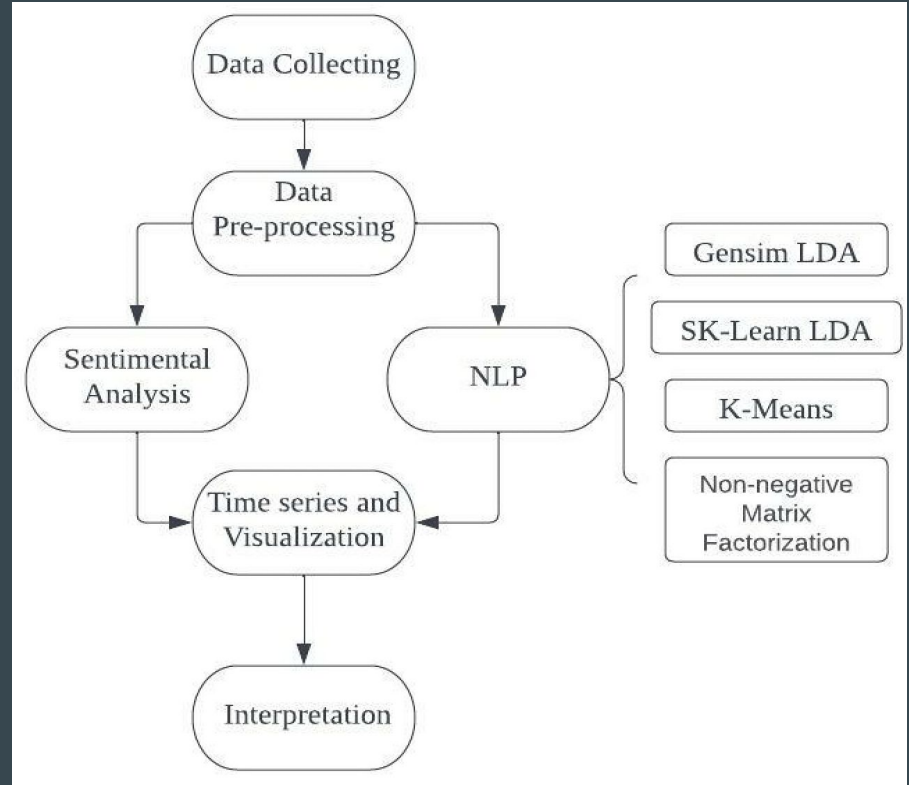# Obama Speech Analysis

• • •

Sihan Wang

Speech analyzing procedures

1. Data collection
2. Data cleaning
3. NLP with different methods
4. Sentimental Analysis
5. Topic Modeling Visualization
6. Conclusion

Data Collecting and Pre-processing

Data Collecting

1. BeautifulSoup to fetch the URL of each speech

'https://www.americanrhetoric.com/speeches/barackobama/weeklyaddresses/barackobamaweekly1.htm'

2. BeautifulSoup to extract the title, date, and content of each speech

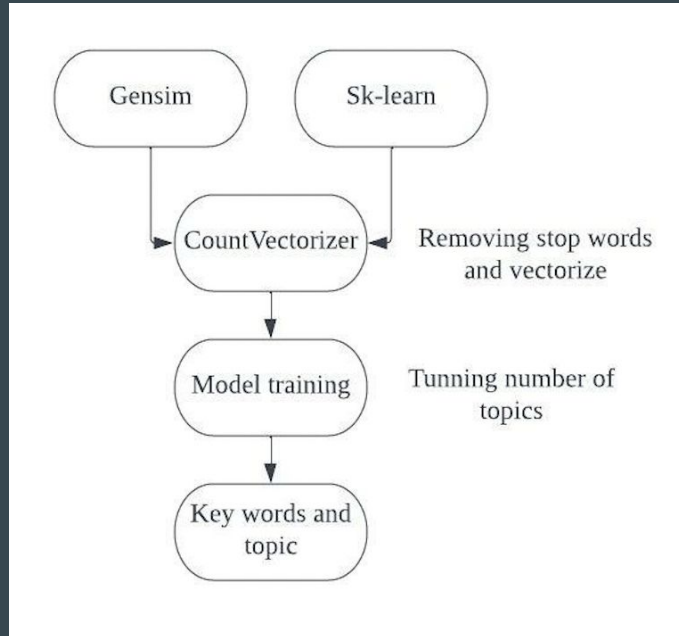|	url	content	title	date
0	https://www.americanrhetoric.com/speeches/bara...	1 Thank you, thank you. My fellow citizens: I...	FIRST PRESIDENTIAL INAUGURAL ADDRESS	20 Jan 2009
1	https://www.americanrhetoric.com/speeches/bara...	We begin this year and this Administration i...	First Presidential Weekly Speech	24 Jan 2009
2	https://www.americanrhetoric.com/speeches/bara...	Mr. Melhem: Mr. President, thank you for this...	Al-Arabiya Television Interview	26 Jan 2009
3	https://www.americanrhetoric.com/speeches/bara...	Thank you, Tim, for your hard work on this i...	Executive Compensation Salary Cap Speech	04 Feb 2009
4	https://www.americanrhetoric.com/speeches/bara...	Good evening, everybody. Please be seated. B...	First Prime Time Press Conference	09 Feb 2009

Data Pre-processing

1. Regular expression and nltk tokenizer

    – Remove all punctuations, numbers, special characters, then lowercase

2. Stemmer

    – Lower the inflation of words into root words by cutting the prefix and suffix

3. Word Lemmatizer

    – Classify the nouns and verbs in the speeches

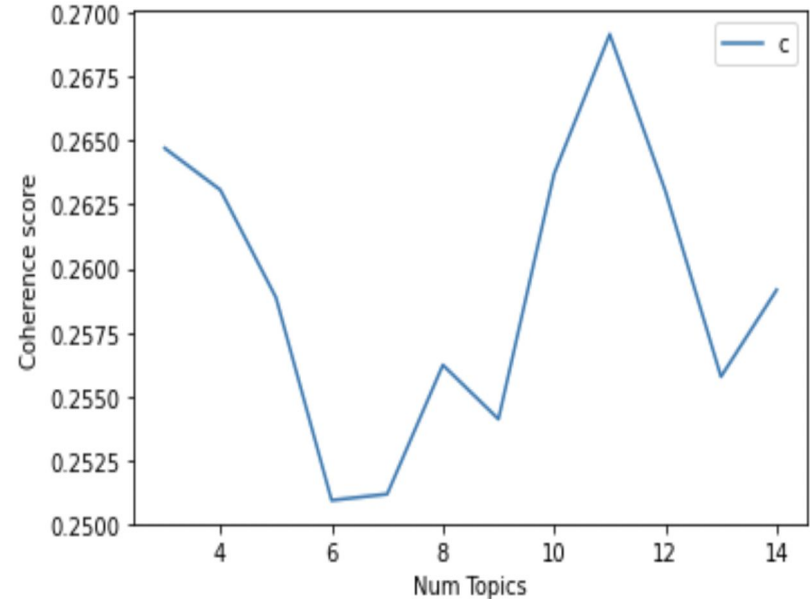|	date	title	content	clean_content	content_nouns
0	20 Jan 2009	First Presidential Inaugural Addre	[Chief Justice John G. \n\t\tRoberts administe...	chief justice john robert administer oath offi...	justice john robert administer oath office s bar...
1	24 Jan 2009	>American Rhetoric: Barack Obama: First Presid...	: First Presidential Weekly Address \n(01-24-0...	presidential weekly address barack obama presi...	address barack addressfirst broadcast washingt...

NLP Topic Modelling

Models: Gensim and Sk-Learn
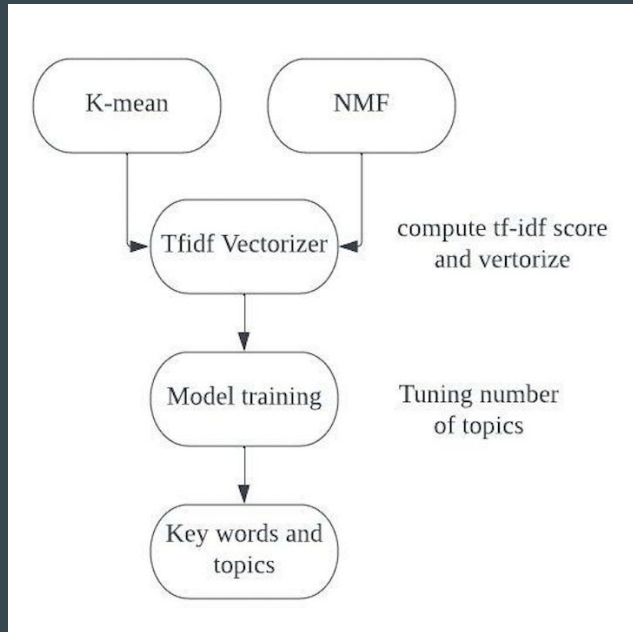
Model procedure

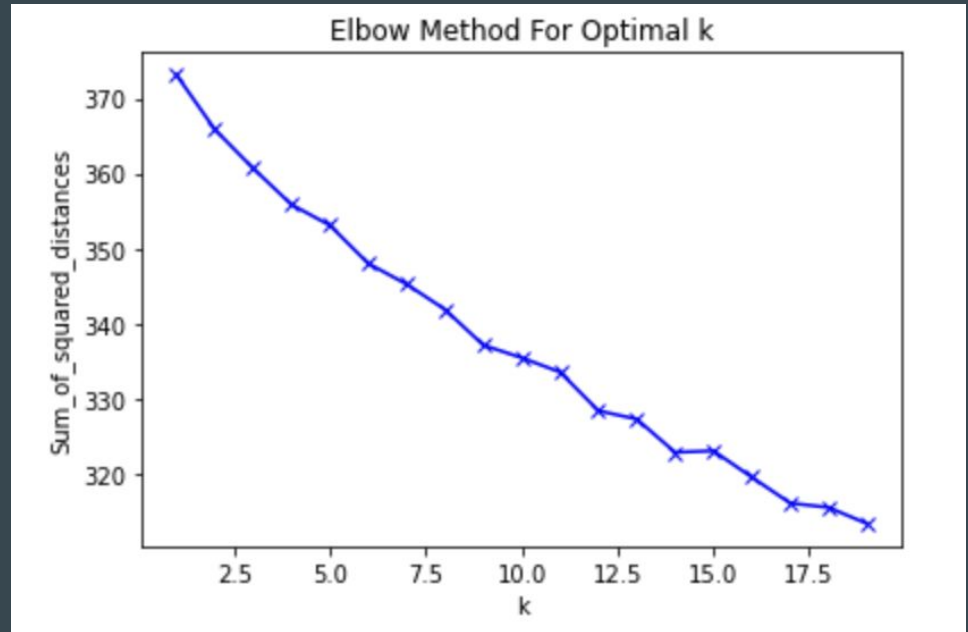Optimal topic number = 11

Models: K-mean and NMF

Model procedure
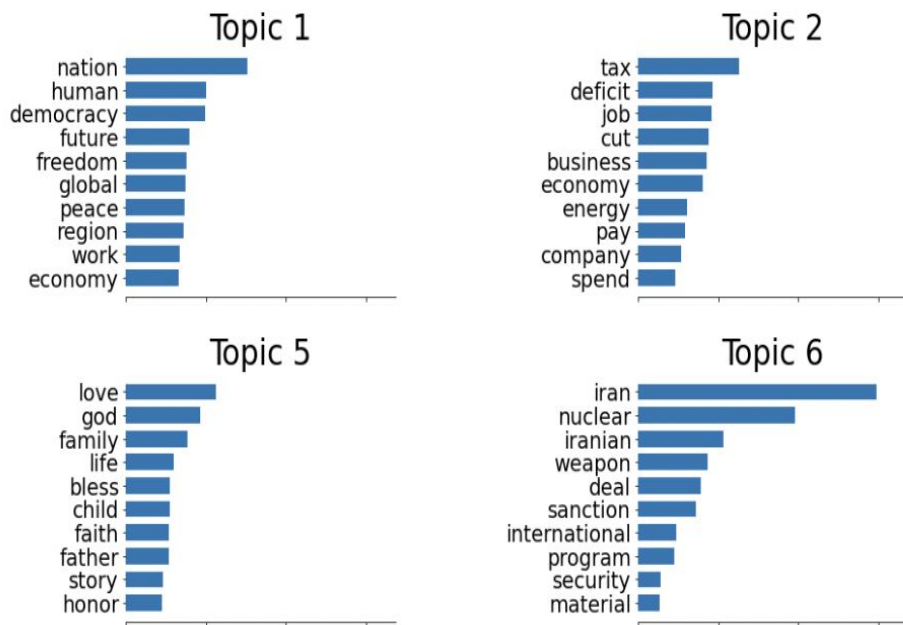
Optimal topic number (no clear answer)

Result of topic modeling

Output of Gensim model

```
[(0,
  '0.009*"security" + 0.009*"work" + 0.00
9*"isil" + 0.007*"terrorist" + '
  '0.007*"military" + 0.007*"continue" +
0.006*"government" + 0.006*"force" + '
  '0.005*"attack" + 0.005*"syria"'),
 (1,
  '0.015*"health" + 0.014*"care" + 0.012
*"work" + 0.011*"insurance" + '
  '0.010*"question" + 0.009*"iran" + 0.00
9*"obama" + 0.007*"deal" + '
  '0.007*"republican" + 0.006*"good"'),
 (2,
  '0.014*"tax" + 0.012*"cut" + 0.011*"gov
ernor" + 0.010*"israel" + '
  '0.009*"romney" + 0.009*"deficit" + 0.0
08*"job" + 0.008*"pay" + 0.007*"work" '
  '+ 0.007*"plan"'),
 (3,
  '0.011*"question" + 0.011*"obama" + 0.0
10*"work" + 0.006*"good" + '
  '0.006*"government" + 0.006*"business"
+ 0.005*"start" + 0.005*"economy" + '
  '0.005*"young" + 0.004*"year"')]
```

Output of NMF model

Sentiment Analysis
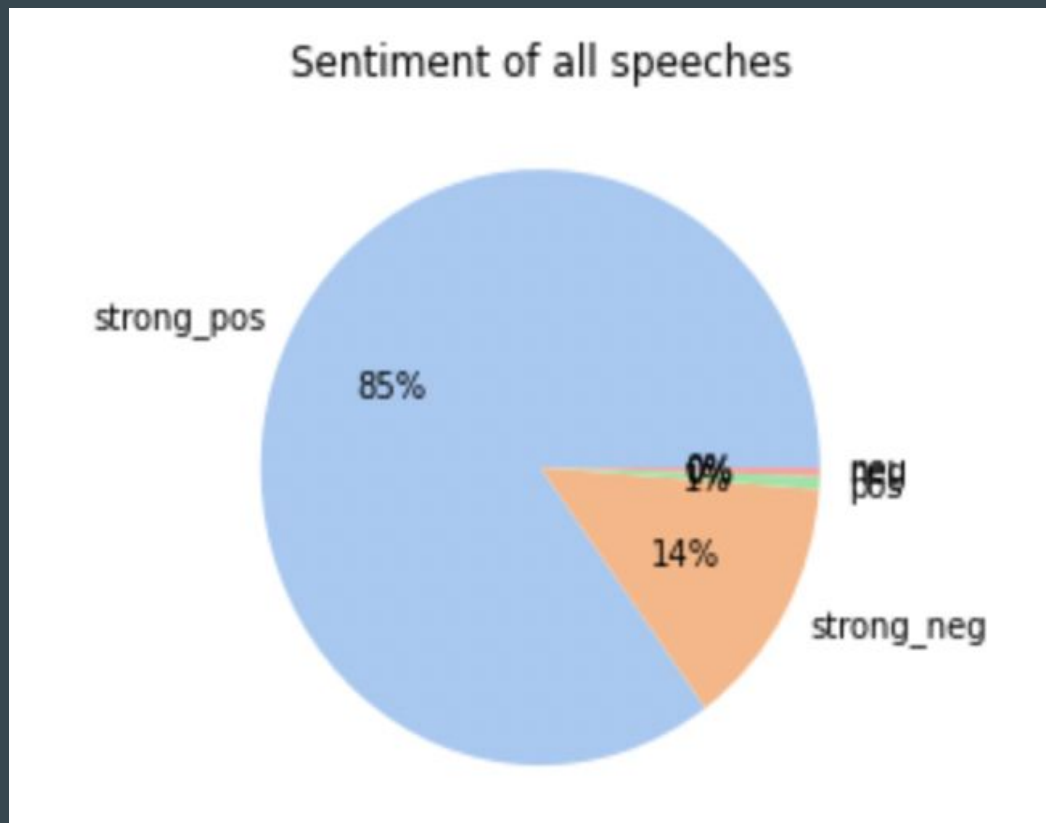
Sentiment analysis

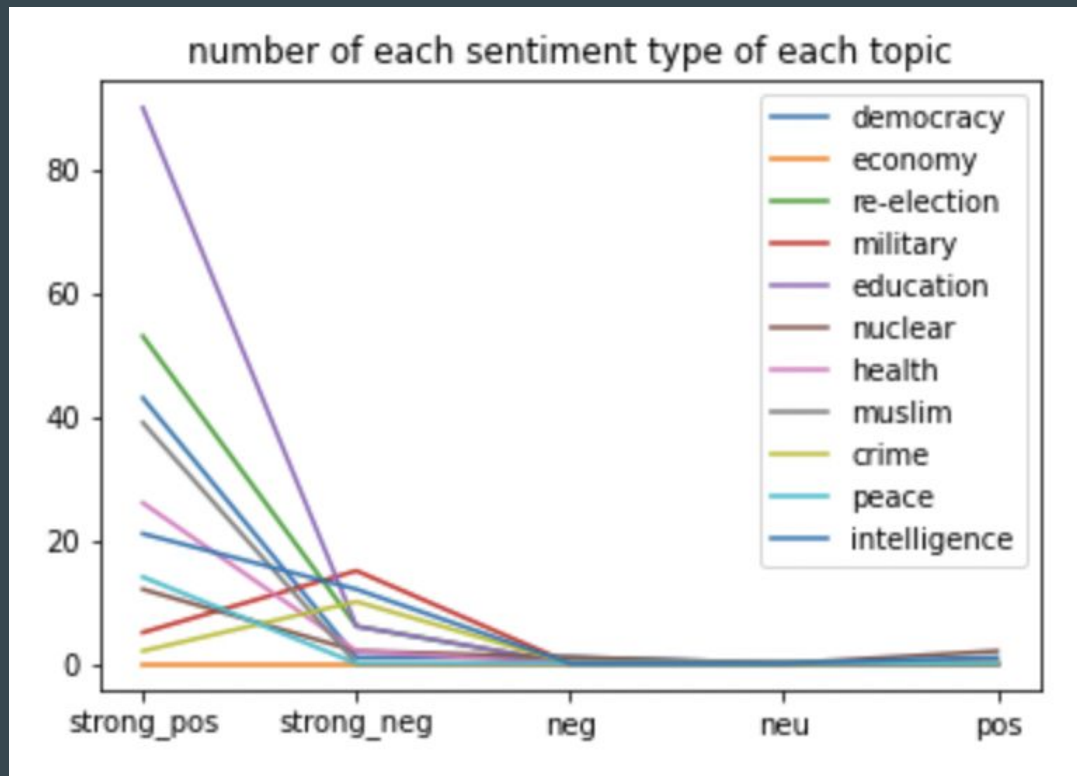# Overall Sentiments

Strong Positive: 364

Positive: 3

Neutral; 0

Negative: 2

Strong Negative: 59

Sentiment of all speeches

strong_pos 85%

0%

1%

14%

strong_neg

Sentiments for each topic

- Most of the topics have more positive sentiment

- Military topic has strongest negative sentiment compare to other topics

- When talking about democracy, health care, labour, economy and education, Obama had more optimistic tone

number of each sentiment type of each topic

Legend: democracy, economy, re-election, military, education, nuclear, health, muslim, crime, peace, intelligence

Topic visualization

Topics

1. Democracy and Human Rights Freedom
2. Economy and job opportunities
3. Obama's second presidential election
4. Military force and threat to Muslims in Iraq, Isil etc.
5. Education and family
6. Nuclear war in Iran
7. Health Care issues
8. Criminal Law and violence
9. Intelligence security and surveillance

Topic1

Topic2

Topic3

Topic4

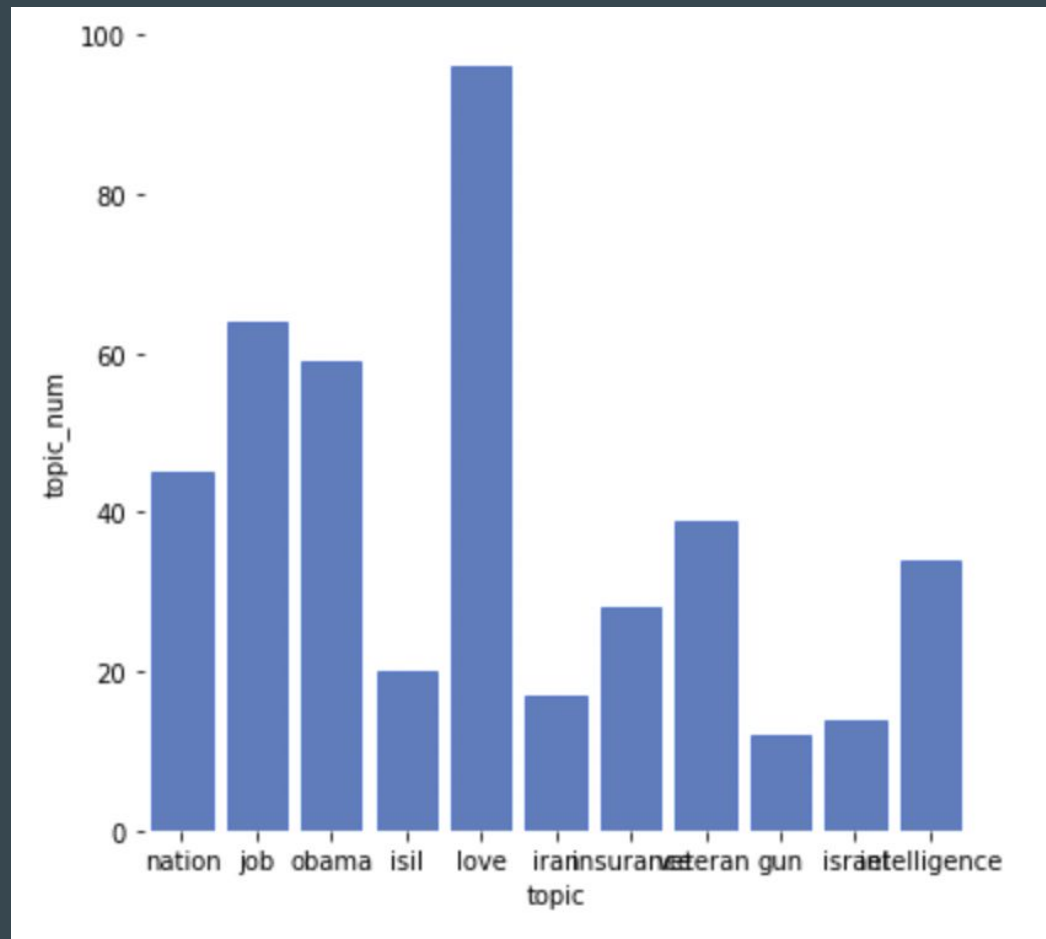Topic5

Topic6

Topic7

Topic8
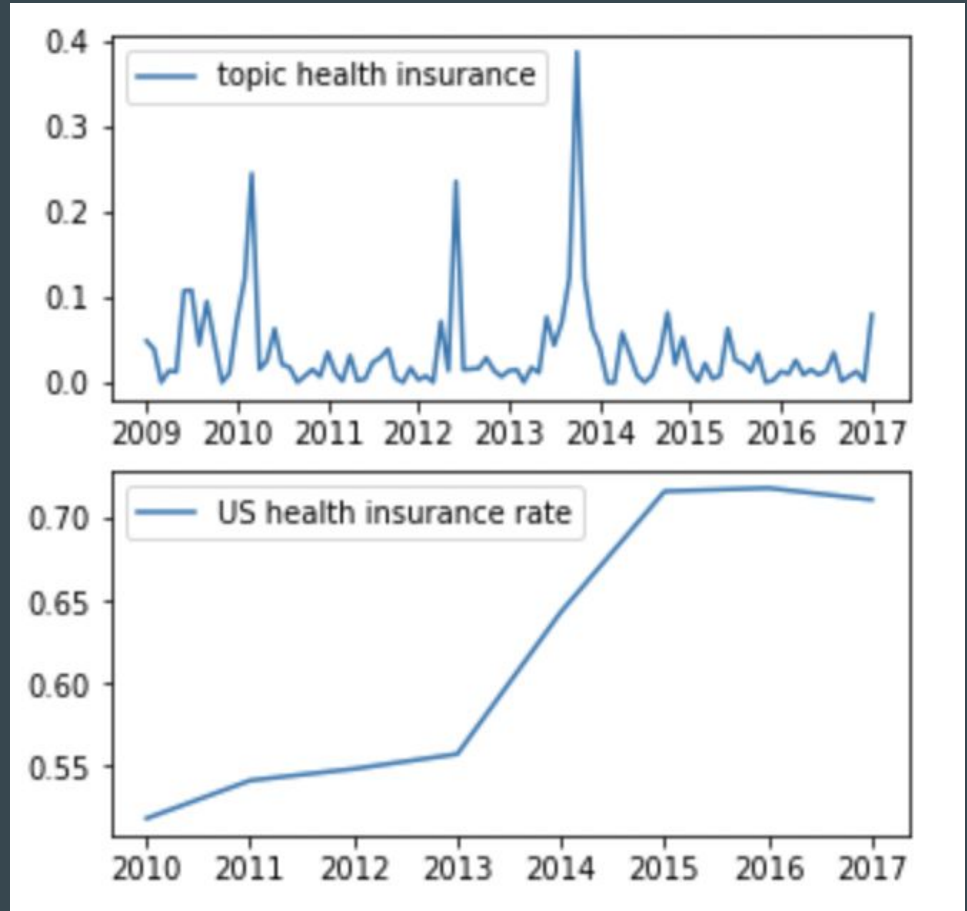
Topic9

Count Topic

Result of NMF

Health Care

Important events happened between 2011-2013:

1. Health Care Reform Bill
2. Affordable Care Act

Obama's speech had a significant influence on promoting health care in the United States, which was fairly reasonable given that health care was his major political accomplishment.
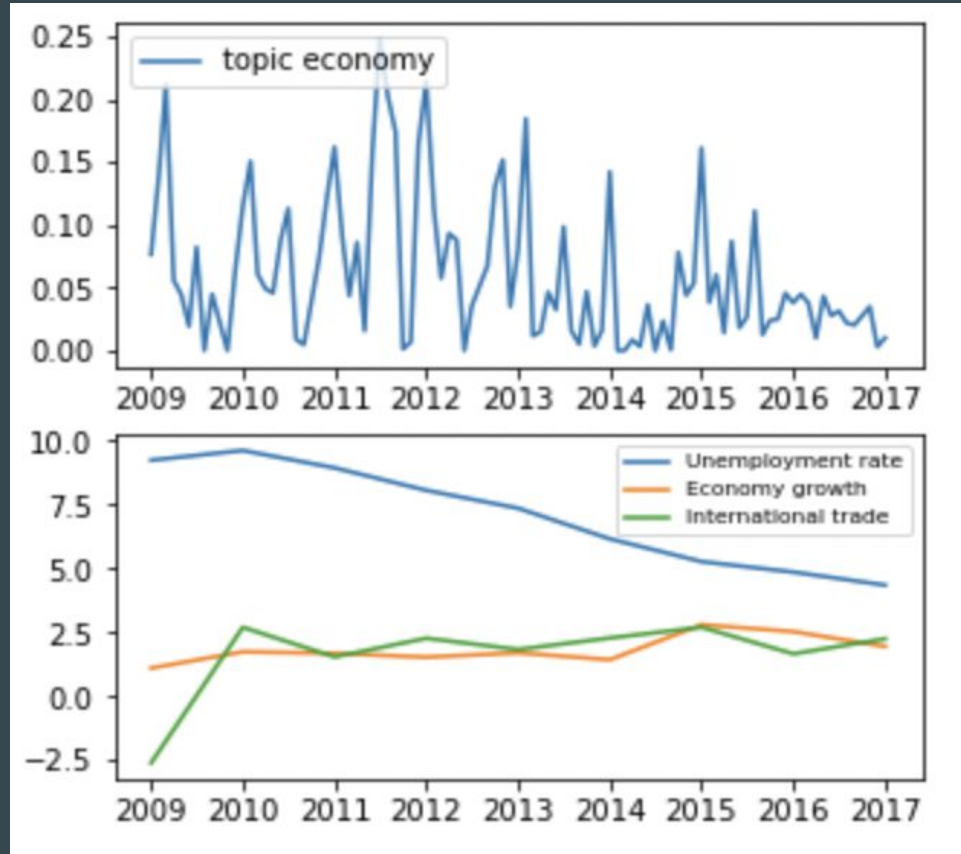
Economy and Employment

Important events happened between 2009 - 2012:

1. American Recovery and Reinvestment
2. American Job Act

Another of Obama's accomplishments is to the U.S. economy and jobs. As a result of his proposals and speeches, the unemployment rate has greatly decreased, and the country's economy and global trade have slowly increased.
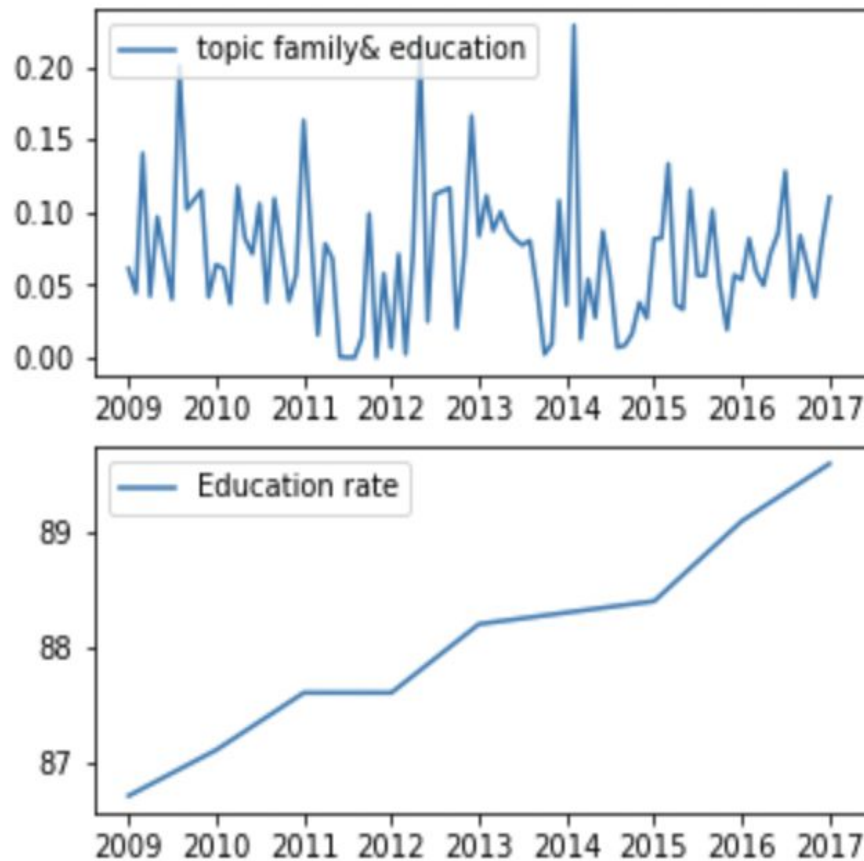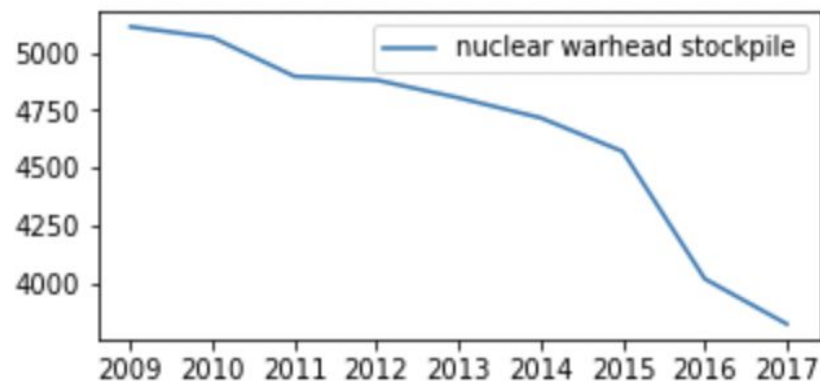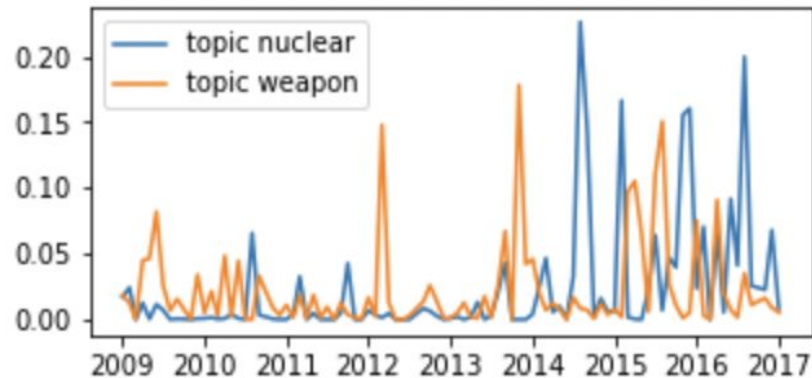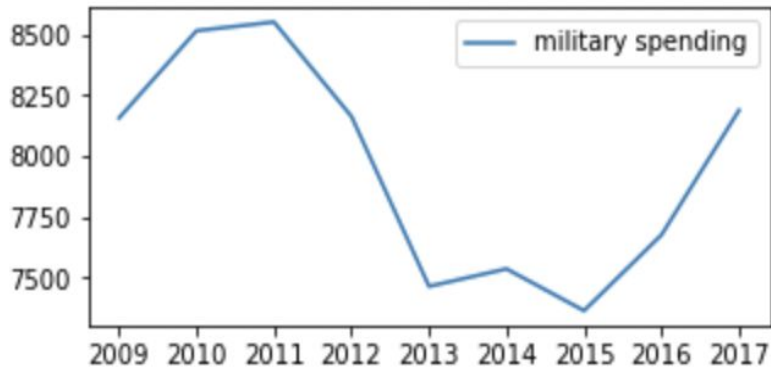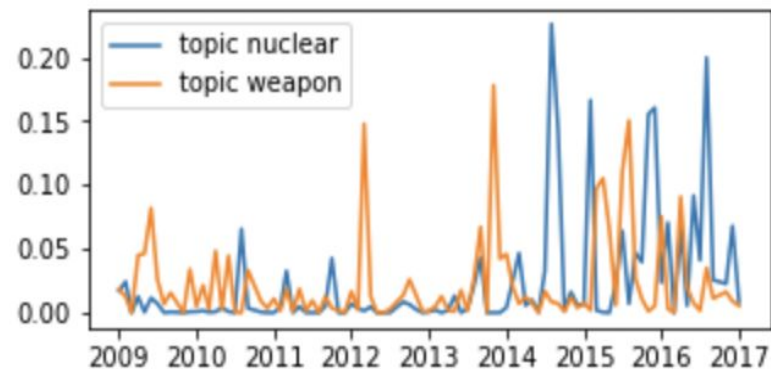
Education

Important events happened between 2009-2014:

1. Veteran education reform
2. No child left behind law

Children in America now have more access to school, and the advancement of American education has been accelerated by Obama's remarks and actions to bring about official education reform.

Military Spending

Conclusion

Summary

- Data preparation: BeautifulSoup, and NLTK
- Topic Modelling: Gensim LDA, Sk-Learn LDA, K-Means, and NMF
- Sentimental Analysis: Textblob, and SentimentIntensitAnalyzer
- 428 Obama speeches can be classified into approximately 9 clusters
- Most of his speeches are about democracy, economy, health care
- Obama's speeches are more positive
- External Factors: Military costs and Nuclear safety, Education, Economy and employment, and health care.

Thanks for listening