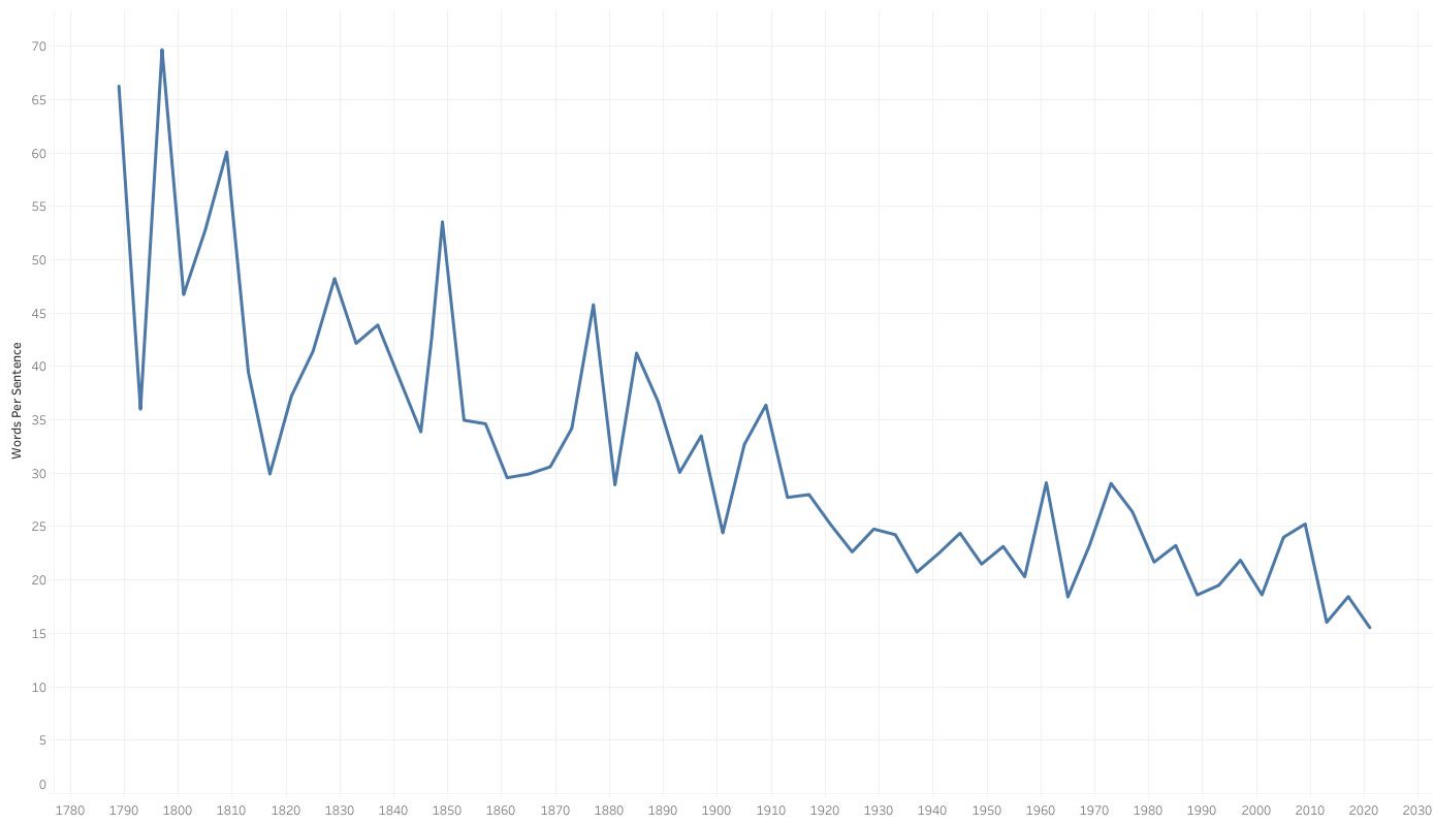
Total Words over Time



Exploratory Data Analysis: *Sentence Length*

Sharp decline in words per sentence over time

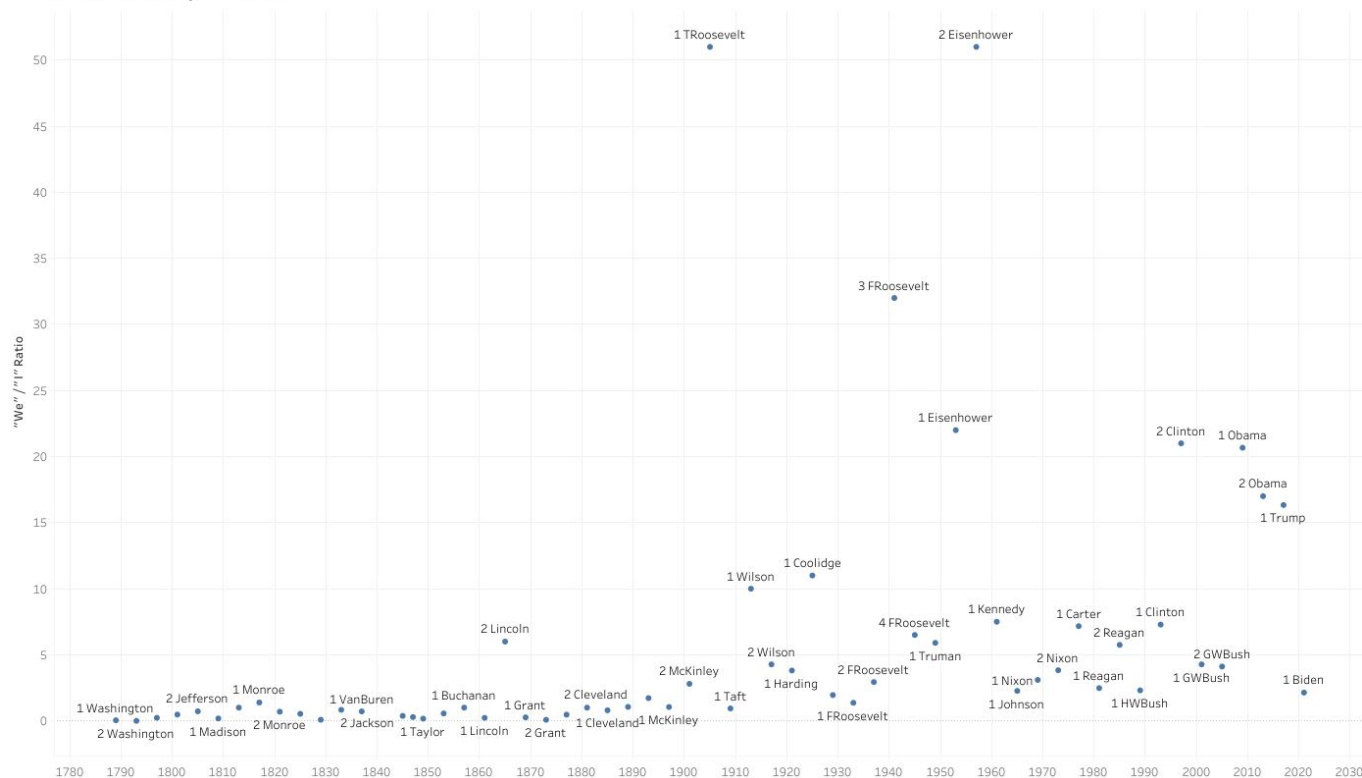
Words per Sentence



Exploratory Data Analysis: “We” vs. “I”

General increase in “We” vs. “I” over time

“I” to “We” Ratio by President



Workflow Process

- Exploratory Data Analysis
 - Address Word Length
 - Words / Sentence
 - “I” vs. “We”
- First Round of Topic Modelling
 - Removed default stop words
 - Added additional stop words
 - Lemmatized
 - Removed words that were not nouns, adjective, verbs, adverbs
 - Modelled with LDA and NMF
 - Optimized for min_df / max_df, ngram_range
 - Tested various number of topics
- Second Round of Topic Modelling
 - Tried CorEx
 - Continue with NMF
 - Tested from the sentence and paragraph level

Workflow Process

- Exploratory Data Analysis
 - Address Word Length
 - Words / Sentence
 - “I” vs. “We”
- First Round of Topic Modelling
 - Removed default stop words
 - Added additional stop words
 - Lemmatized
 - Removed words that were not nouns, adjective, verbs, adverbs
 - Modelled with LDA and NMF
 - Optimized for min_df / max_df, ngram_range
 - Tested various number of topics
- Second Round of Topic Modelling
 - Tried CorEx
 - Continue with NMF
 - Tested from the sentence and paragraph level

Workflow Process

- Exploratory Data Analysis
 - Address Word Length
 - Words / Sentence
 - “I” vs. “We”
- First Round of Topic Modelling
 - Removed default stop words
 - Added additional stop words
 - Lemmatized
 - Removed words that were not nouns, adjective, verbs, adverbs
 - Modelled with LDA and NMF
 - Optimized for min_df / max_df, ngram_range
 - Tested various number of topics
- Second Round of Topic Modelling
 - Tried CorEx
 - Continue with NMF
 - Tested from the sentence and paragraph level

Workflow Process

- Exploratory Data Analysis
 - Address Word Length
 - Words / Sentence
 - “I” vs. “We”
- First Round of Topic Modelling
 - Removed default stop words
 - Added additional stop words
 - Lemmatized
 - Removed words that were not nouns, adjective, verbs, adverbs
 - Modelled with LDA and NMF
 - Optimized for min_df / max_df, ngram_range
 - Tested various number of topics
- Second Round of Topic Modelling
 - Tried CorEx
 - Continue with NMF
 - Tested from the sentence and paragraph level

Workflow Process

- Exploratory Data Analysis
 - Address Word Length
 - Words / Sentence
 - “I” vs. “We”
- First Round of Topic Modelling
 - Removed default stop words
 - Added additional stop words
 - Lemmatized
 - Removed words that were not nouns, adjective, verbs, adverbs
 - Modelled with LDA and NMF
 - Optimized for min_df / max_df, ngram_range
 - Tested various number of topics
- Second Round of Topic Modelling
 - Tried CorEx
 - Continue with NMF
 - Tested from the sentence and paragraph level

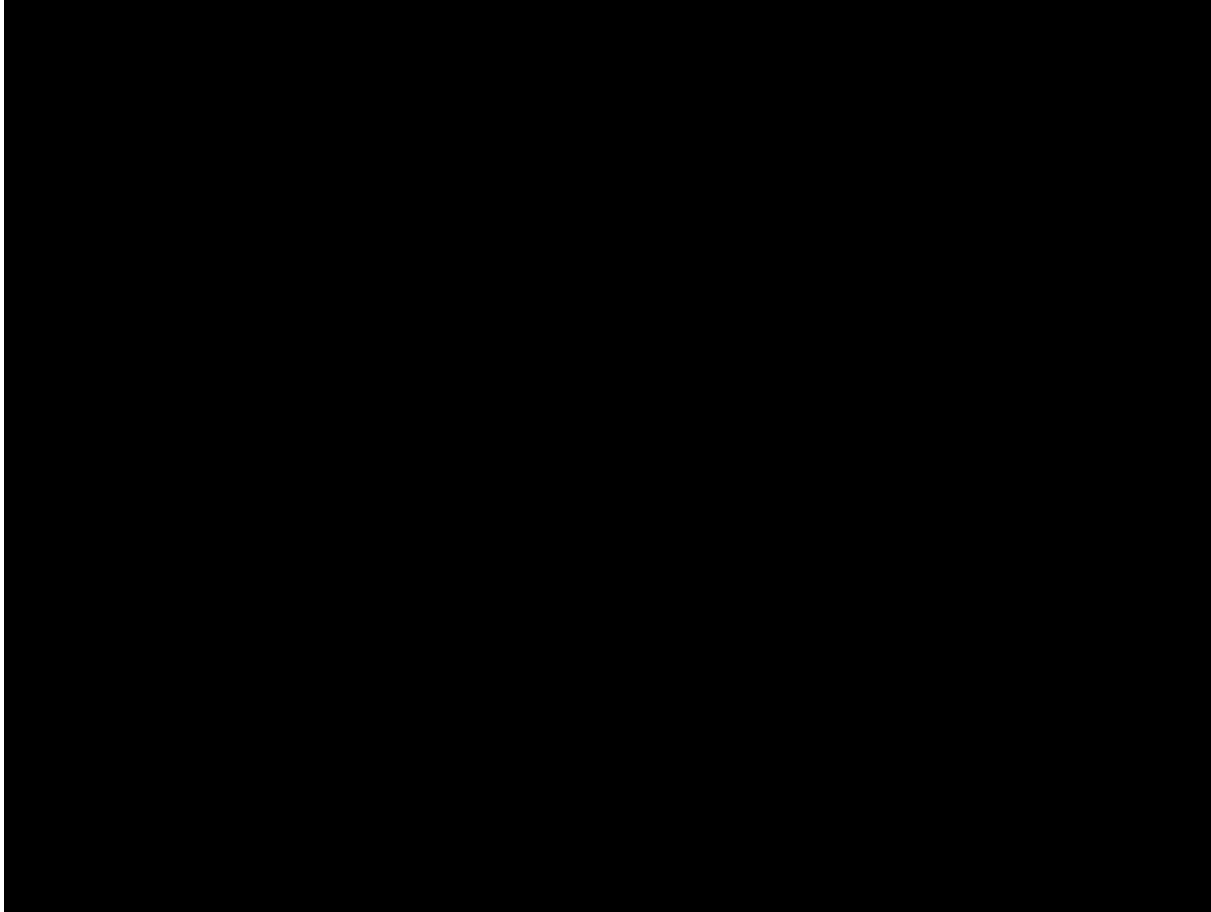
Workflow Process

- Exploratory Data Analysis
 - Address Word Length
 - Words / Sentence
 - “I” vs. “We”
- First Round of Topic Modelling
 - Removed default stop words
 - Added additional stop words
 - Lemmatized
 - Removed words that were not nouns, adjective, verbs, adverbs
 - Modelled with LDA and NMF
 - Optimized for min_df / max_df, ngram_range
 - Tested various number of topics
- Second Round of Topic Modelling
 - Tried CorEx
 - Continue with NMF
 - Tested from the sentence and paragraph level

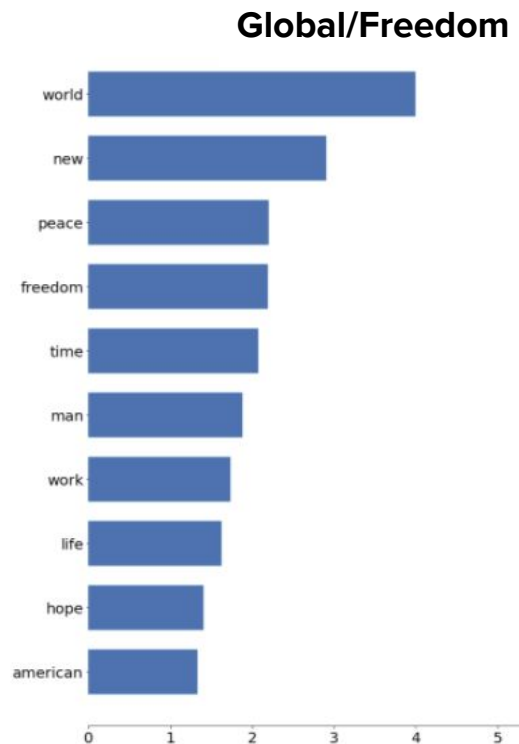
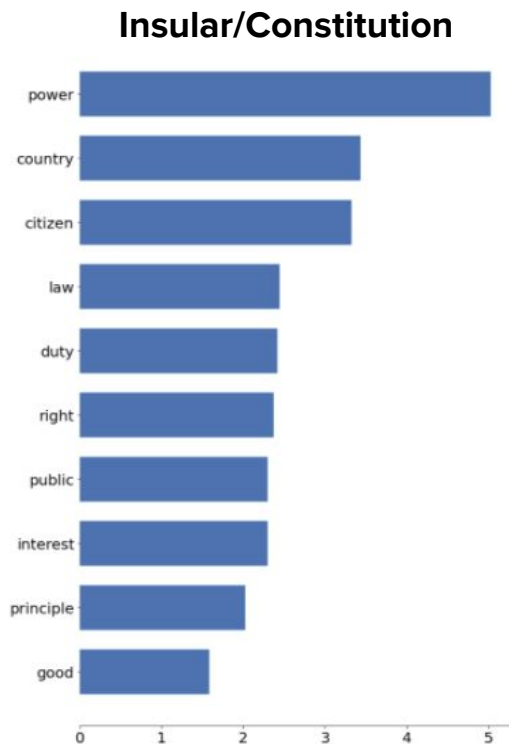
Workflow Process

- Exploratory Data Analysis
 - Address Word Length
 - Words / Sentence
 - “I” vs. “We”
- First Round of Topic Modelling
 - Removed default stop words
 - Added additional stop words
 - Lemmatized
 - Removed words that were not nouns, adjective, verbs, adverbs
 - Modelled with LDA and NMF
 - Optimized for min_df / max_df, ngram_range
 - Tested various number of topics
- Second Round of Topic Modelling
 - Tried CorEx
 - Continue with NMF
 - Tested from the sentence and paragraph level

Topic Modelling: 2 Components



Topic Modelling: 2 Components



Topic Modelling:



McKinley onward, inaugural address topics switch from US centric and upholding constitutional rights to the world and peace.

Topic Modelling: 2 Components

US abruptly becomes a world power by the end of the 19th century

Early 1890s: no interest in foreign affairs

1898: US wins the Spanish American War and annexes Puerto Rico, Guam, and the Philippines from Spain (McKinley)

1898: U.S. annexes Hawaii

1899: Out of fear of being shut out of trade with China, the Roosevelt Corollary to the Monroe Doctrine declare the U.S. would exercise “international police power” in the Western Hemisphere



Workflow Process

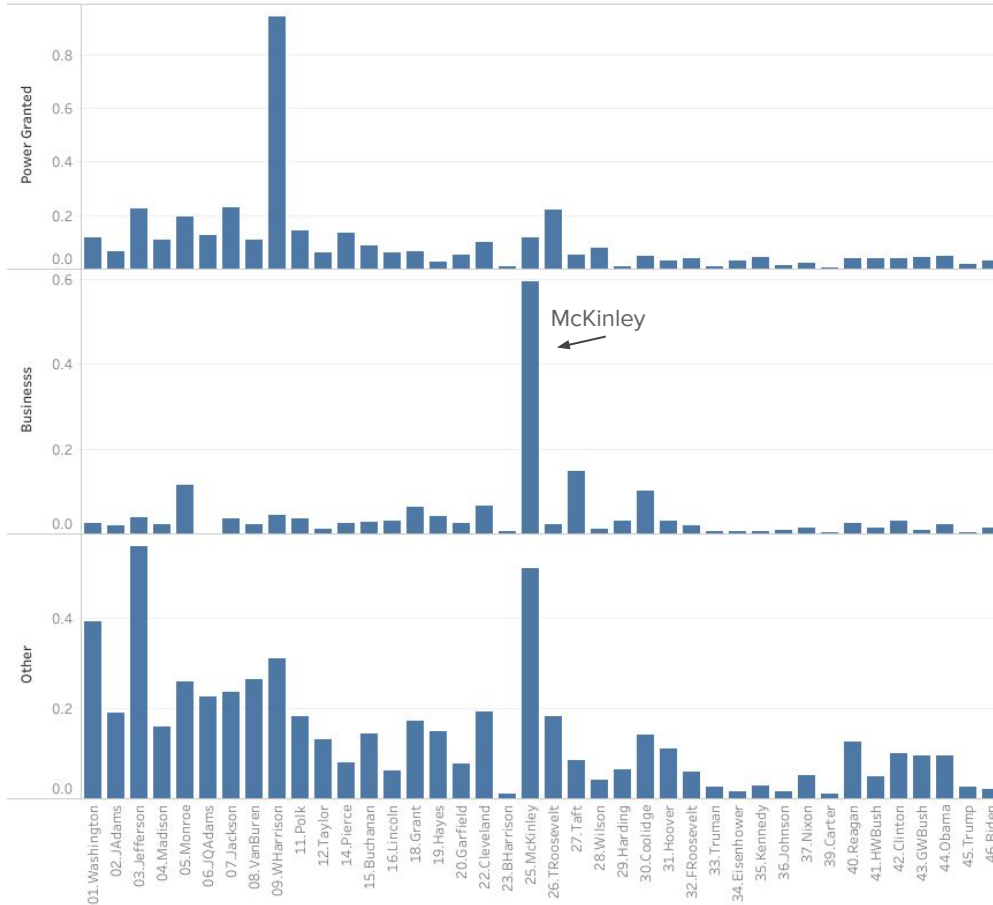
- Exploratory Data Analysis
 - Address Word Length
 - Words / Sentence
 - “I” vs. “We”
- First Round of Topic Modelling
 - Removed default stop words
 - Added additional stop words
 - Lemmatized
 - Removed words that were not nouns, adjective, verbs, adverbs
 - Modelled with LDA and NMF
 - Optimized for min_df / max_df, ngram_range
 - Tested various number of topics
- Second Round of Topic Modelling
 - Tried CorEx
 - Continue with NMF
 - Divided text at the sentence and paragraph level

14 Component Grouping

Topic Name	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6
other	country	find	much	high	part	many
other	government	island	already	order	inhabitant	self
other	member	power	however	institution	separate	principle
power granted	power	executive	act	judiciary	well	want
power granted	power	control	state	executive	government	officer
power granted	power	grant	sovereignty	grant power	right	possess
American spirit	spirit	liberty	power	character	government	free
American spirit	party	whole	liberty	interest	spirit	country
protecting us interests	right	free	respect	interest	American	foreign
protecting us interests	revenue	interest	protection	duty	home	equally
world freedom	world	new	freedom	American	work	time
upholding amendments	law	man	enforce	pass	amendment	support
war/peace/internal	party	political	peace	time	war	government
preserve values	spirit	honor	interest	preserve	love	wish
war/peace/external	peace	policy	war	world	foreign	treaty
political agitation	institution	interest	never	subject	political	agitation
war	war	force	invasion	power	time	naval
public service	public	duty	good	service	office	economy
business	revenue	business	well	party	system	country
upholding constitution	time	hope	believe	constitutional	good	right

20 Component Grouping: *Power Granted, Business, Other*

Power Granted, Business, Other

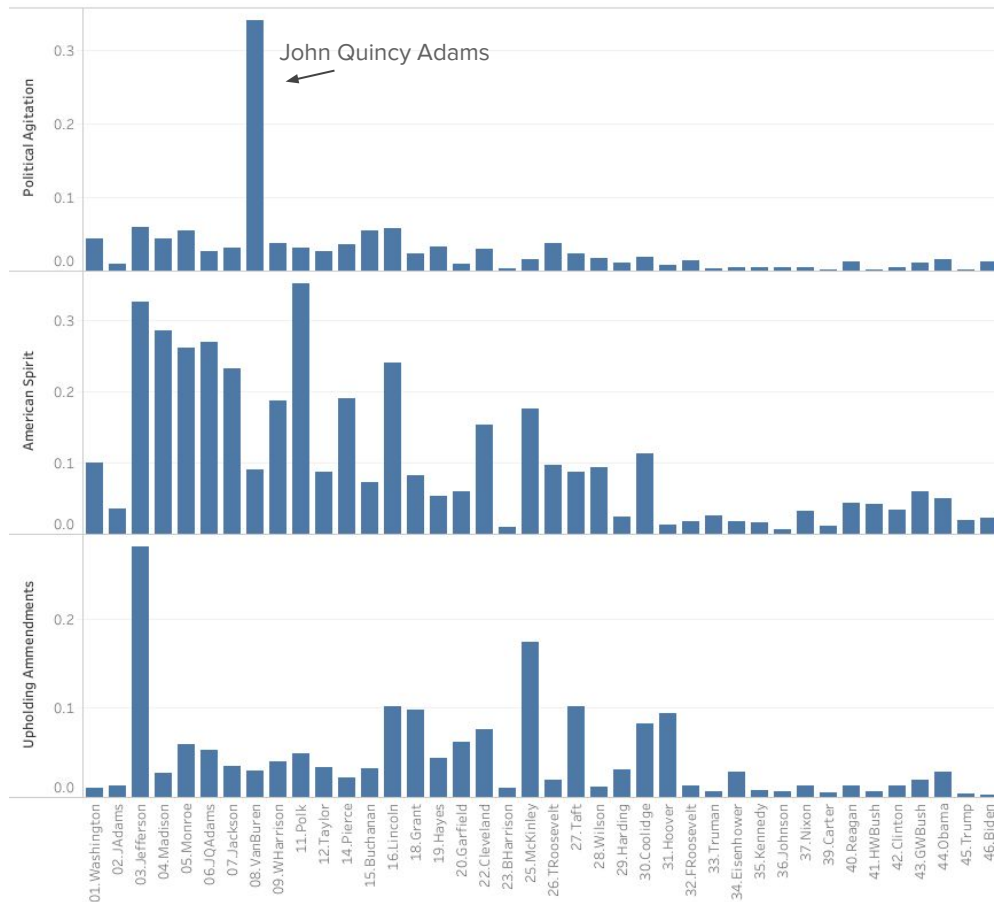


Decline in discussion on power granted to president and branches of government

Proponent of big business through protectionism and high tariffs

20 Component Grouping: *Agitation, Spirit, Amendments*

Political Agitation, American Spirit, Upholding Amendments



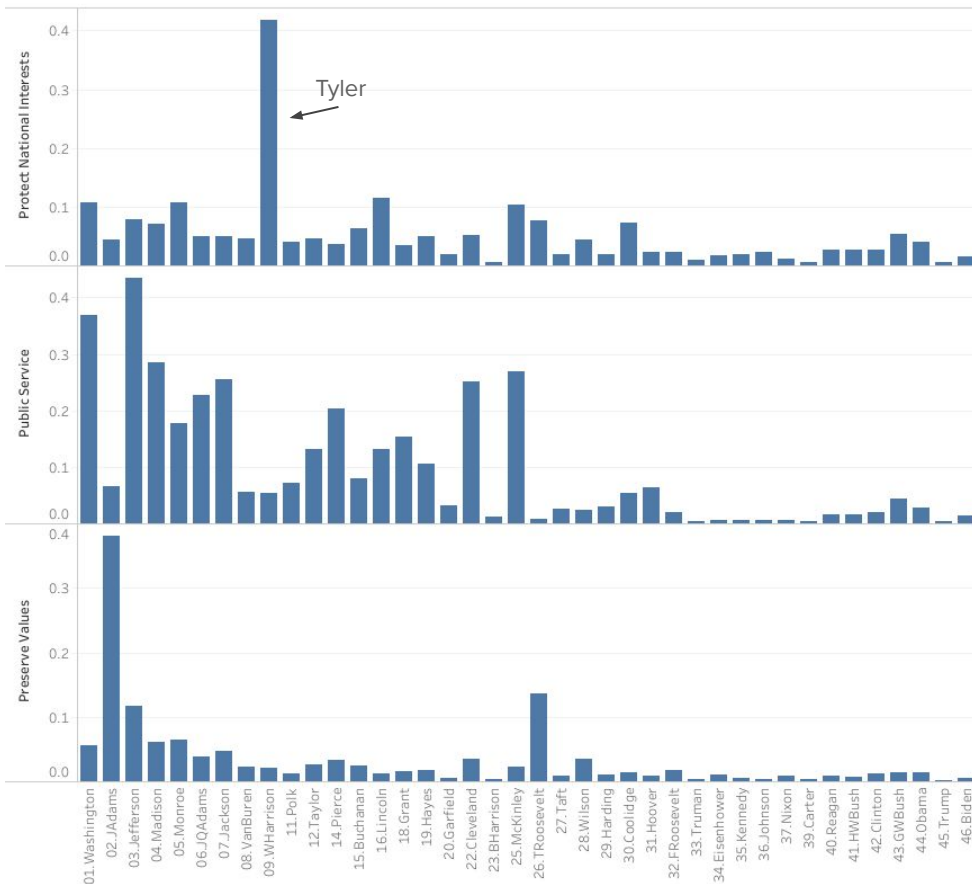
There was no majority electoral college vote and John Quincy Adams was elected by the House of Representatives. There was fear of potential violence and he tried to heal electoral divides in his address.

Decline in words related to the founding principles of the U.S. ("spirit", "liberty", "free")

Increased discussion around amendments after the civil war and WWI

20 Component Grouping: *Interests*, *Public Service*, *Values*

Protect National Interests, Public Service, Preserve Values



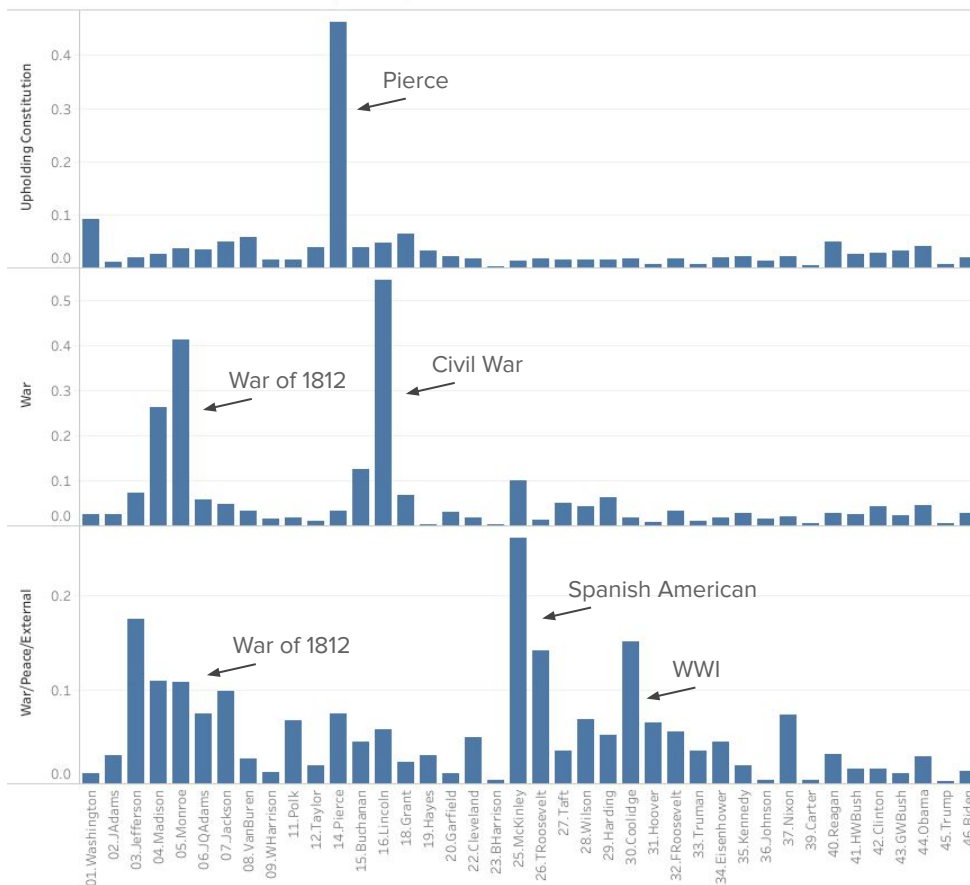
Stressed how he would scale back “monarchical” Jacksonian democracy, for fear of violence and decline (i.e. Cromwell, Caesar).

Decline in discussion on president’s public duty

Decline in discussion on preserving values

20 Component Grouping: *Constitution, War, Foreign Affairs*

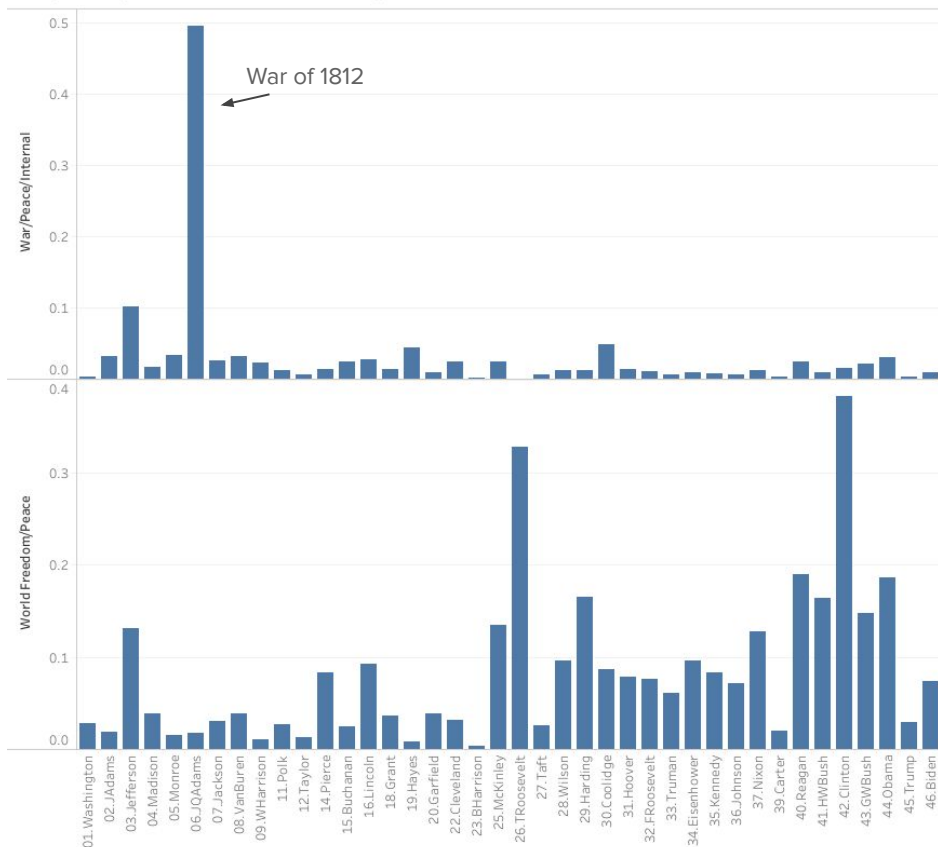
Uphold Constitution, War, War/Peace/External



Stresses states rights and the balance between the federal government; harbinger of Civil War

20 Component Grouping: *War/Peace*, *World Peace/Freedom*

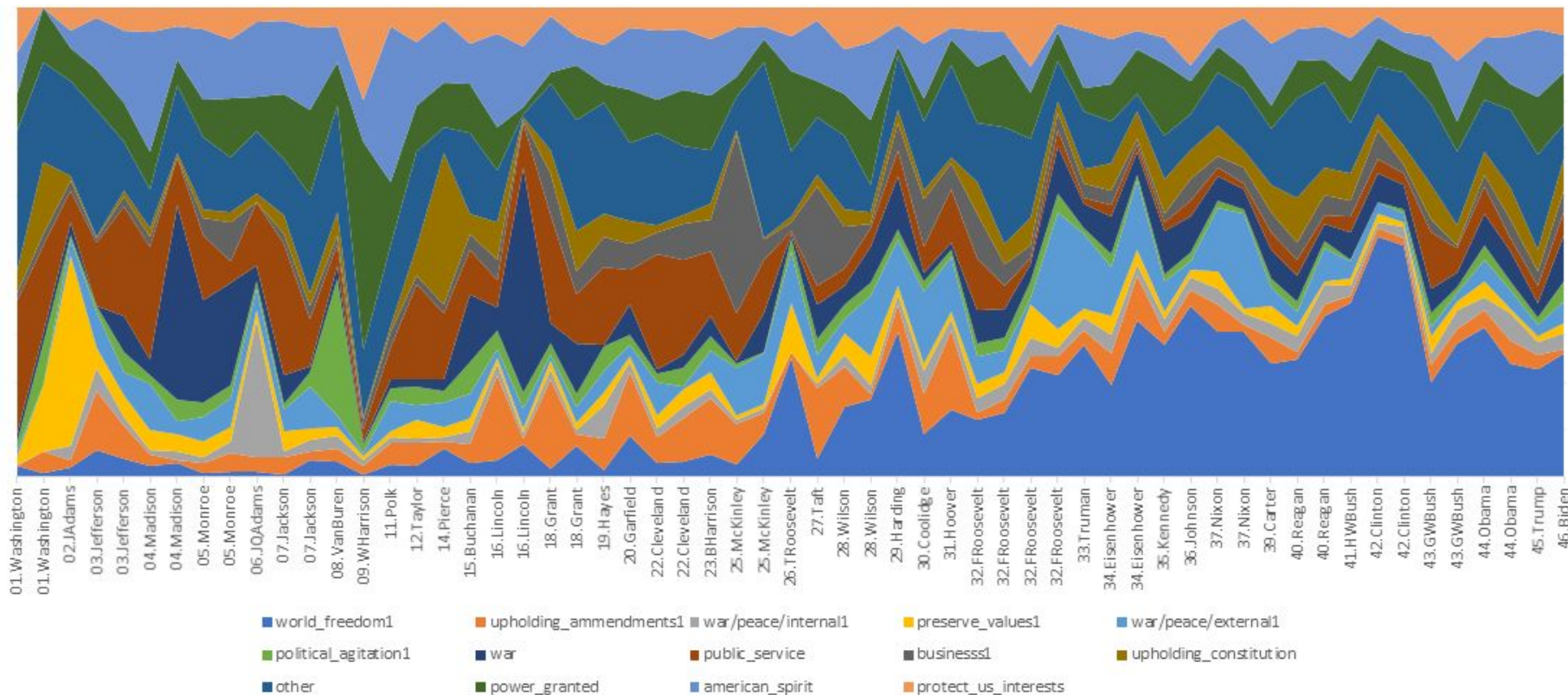
War/Peace/Internal, World Freedom/Peace



Increased global focus following McKinley

20 Component Grouping: *Share*

Share of Topics over Time



Appendix

Exploratory Data Analysis: *Sentiment Analysis*

Positive Sentiment Score for all Inaugural Addresses

