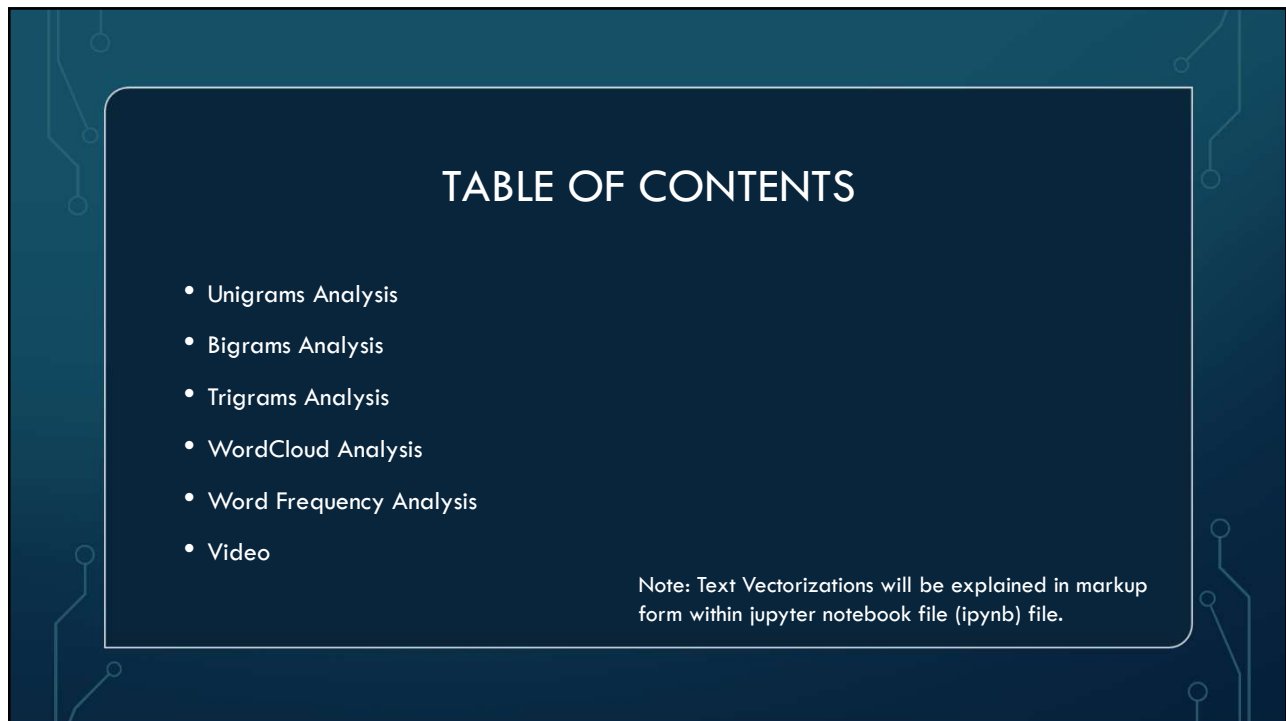


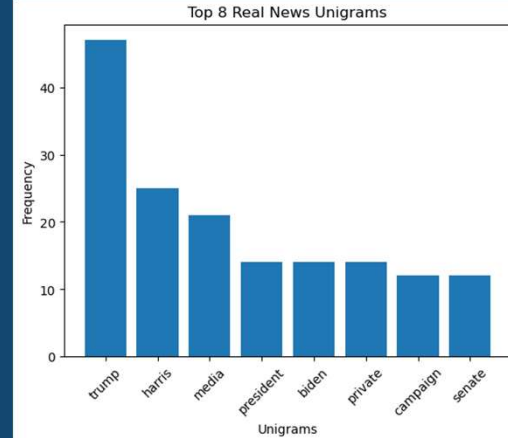
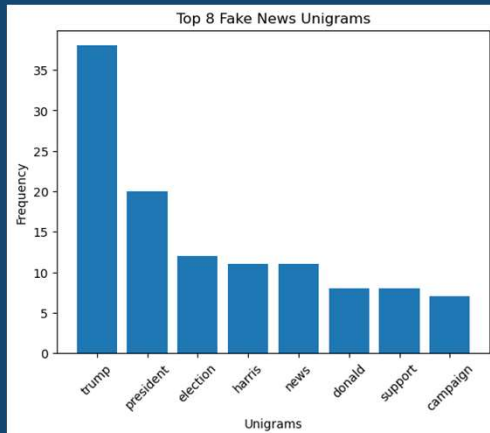


1



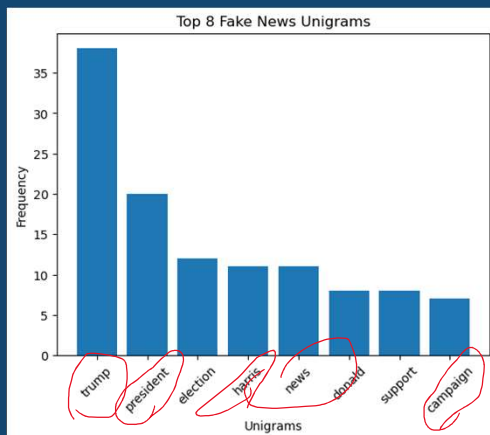
2

## UNIGRAMS

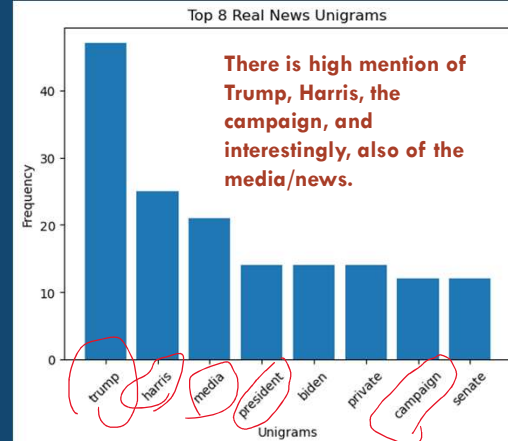


3

## UNIGRAMS

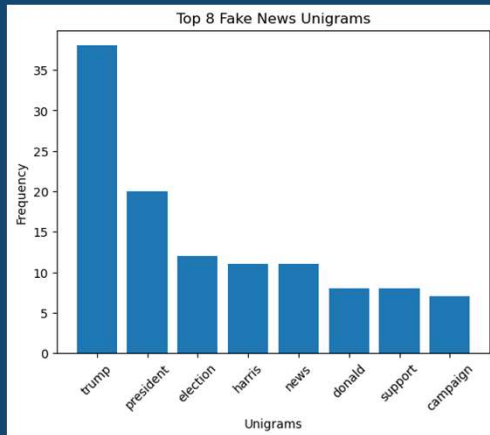


## Similarities

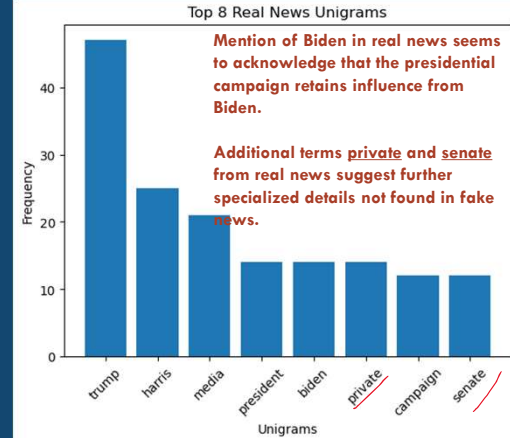


4

## UNIGRAMS

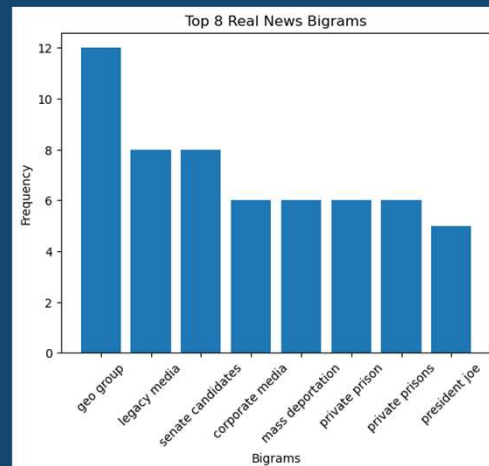
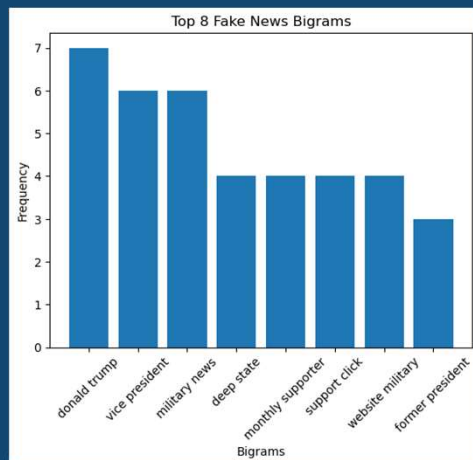


## Differences



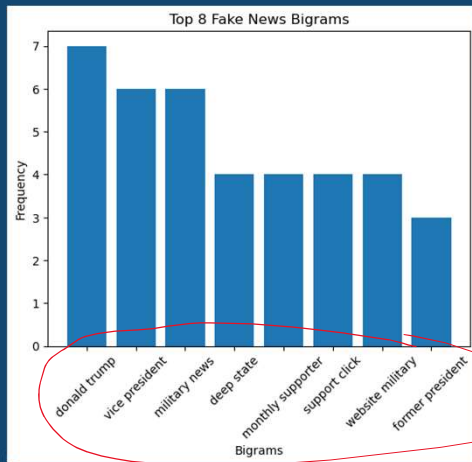
5

## BIGRAMS

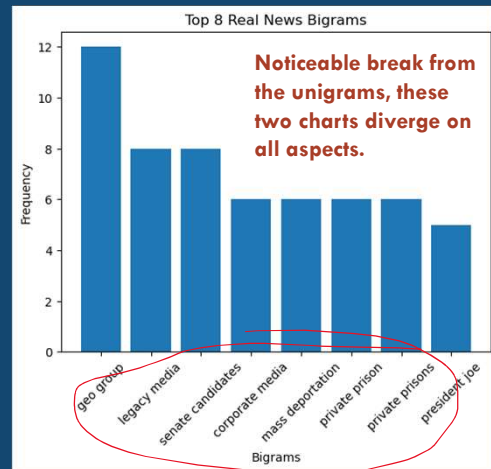


6

## BIGRAMS

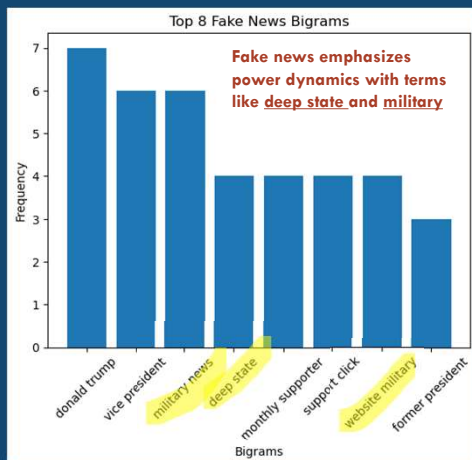


## Similarities

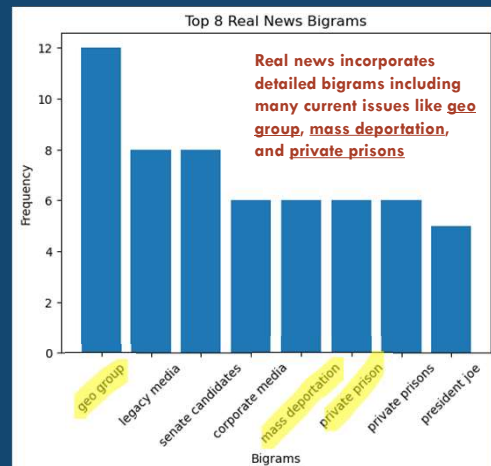


7

## BIGRAMS

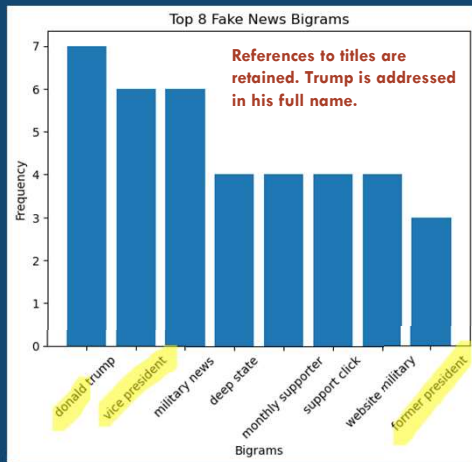


## Differences

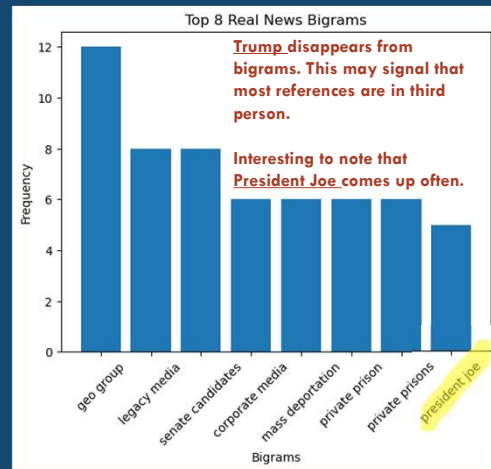


8

## BIGRAMS

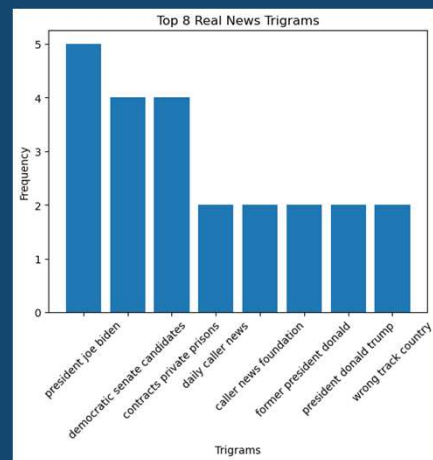
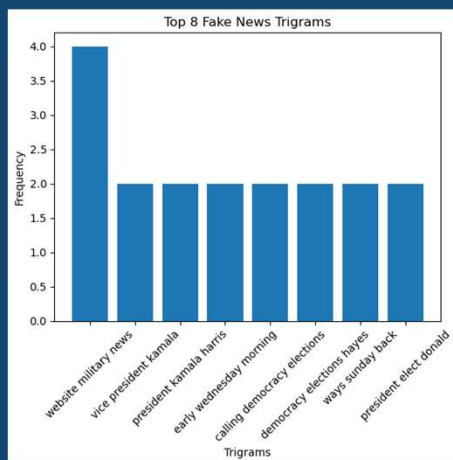


## Differences



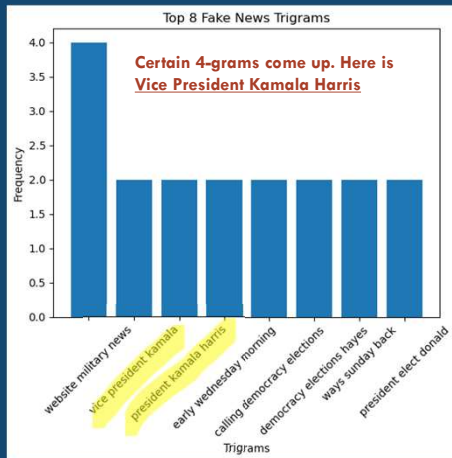
9

## TRIGRAMS

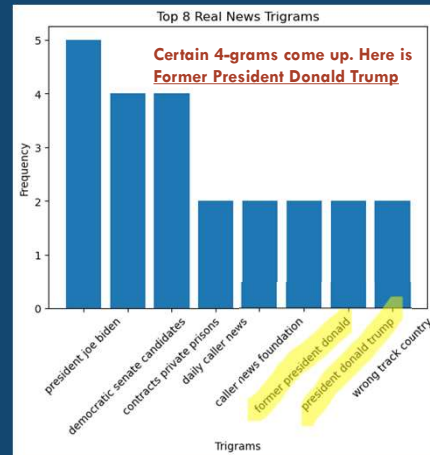


10

## TRIGRAMS

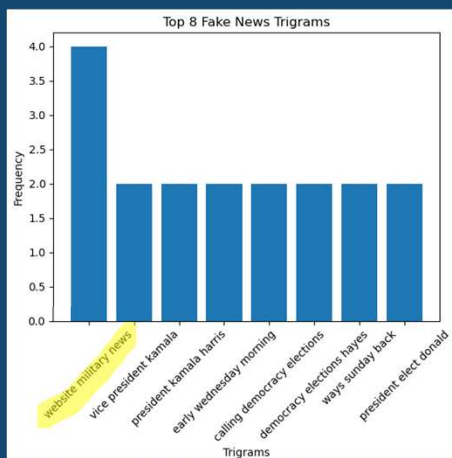


## Similarities

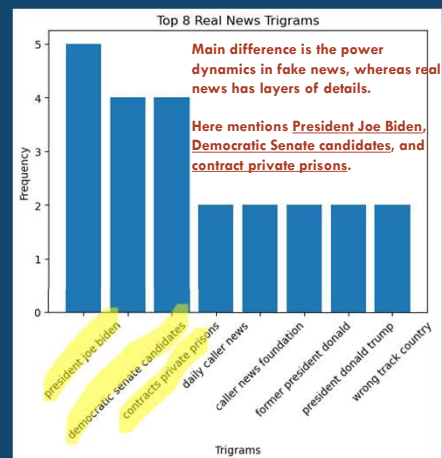


11

## TRIGRAMS



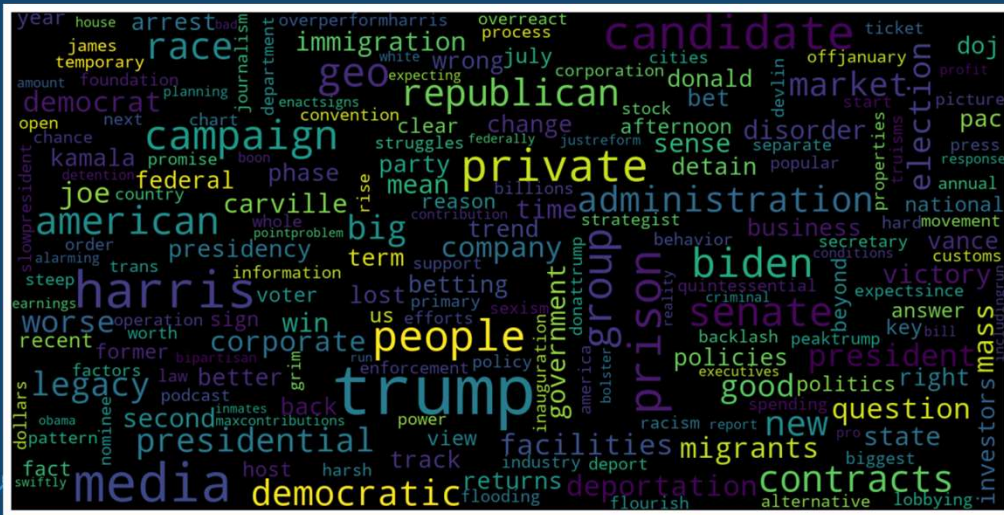
## Differences



12

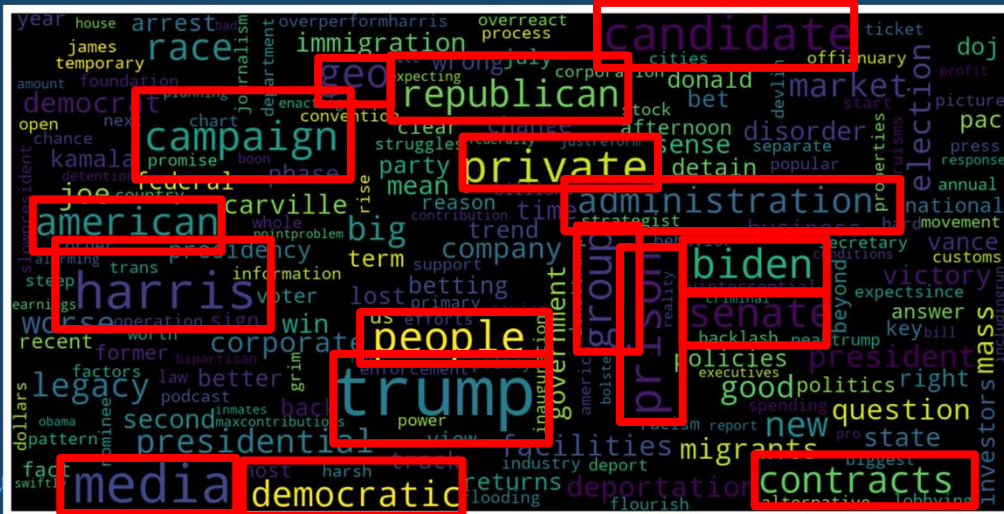


## WORDCLOUD – REAL NEWS



13

## WORDCLOUD – REAL NEWS



Notable words include:


Trump  
Prison  
Harris  
Candidate  
Biden  
Republican  
Campaign  
Democratic  
Media  
Contracts  
Senate  
Geo  
Group  
American  
People

14

# WORDCLOUD – FAKE NEWS

Notable words include:

- Trump
- President
- Election
- Campaign
- Harris
- Donald
- Vice
- State
- New
- Voters
- Support
- US



expected kill spoke free harris website issue commentelon fight documents overreach vote democrac race monthly campaign victory country department earlier popular republicans strong difference video kamala show big wind american institutions weeks federal rated celebrities paypal senate facebook new america support justice corrected zuckerberg media anti senior point staffnews click among former peaceful doj children holding work final presidency early billionaire office men advertisers speech state vice president election support november first transfer win changer power leader people mag human matter nation california years lastjack fall bathrooms majorityelon others black legal ahead chuck a civil united lost policies presidential candidate control interview voters georgia lost

8



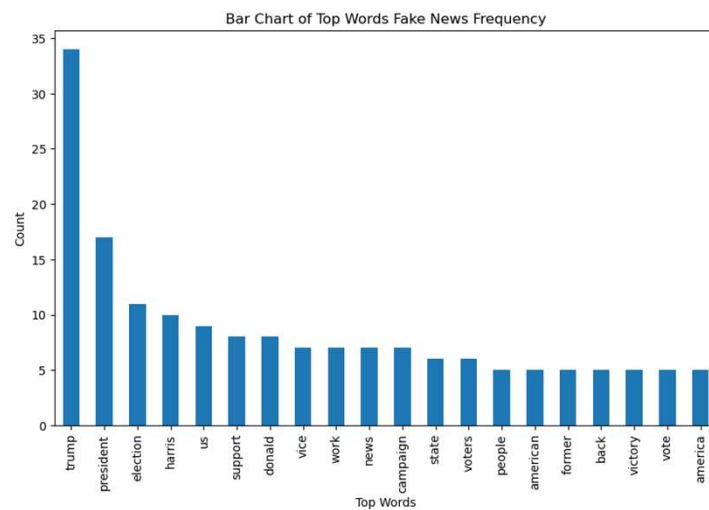
## WORDCLOUD

### Noticeable Differences

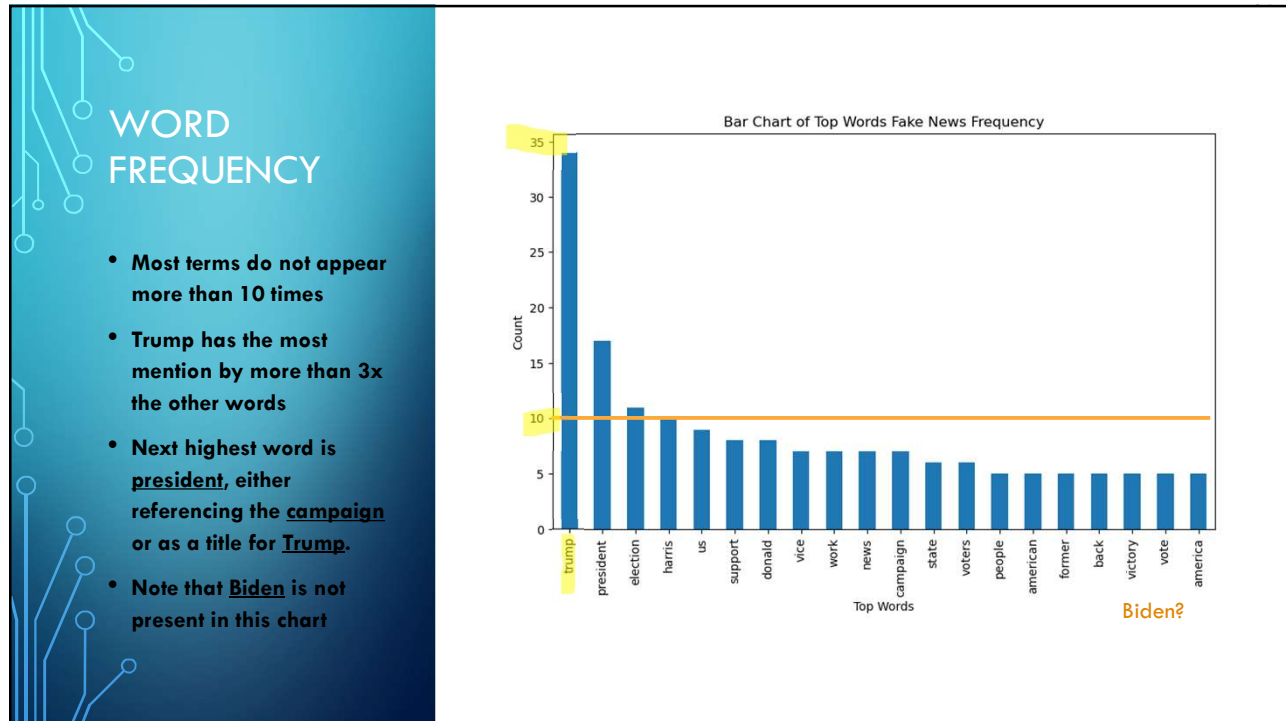
- Harris has been de-emphasized in Fake News
- Trump retains similar or equal emphasis in both Fake News and Real News
- Less words are as pronounced in Fake News than Real News
- Both news highlight major current events like the election campaign
- Real News has more detailed terms like Geo Group, Contracts, and Prison

17

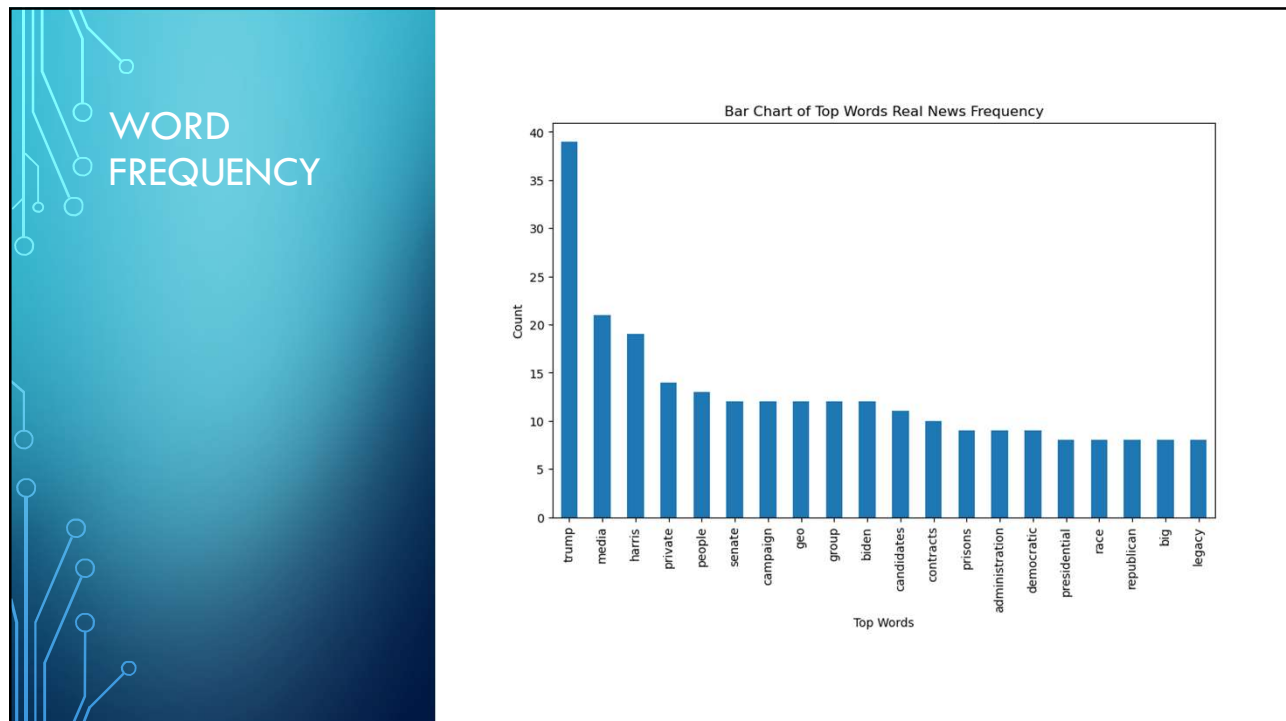
## WORD FREQUENCY



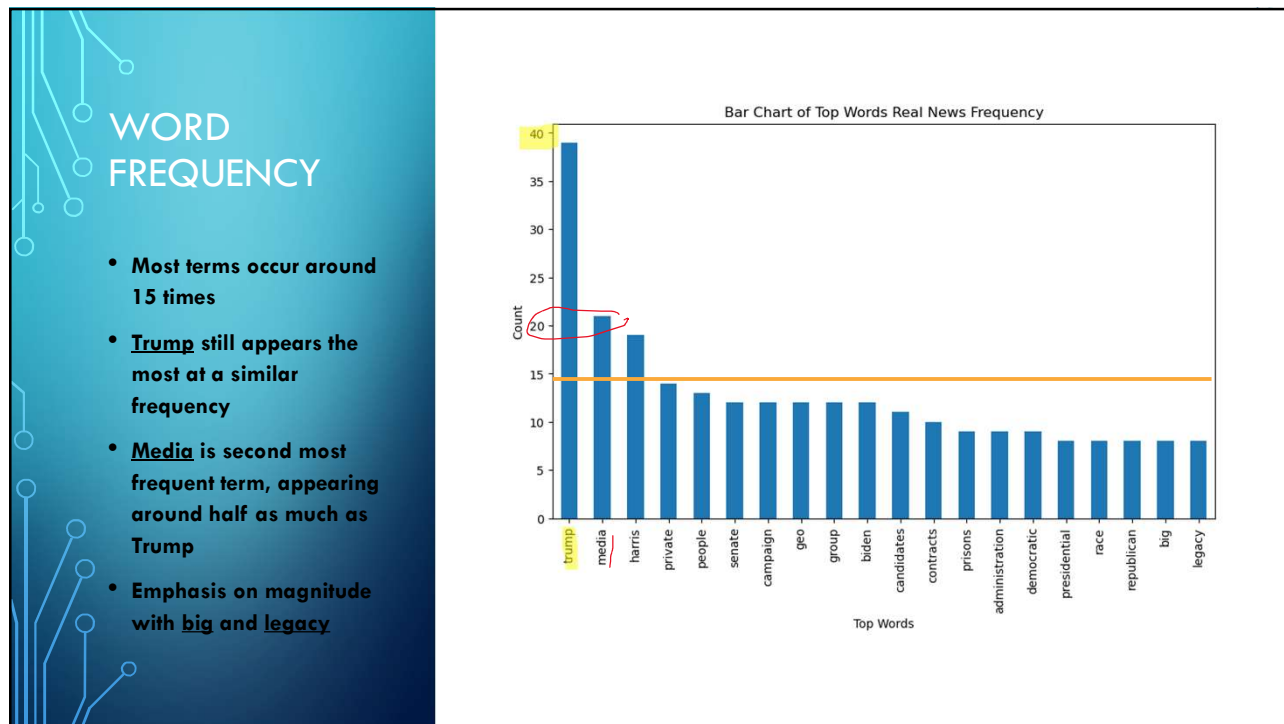
18



19



20




21

## SUMMARY OF FINDINGS

	Unigrams	Bigrams	Trigrams	WordCloud	Frequency Chart
Fake News	Nothing outside primary election	Mentions power dynamics 'military', 'deep state'. Emphasizes titles like 'vice president'	Mentions 'website military news'	Fewer words stand out. Little mention of terms outside of campaign-election.	Sole emphasis on Trump. No mention of issues. No mention of Biden.
Real News	Mentions 'Biden', 'senate', 'private'	Issues like 'mass deportation', 'private prison', 'mass deportation'	Mentions 'President Joe Biden', 'democratic senate candidates', 'contracts private prison'	More words stand out. More mention of key issues.	Emphasis on 'media' and 'harris'. Other terms are more pronounced.
Both	Mention of election and campaign	No similarities	Address candidates by their current title	'Trump' mentioned at similar frequency	'Trump' is the leading term

22



ANALYSIS		
	Fake News	Real News
Topics	Emphasis on major current events.	Has a good mix of events and issues.
Persons	Only key individuals are mentioned. Big personalities like 'Trump' are emphasized.	Mentions all parties relevant to the issue/news.
Content	Appeal to universal issues like power	Attempts to paint a more cohesive picture of issues at hand
Substance	Little details	A lot of details

23



## IDENTIFYING FAKE NEWS

- Fake News tend to have little substance. This is likely to happen because details provide challenges to its authenticity. As the content becomes more detailed, it gets more susceptible to being fact checked.
- Fake News do not have a variety of topics. Targeting a few key topics with a high audience base is more effective to disseminating false information. Specialized content are unlikely to attract much attention.
- Fake News attempts to sway emotions. While there is an emotional appeal for real news as well, it is not as obvious.
- Fake News attempts to capitalize on major personalities. This is likely because there is strong appeal for the average reader to be curious and interested in famous personalities.

24

## ADDITIONAL LINKS

- [https://scikit-learn.org/1.5/modules/generated/sklearn.feature\\_extraction.text.TfidfVectorizer.html](https://scikit-learn.org/1.5/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html)
- <https://kavita-ganesan.com/tfidftransformer-tfidfvectorizer-usage-differences/>
- <https://stackoverflow.com/questions/17839973/constructing-dataframe-from-values-in-variables-yields-valueerror-if-using-all>
- <https://scikit-learn.org/dev/modules/generated/sklearn.neighbors.KNeighborsClassifier.html#sklearn.neighbors.KNeighborsClassifier>
- <https://stackoverflow.com/questions/74406936/knn-python-implementation>