Hari's Hooligans

Sentiment Analysis on Articles About the 2024 Election

A. Chyzh, L. Colak, and E. Shin
H. Nathan

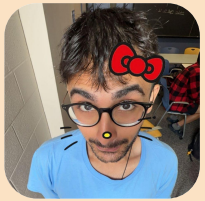
Don't use our slide show
Or else
We will sue

August 2025

Ezra

Rice lover

Watched youtube for the entire program



Latif

Baklava lover

Took 5 hour lunch breaks

Siia

Vodka lover

Phone addict

The “Don”

Before U of R grad program

After the U of R grad program

Background & Data

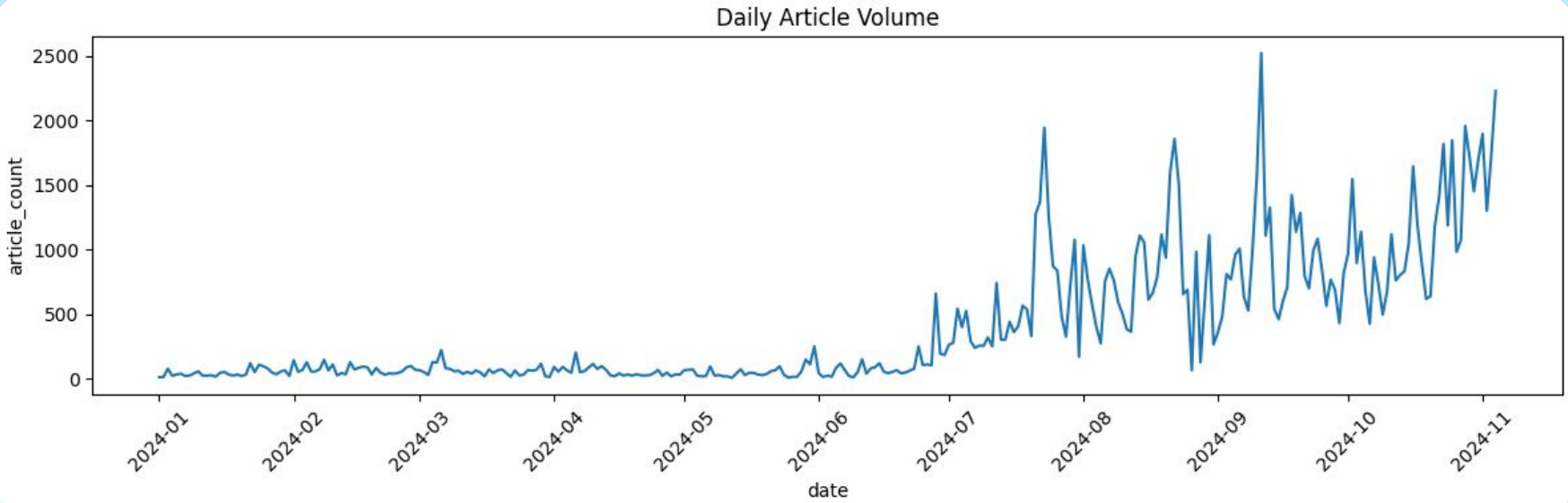
Request → GDELT 2.0 → Article URLs → 

→ Article Text

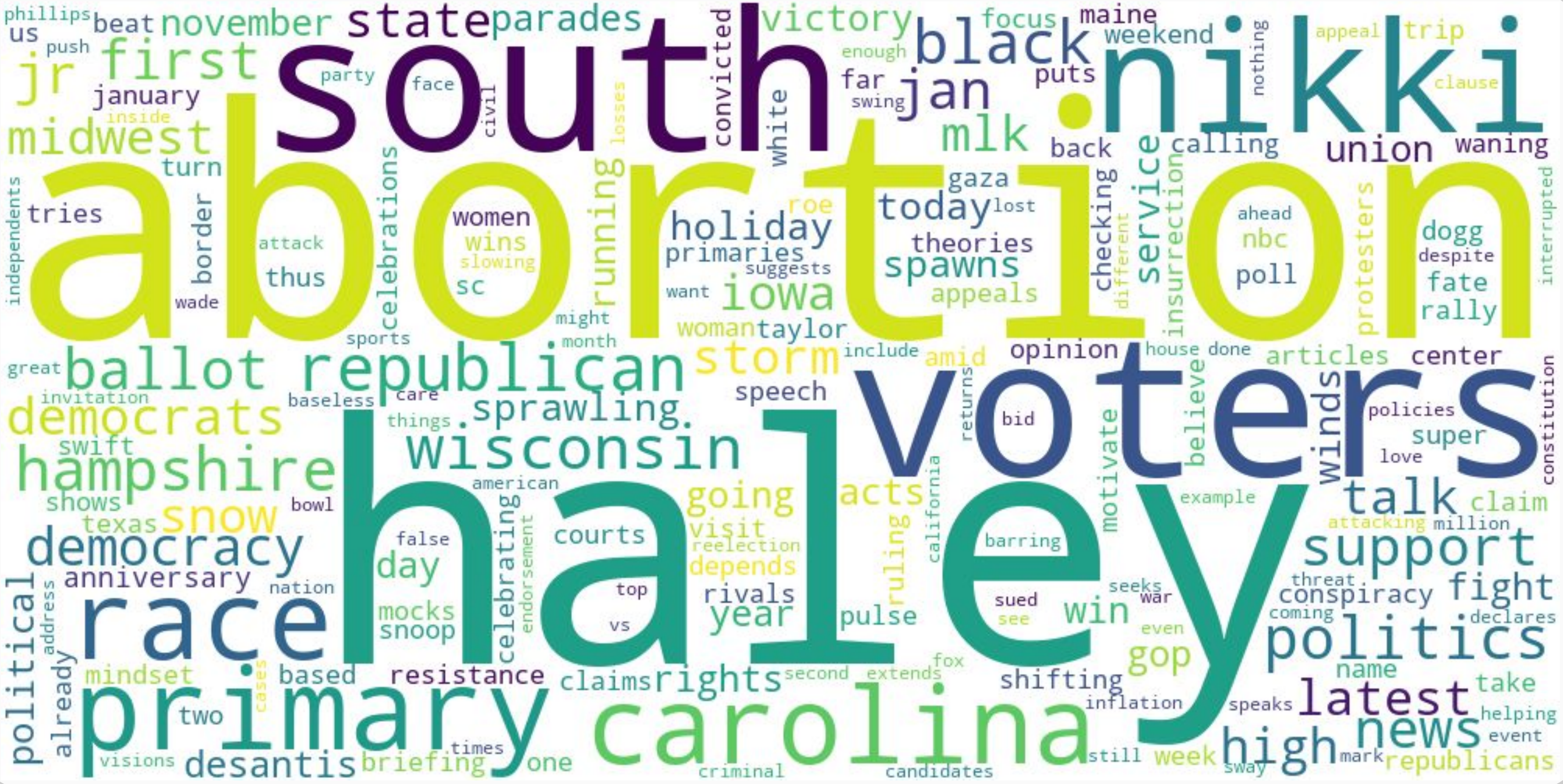
```
params = {
    "query": "harris AND trump AND biden AND sourcelang:english AND theme:ELECTION AND sourcecountry:US",
```

121,689 Articles from between 2024-01-01 to 2024-11-04
From 3225 different news sources

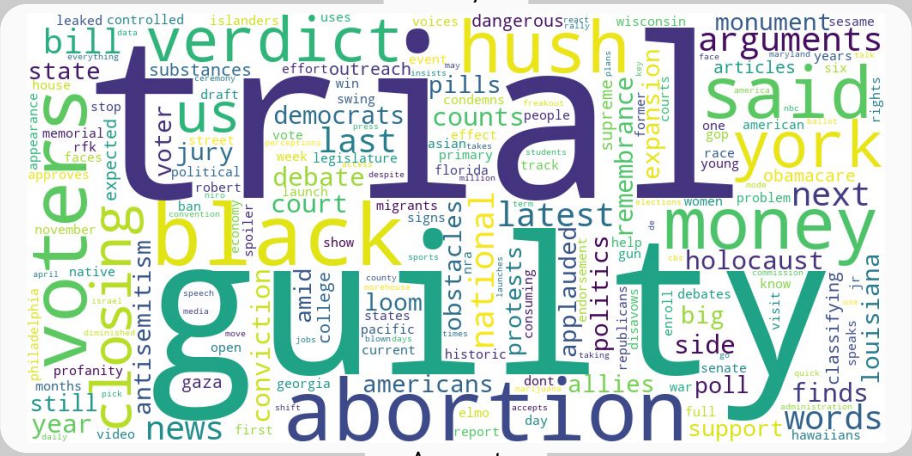
Background & Data



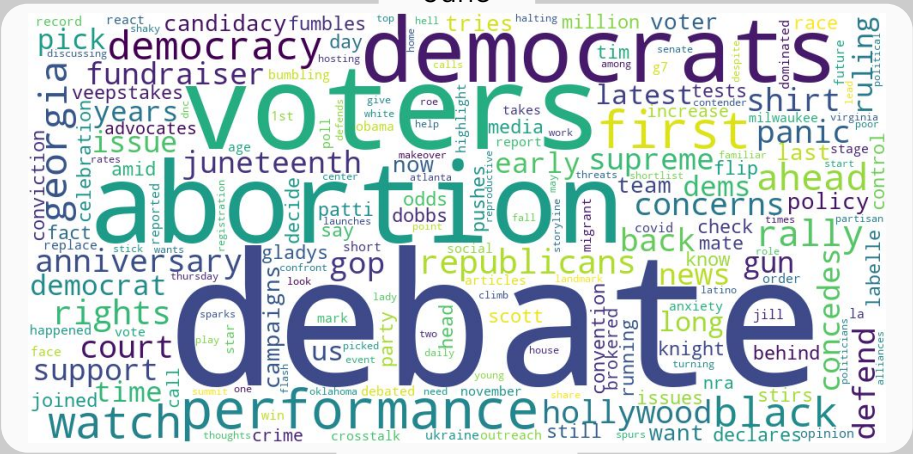
January


abortion
haley
voters
south
nikki
carolina
primary
race
republican
wisconsin
hampshire
ballot
democrats
democracy
black
state
victory
first
midwest
jr
november
parades
jan
mlk
iowa
holiday
spawns
storm
support
politics
latest
news
high
talk
fight
conspiracy
union
waning
calling
back
puts
weekend
maine
focus
trip
appeal
clause
nothing
enough
far
swing
convicted
white
running
service
insurrection
checking
protesters
interrupted
dogg
despite
fate
rally
poll
nbc
ahead
today
theories
gaza
lost
roe
primaries
suggests
want
appeals
woman
taylor
include
amid
opinion
house
done
articles
center
women
wins
slowing
thus
sc
attack
celebrations
border
turn
tries
wade
independents
january
inside
party
face
push
us
beat
phillips
civil
losses
might
month
sports
great
invitation
baseless
care
things
sprawling
speech
swift
shows
texas
snow
anniversary
political
address
already
mindset
based
resistance
two
desantis
briefing
one
times
visions
cases
nation
top
vs
false
bowl
american
celebrating
endorsement
courts
going
visit
reelection
depends
acts
day
mocks
snoop
claims
rights
year
rivals
pulse
ruling
california
win
gop
shifting
motivate
believe
winds
returns
bid
policies
super
love
example
claim
attacking
million
barring
threat
declares
coming
seeks
war
sued
see
even
fox
second
extends
inflation
name
take
speaks
helping
event
mark
sway
still
week
republicans
candidates
criminal
constitution
different

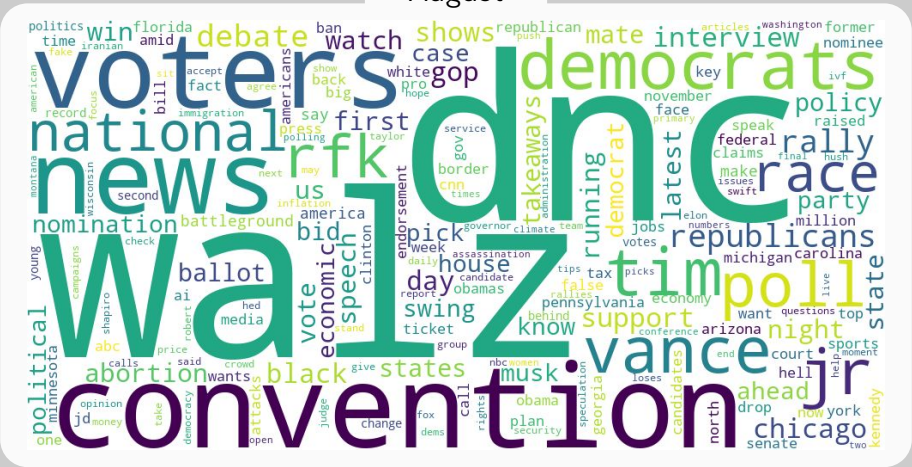
May



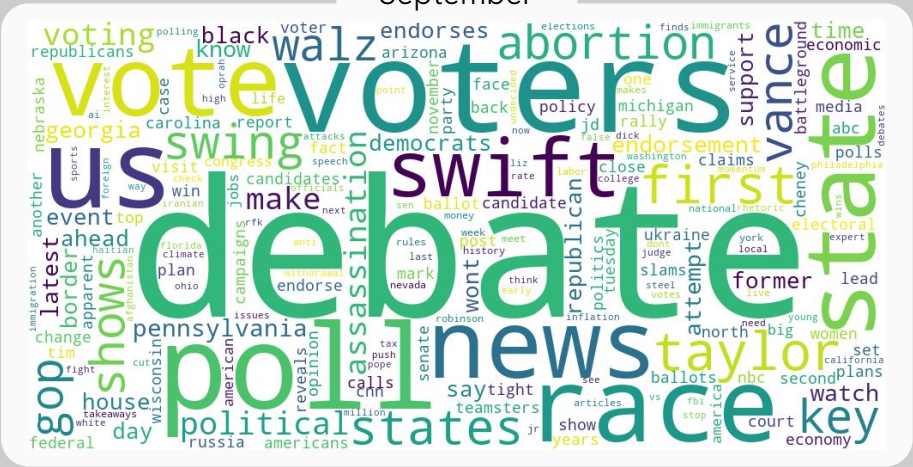
June



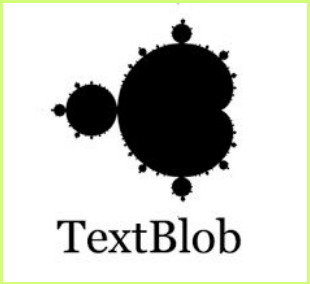
August



September



Sentiment Analysis



Polarity: how positive or negative a piece of text is

Averages individual word polarities (word-by-word analysis)

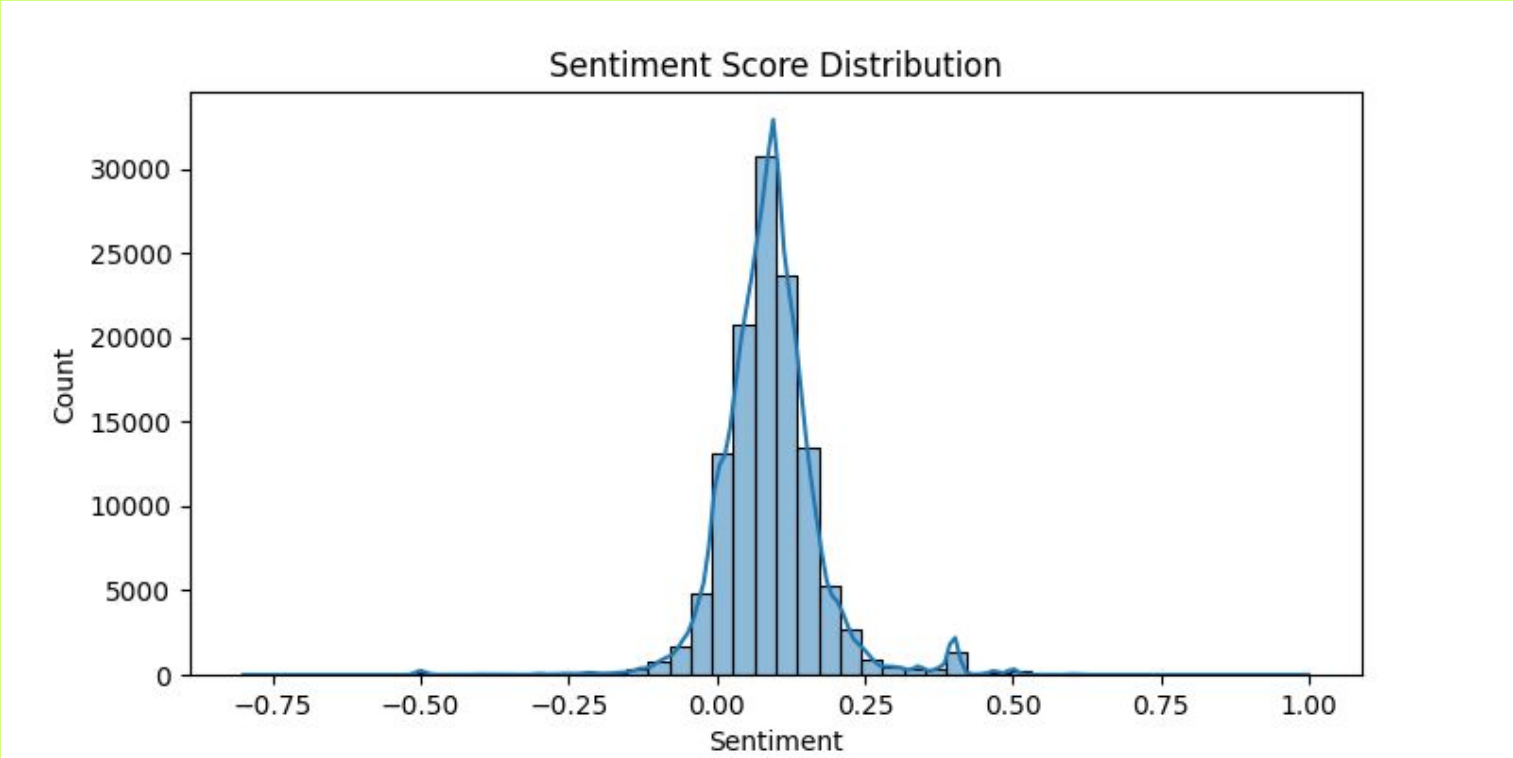
Bias Tagging:

```
df_day['bias'] = "Neutral"

df_day.loc[df_day['more dem'] & (df_day['sentiment'] > 0), 'bias'] = "Pro-Democrat"
df_day.loc[df_day['more rep'] & (df_day['sentiment'] > 0), 'bias'] = "Pro-Republican"
df_day.loc[df_day['more dem'] & (df_day['sentiment'] < 0), 'bias'] = "Anti-Democrat"
df_day.loc[df_day['more rep'] & (df_day['sentiment'] < 0), 'bias'] = "Anti-Republican"
```

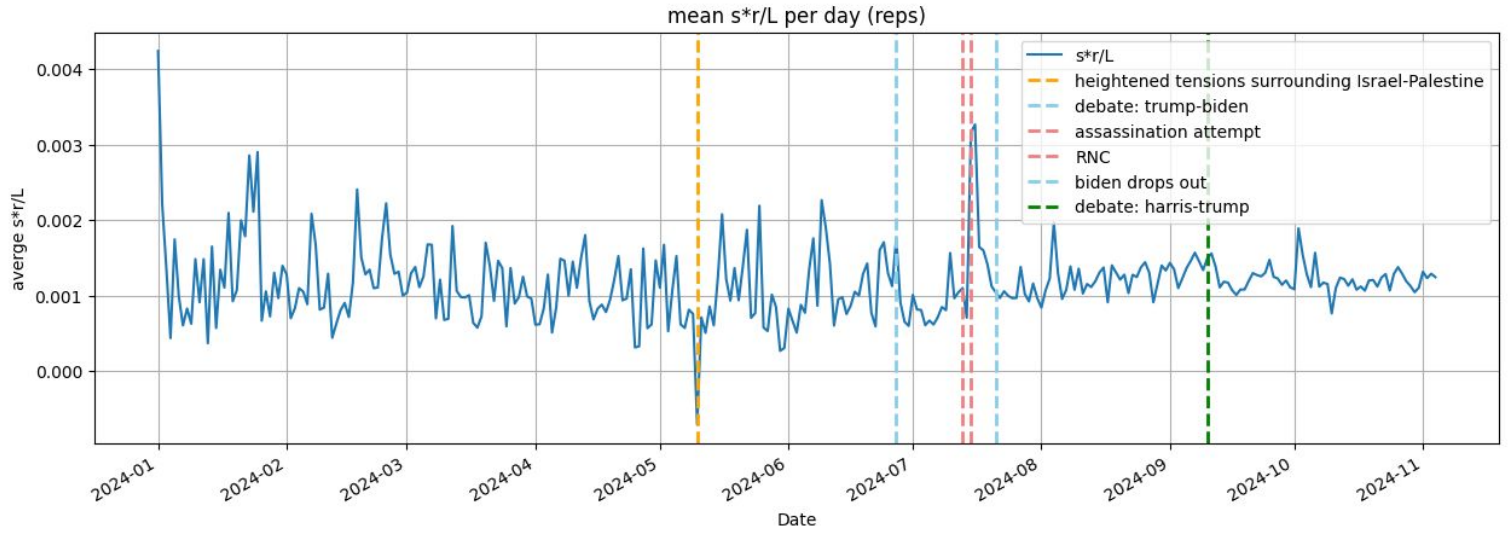
```
republican_keywords = ["Trump", "Vance", "GOP", "Republican", "conservative", "MAGA", "right-wing"]
democrat_keywords = ["Biden","Harris", "Waltz", "Democrat", "liberal", "progressive", "left-wing"]
```

Sentiment Analysis



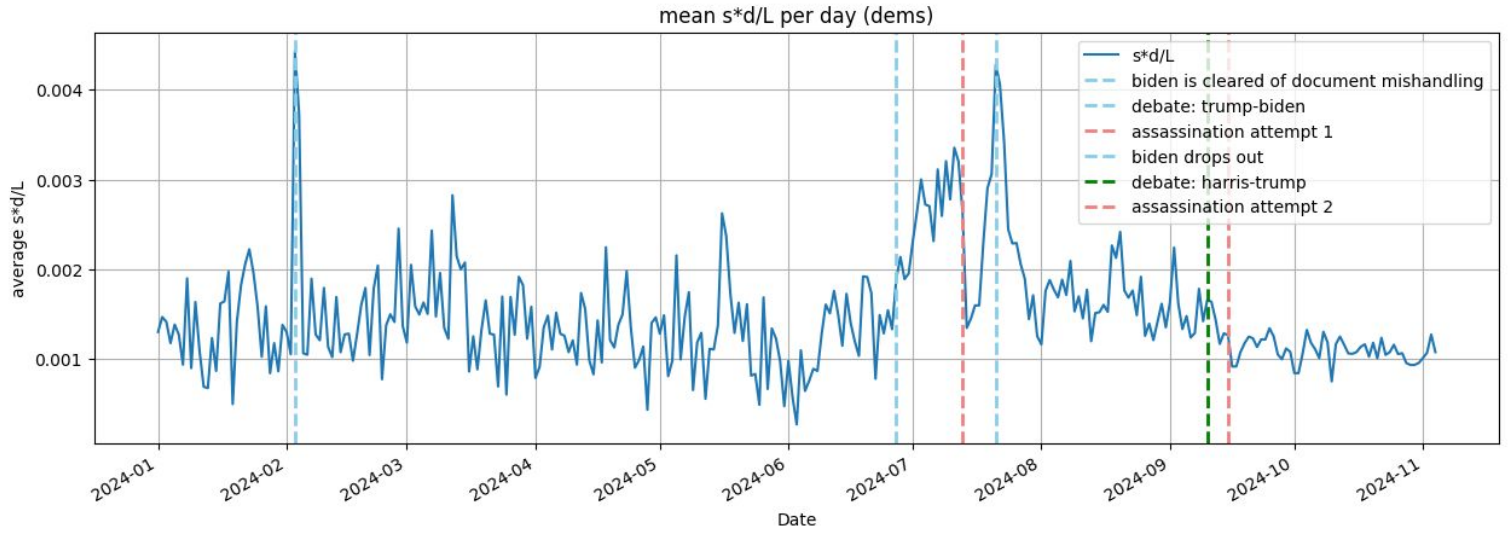
SRDL Scores™

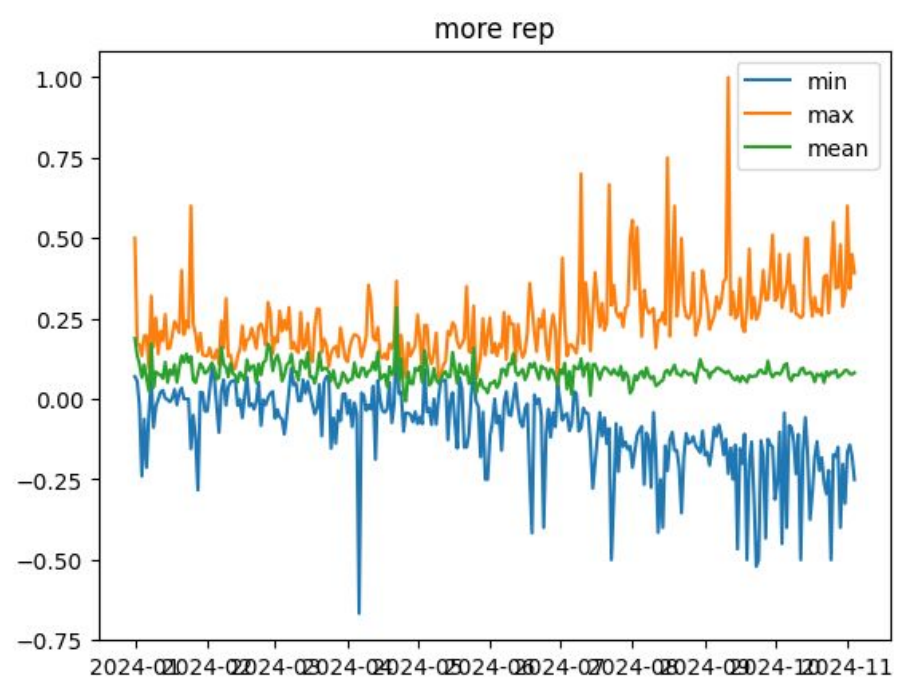
$$(sentiment) \times \frac{(number\ of\ party\ words\ in\ the\ article)}{(length\ of\ the\ article)} = s \times \frac{(r|d)}{L}$$

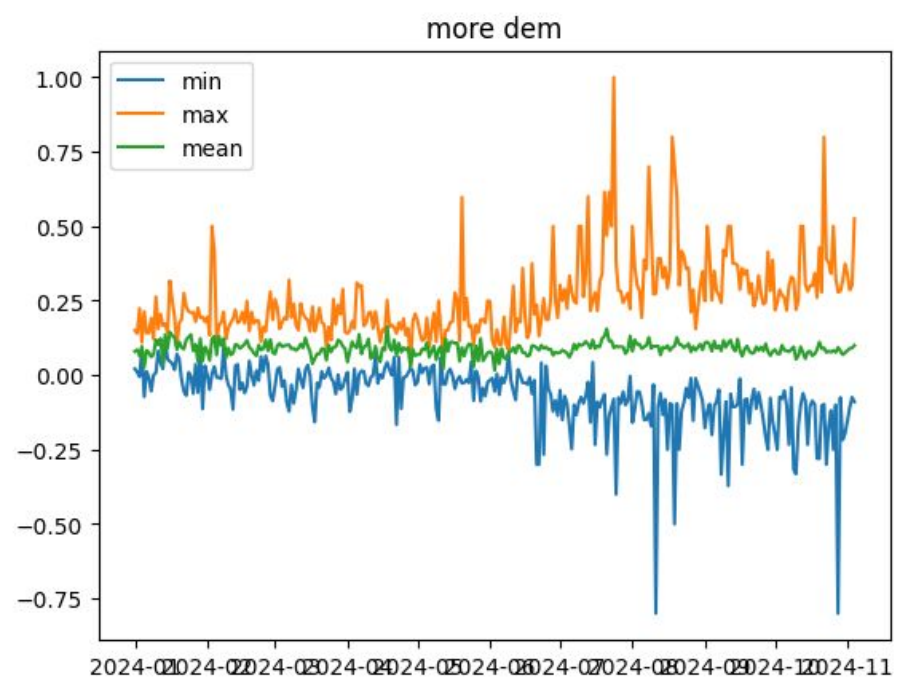


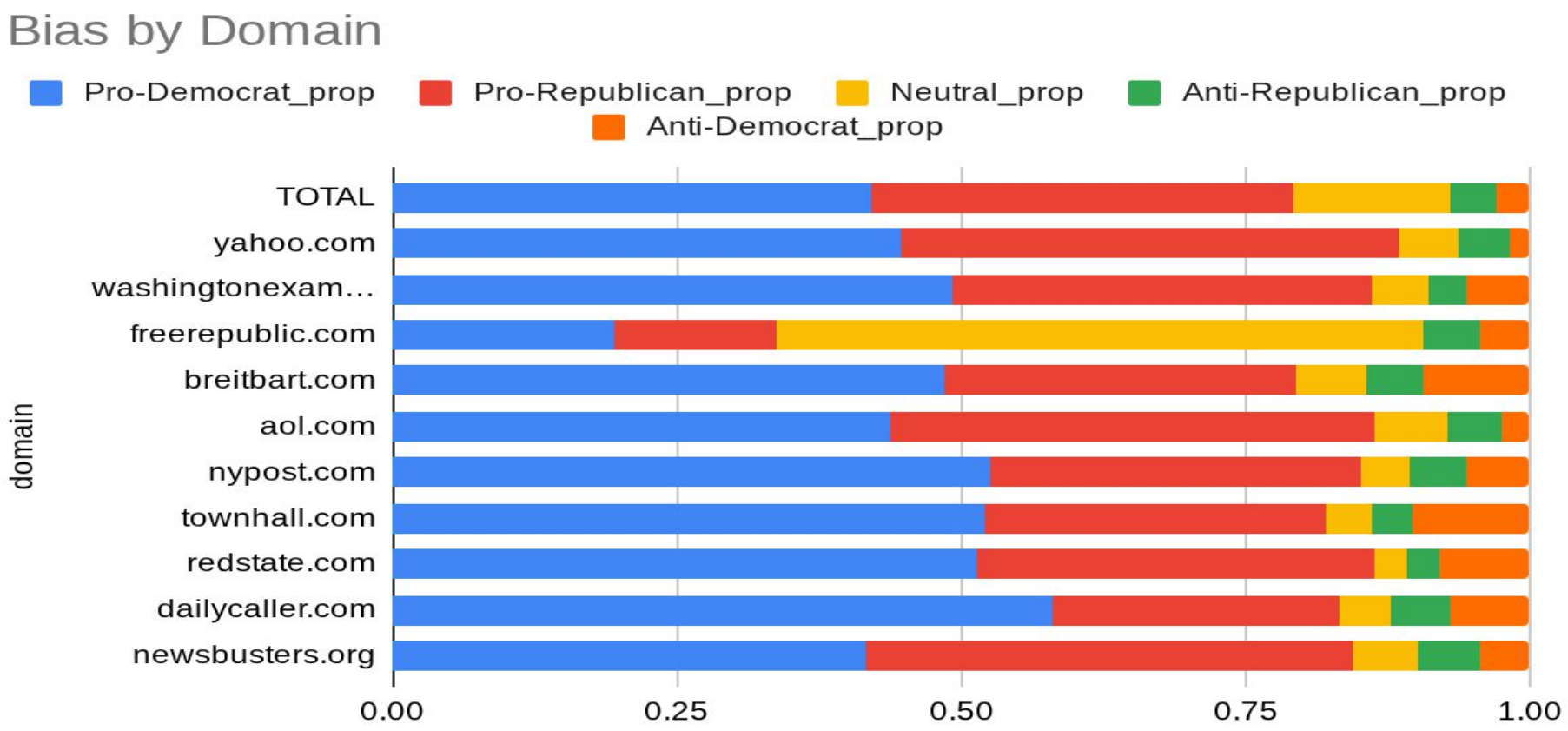
SRDL Scores™

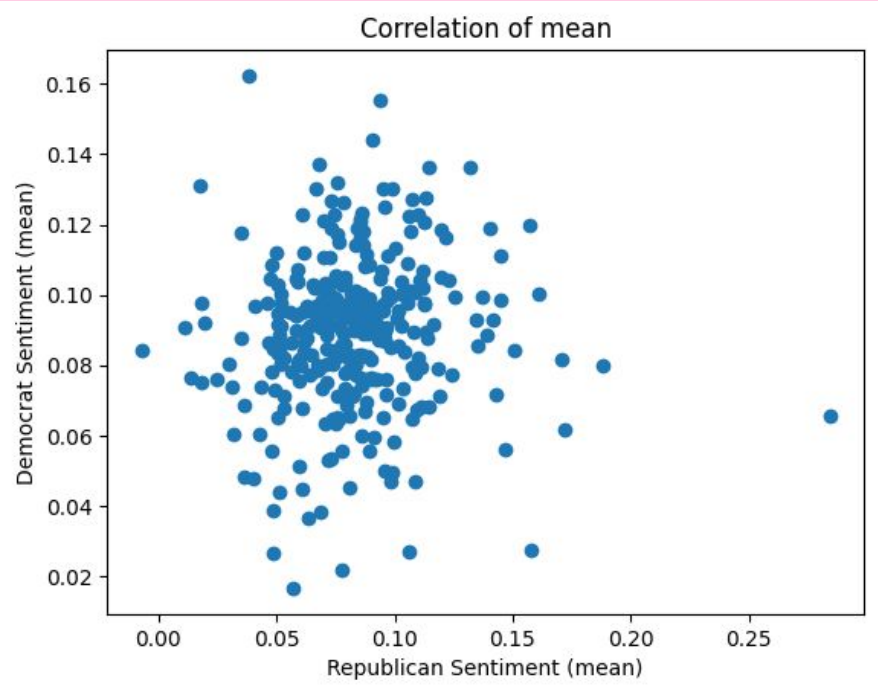
$$(sentiment) \times \frac{(number\ of\ party\ words\ in\ the\ article)}{(length\ of\ the\ article)} = s \times \frac{(r|d)}{L}$$



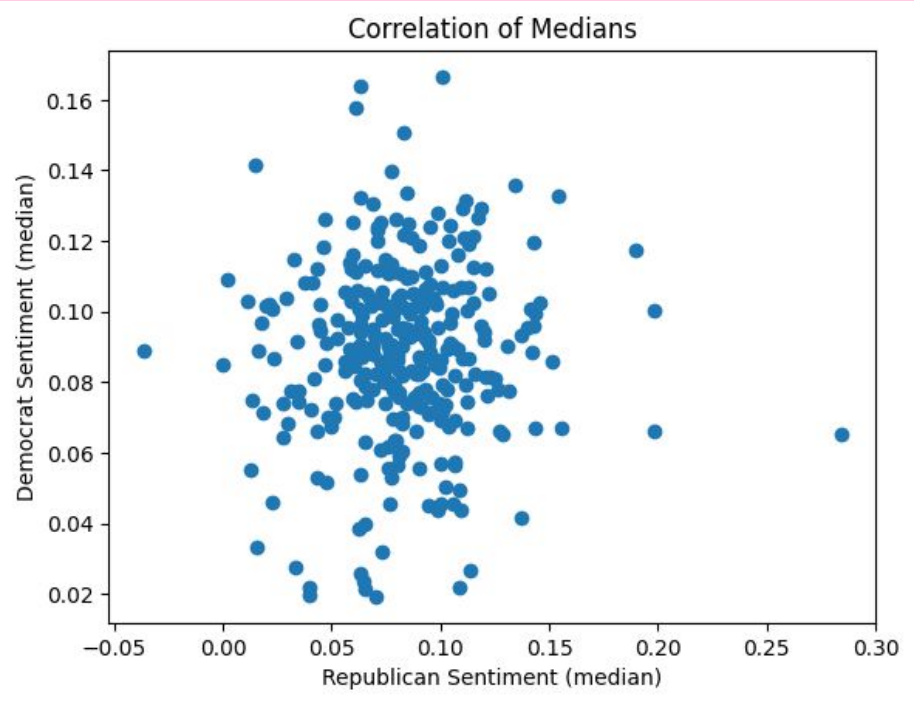

more rep
min
max
mean
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
2024-11


more dem
min
max
mean
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
2024-11

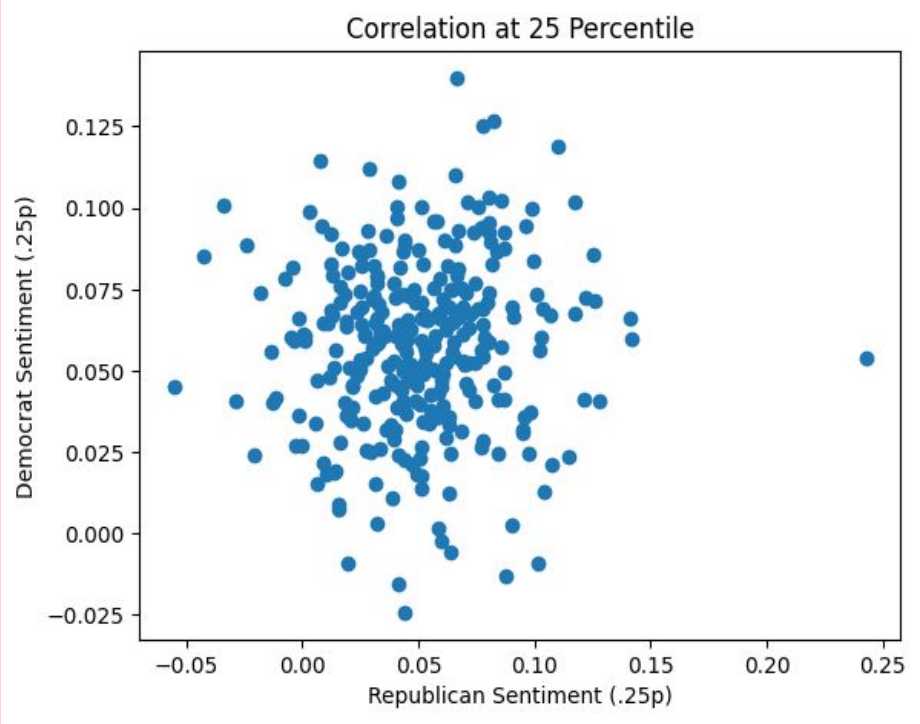

Bias by Domain
Pro-Democrat_prop
Pro-Republican_prop
Neutral_prop
Anti-Republican_prop
Anti-Democrat_prop
domain
TOTAL
yahoo.com
washingtonexam...
freerepublic.com
breitbart.com
aol.com
nypost.com
townhall.com
redstate.com
dailycaller.com
newsbusters.org
0.00
0.25
0.50
0.75
1.00



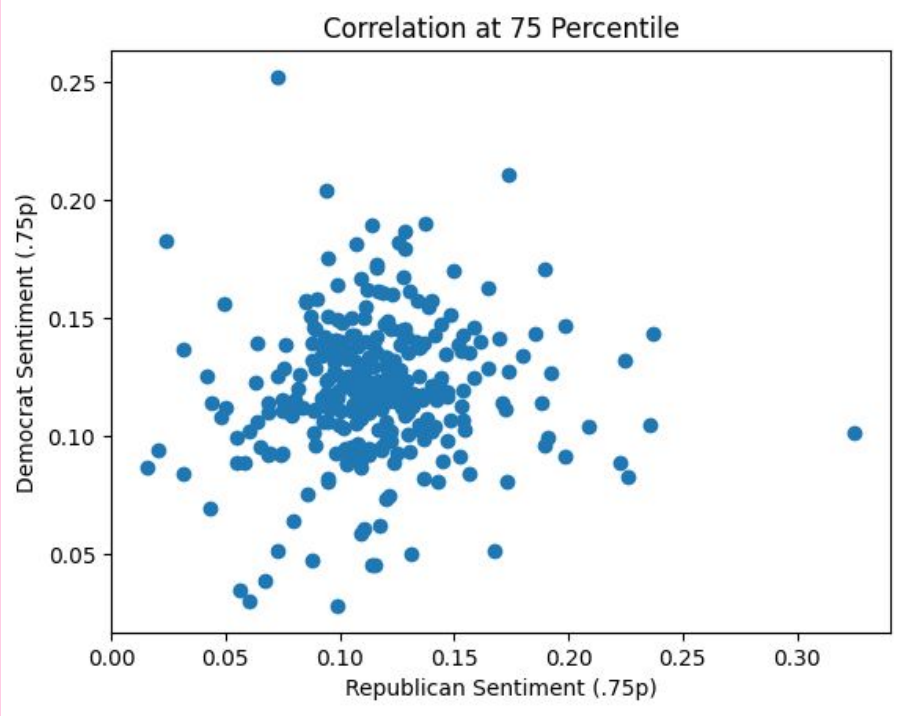
Correlation Coefficient: 0.06



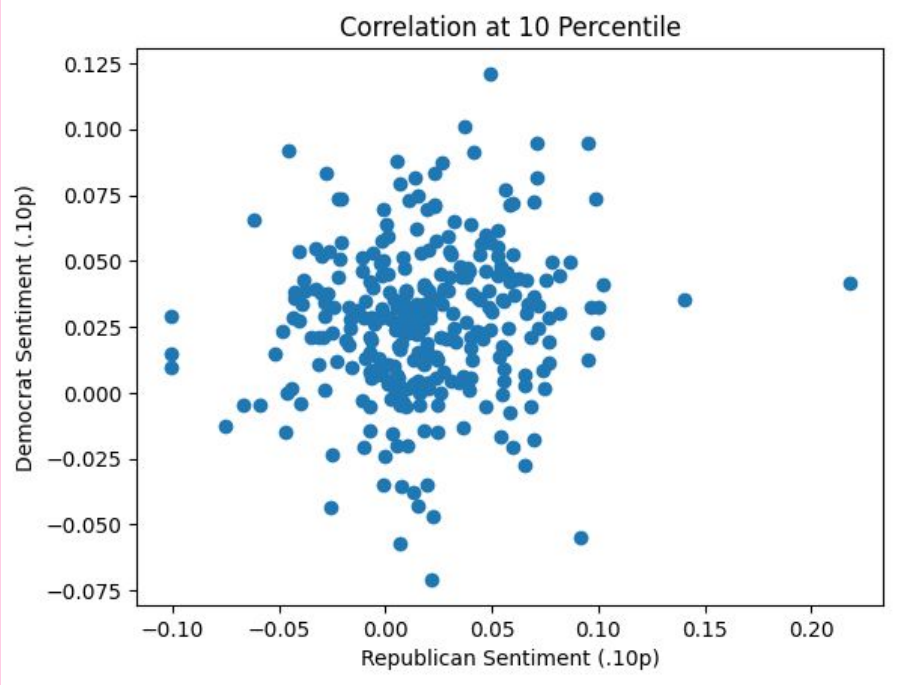
Correlation Coefficient: 0.05



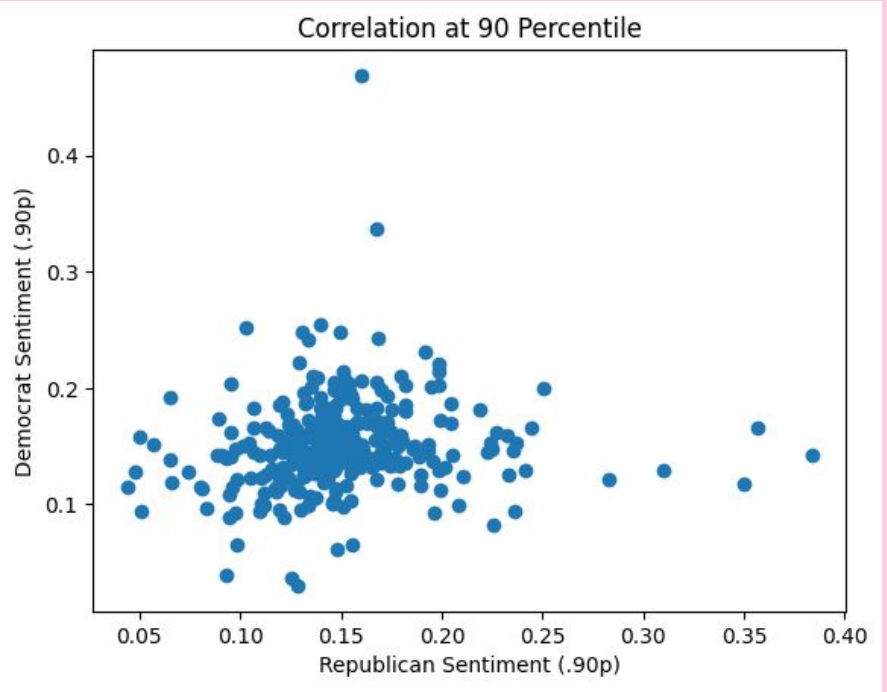
Correlation Coefficient: 0.06



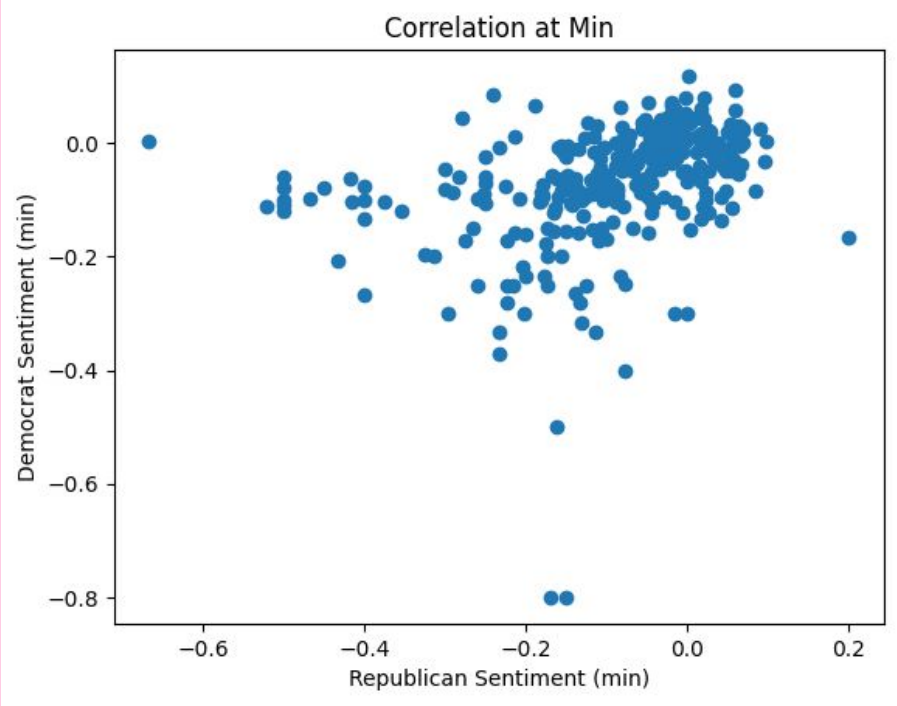
Correlation Coefficient: 0.08



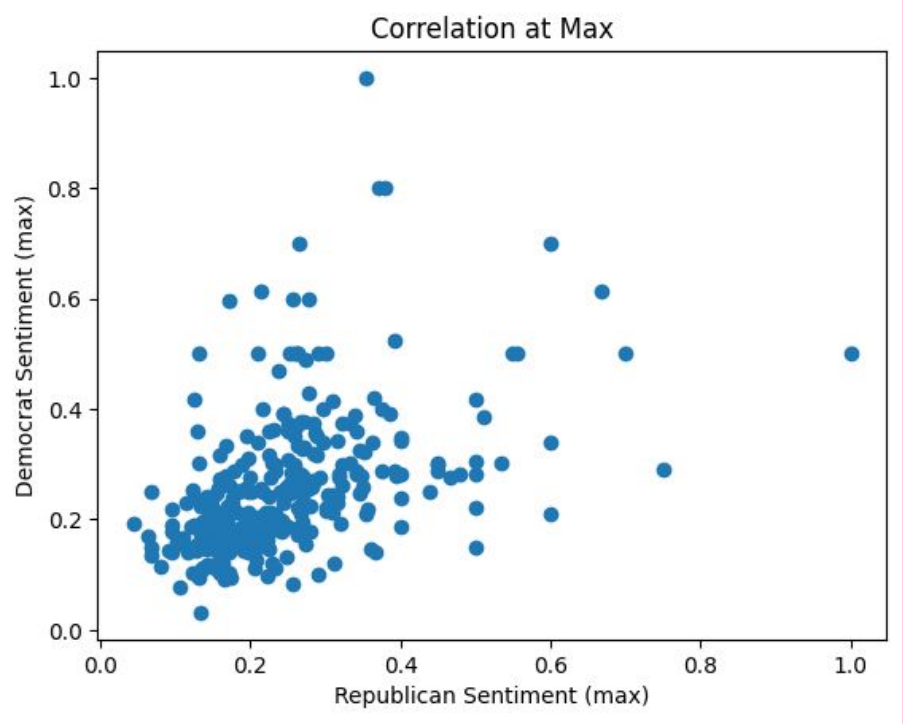
Correlation Coefficient: 0.12



Correlation Coefficient: 0.12



Correlation Coefficient: 0.33



Correlation Coefficient 0.47

Limitations and Issues

Gdelt

Difficult to query

- Slightly mismatched article dates
- Unable to query data before 2017
- Takes a long time

Textblob

Very basic textual analysis

- Limited awareness of context
- No ability to understand sarcasm or irony
- Issues with detecting negative sentiment

SRDL Scores™

Statistically questionable?

- Favors shorter articles
- Unrefined technique for determining which party an article is about
- Uses potentially flawed Textblob sentiment analysis

Us

Python? What's that?

- Dumb, dumber, and dumbest sitting in one room
- Vibe coded our way through this

Conclusion

Our sentiment analysis shows strong alignment with real-world political events.

Moving forward, we will

- expand our dataset,
- calculate additional SRDL scores, and
- explore their potential for predicting election outcomes with greater accuracy.

감사합니다!
Дякую!
Teşekkürler!
THANK YOU!!

Any questions?