



r/CMV

changemyview | group 1

r/changemyview

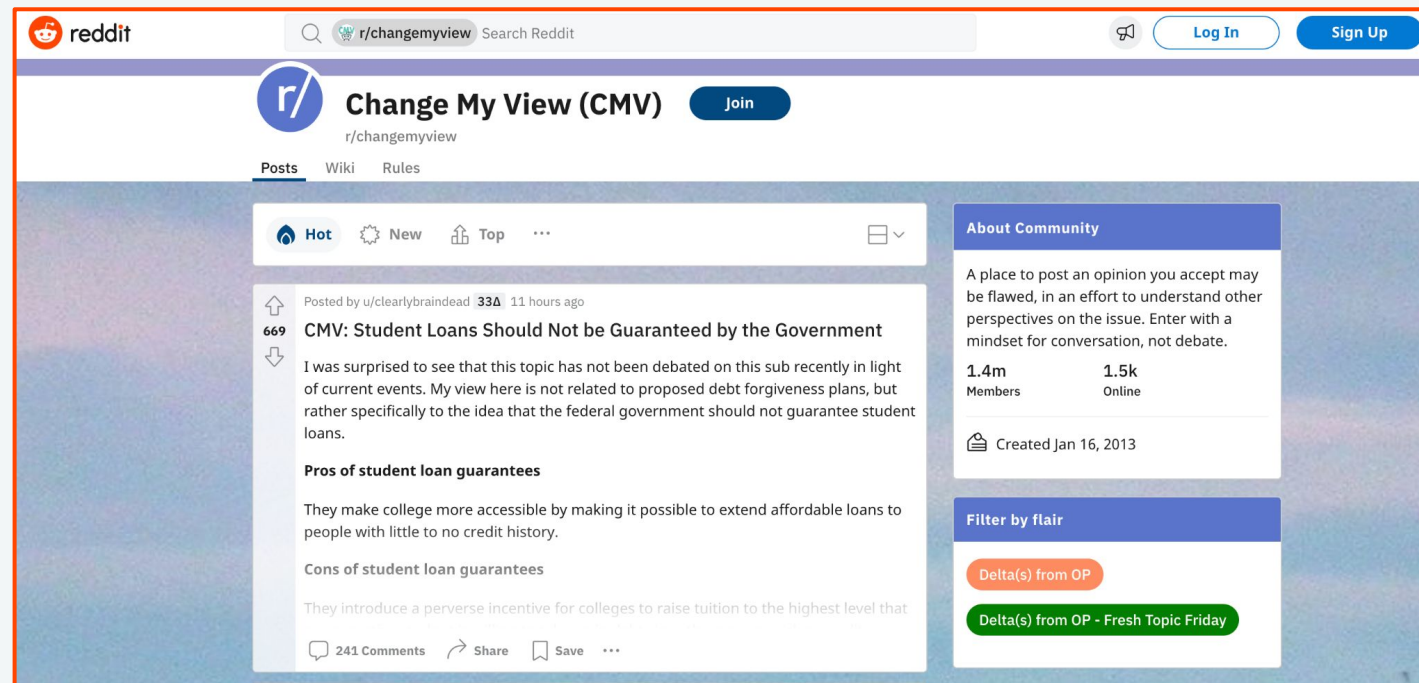


r/About Reddit

Why study Reddit?

Reddit is an open forum for dialogue with:

- moderated discussion,
- no character limit,
- democratic, thus enabling freedom of speech
- open source



ChangeMyView is the subreddit thread analyzed in this project





r/ChangeMyView

What is it?

r/changemyview, also known as Change My View (CMV), is a subreddit where participants discuss various topics for the purpose of understanding opposing views. Topics discussed include politics, media, and popular culture.

Why Delta (Δ): Assignment of a delta symbol indicates that the original poster's views have been changed.

The provision of Delta is the **Dependent Variable**. We aim to understand what factors, if any, drive this view change

How does it work?

Users submit posts containing an opinion

Respondents to post attempt to change the poster's views

Submitters must reply to these challenges in three hours or less

If their view is changed, they can award a delta symbol

Why it exists?



Open-minded debate



Scope for learning new ideas



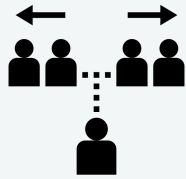
Debiasing own opinions





r/ChangeMyView

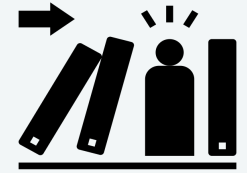
Why should we study this



Understanding what topics are controversial and hotly debated is important in studying view change



Examining how sentiments, engagement, and support influence views provides insights into the role of social influence



Determining correlates of view change can inform behavioral interventions

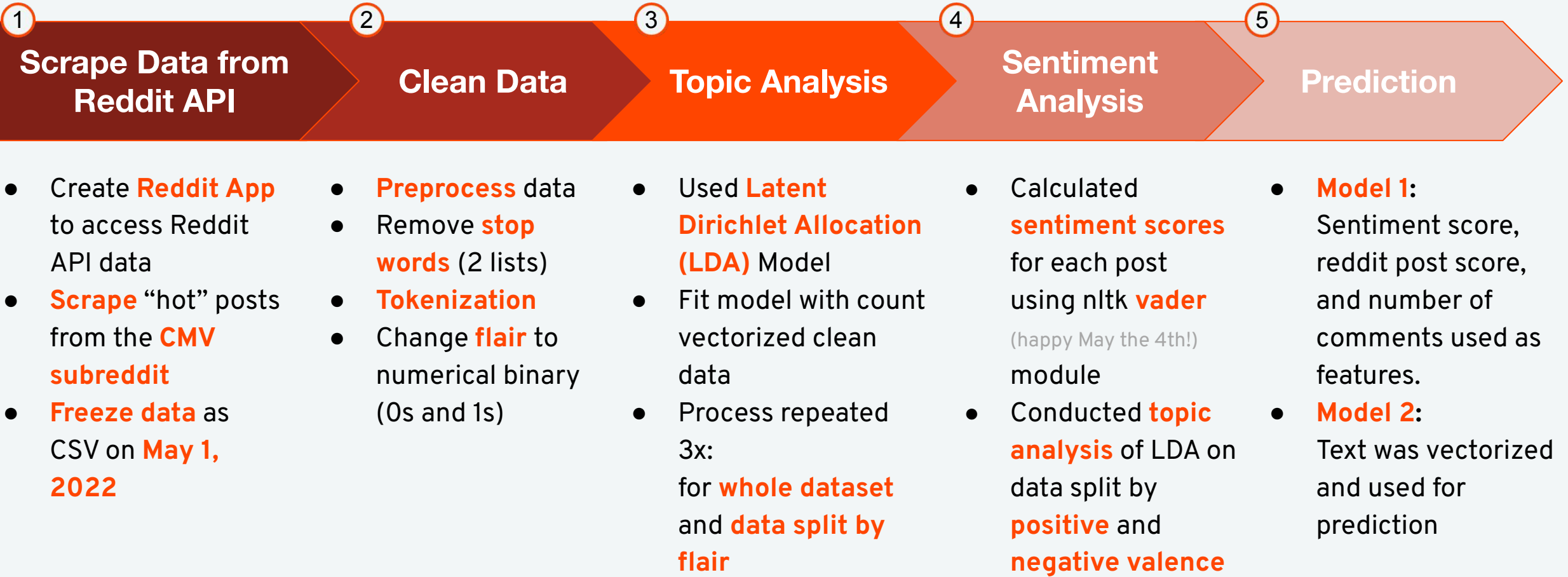
Research Questions

- Which topics do people tend to talk about?
- What topics are people currently more likely to change their opinions on?
- What is the sentiment surrounding these topics?
- How can we use sentiment scores, number of comments, and reddit scores to predict whether a view has been changed?
- How can we predict which topics can result in view change?





r/Process

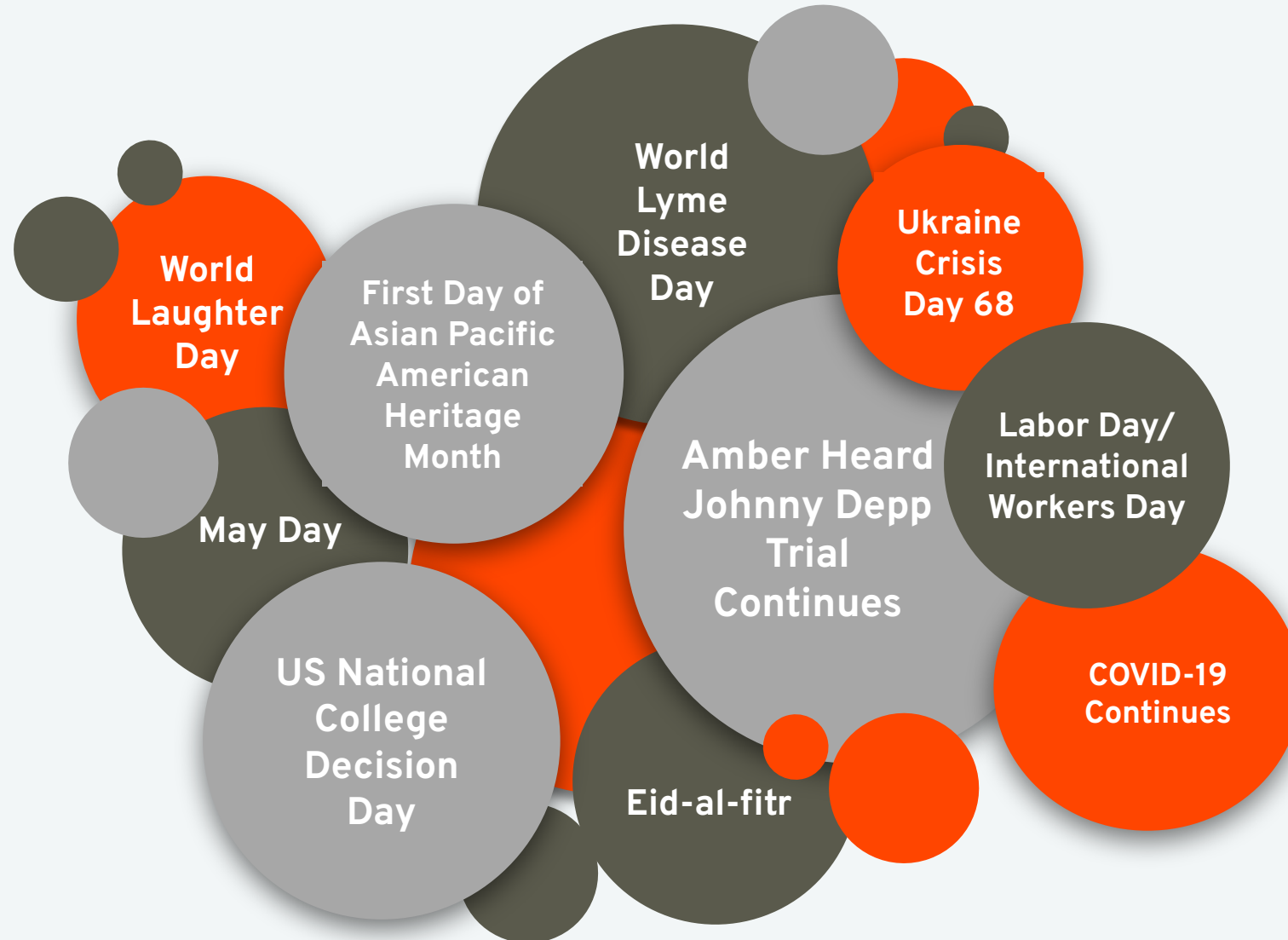


Top 500+ “Hot” Reddit CMV Posts on May 1, 2022

	title	body	flair	score	num_comments	id	url
0	CMV: Student Loans Should Not be Guaranteed by...	I was surprised to see that this topic has not...	None	168	123	ug3oa7	https://www.reddit.com/r/changemyview/comments...
1	CMV: *IF* you believe isolation from lockdown ...	edit note: \nWhile I've already awarded delta...	Delta(s) from OP	440	201	ufqwxc	https://www.reddit.com/r/changemyview/comments...
2	CMV: US Colleges should not waste student's ti...	I went to a very competitive college in the US...	Delta(s) from OP	1892	638	ufdhhz	https://www.reddit.com/r/changemyview/comments...
3	CMV: There is Nothing Wrong with Abortion What...	We do not treat life as something that is sacr...	None	15	85	ug4uhk	https://www.reddit.com/r/changemyview/comments...
4	CMV: Western countries should invest in and su...	Admittedly, I do not know much about the compl...	Delta(s) from OP	1270	140	uf9cyg	https://www.reddit.com/r/changemyview/comments...
...
576	CMV: Financially, things would get better if m...	One big aspect a lot of people seem to skip wh...	Delta(s) from OP	0	31	trogdl	https://www.reddit.com/r/changemyview/comments...
577	CMV: The United States will, at some level, ou...	Let me put my bias and world view on the table...	Delta(s) from OP	0	46	tm2l5k	https://www.reddit.com/r/changemyview/comments...
578	CMV: It is only a matter of time until Russia ...	The view I wish to have changed is that Ukrain...	Delta(s) from OP	1511	862	tjcaig	https://www.reddit.com/r/changemyview/comments...
579	CMV: Philosophy is the most important academic...	I think that philosophy is an incredibly impor...	Delta(s) from OP	21	99	tapr2x	https://www.reddit.com/r/changemyview/comments...
580	CMV: The abuse of Reddit's revamped block feat...	For those who aren't aware, a few weeks ago Re...	None	80	214	sud1gy	https://www.reddit.com/r/changemyview/comments...
581 rows x 7 columns							

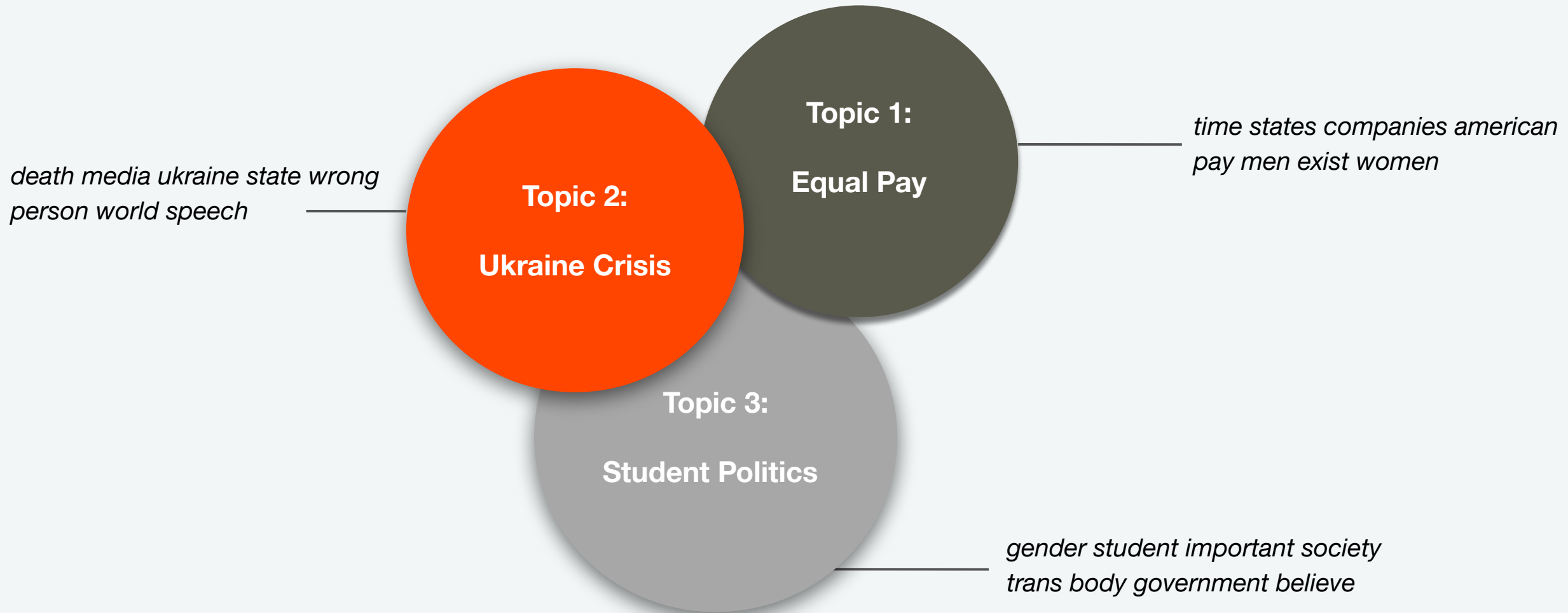


r/What Happened on May 1, 2022?





r/Topic Analysis | Overall



r/Topic Analysis | Delta vs No-Delta

*pay exist women states
companies right ukraine elon*

Topic 1:
Pay Equality

Topic 2:
Polarized Media

Topic 3:
Sports

*gender reason useless media
time wing death trans*

*men american important believe
society sports champion body*

Posts with Delta

*companies student government reason
russia society loan human*

Topic 1:
Student Loans

Topic 2:
Western Ideals

Topic 3:
Women's Rights

*world state america sexual
time student western wrong*

*men gender death change women
body ukraine society*

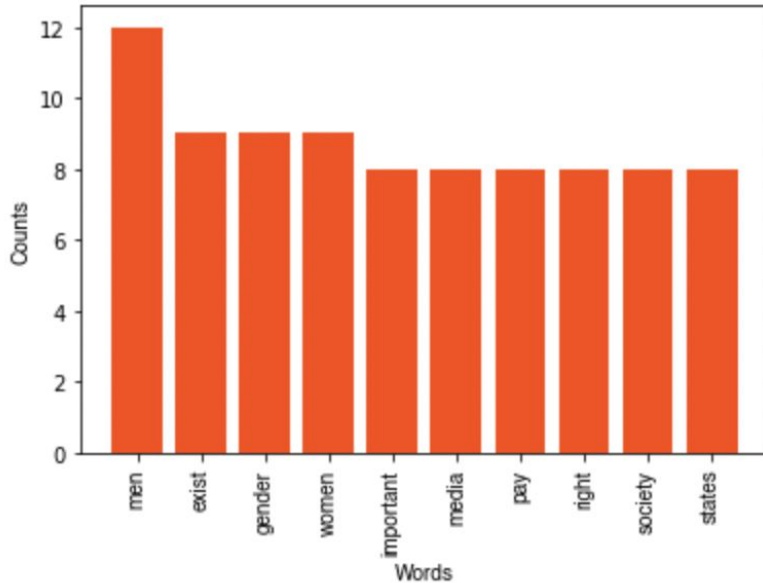
Posts with no-Delta





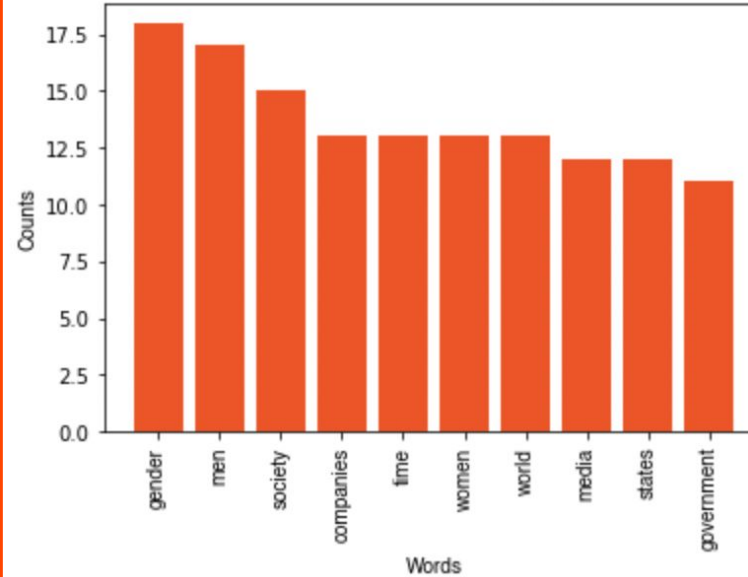
r/Common Words | Delta vs No-Delta

10 Most Common Words for Delta awarded



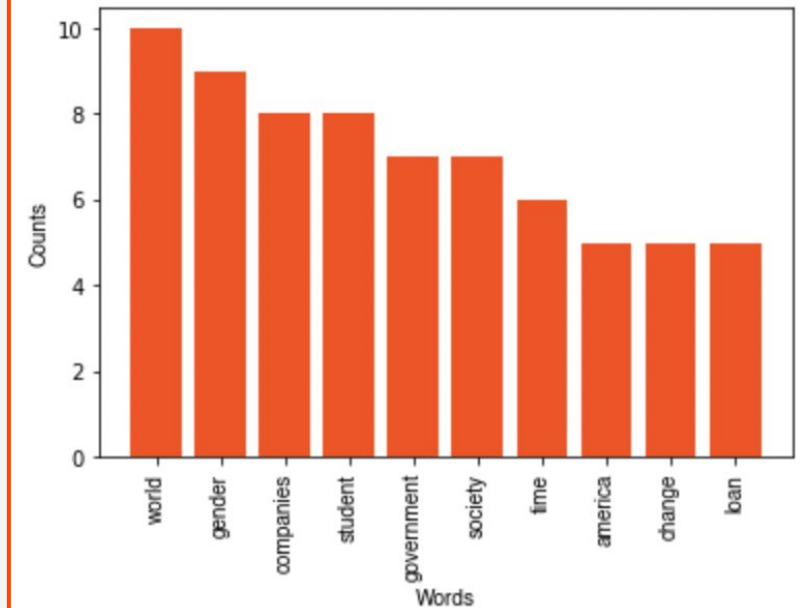
Posts with Delta

10 Most Common Words for Complete Dataset



Overall Posts

10 Most Common Words for No Delta awarded



Posts with no-Delta





r/Implications



The topics covered overall correlate with what happened in the world on May 1st

*Equal Pay, Ukraine Crisis,
Student Politics*



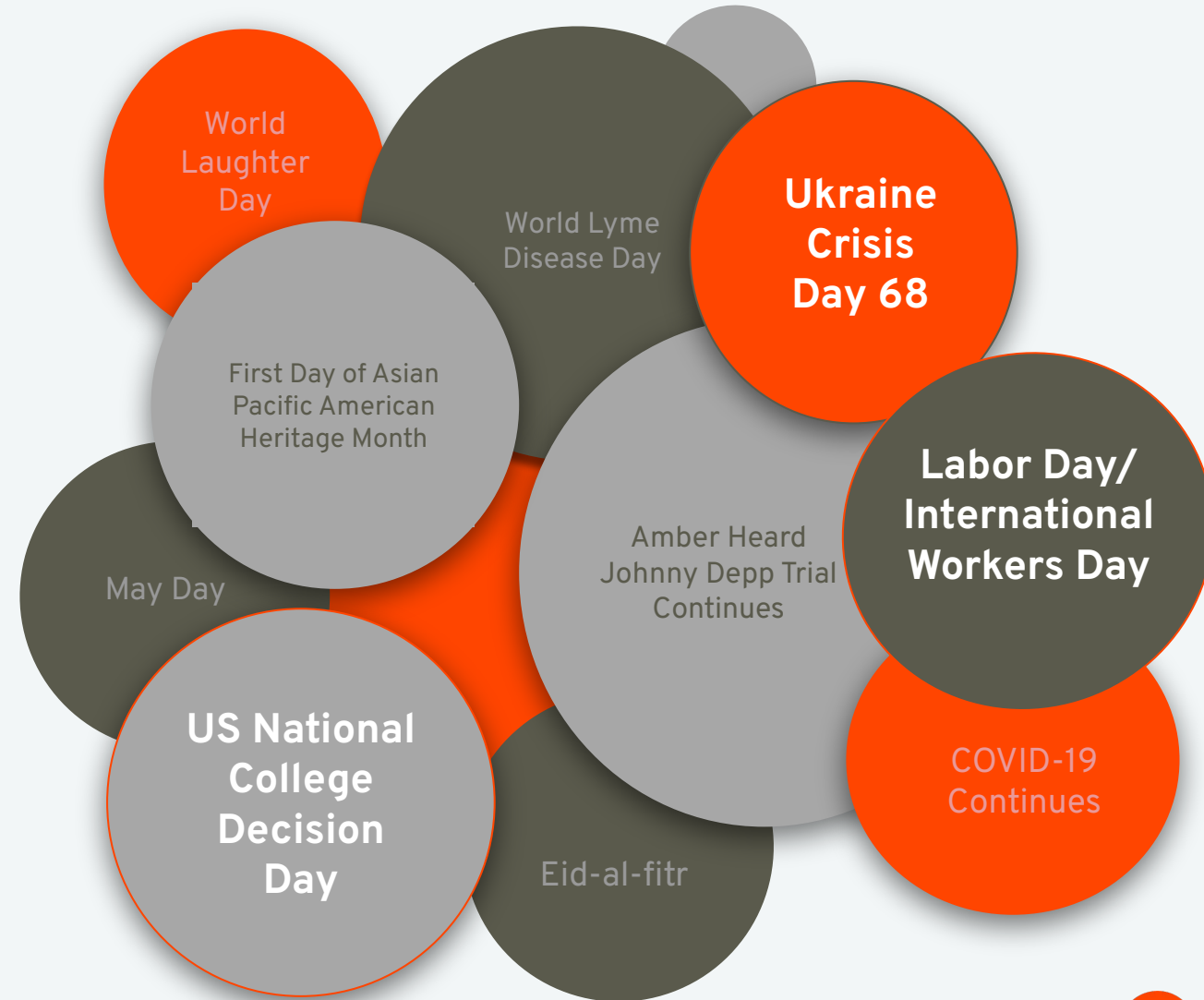
Deltas are awarded to current events or topics you discuss on a first date

Sports, Media, Elon



Deltas are not awarded regarding long held values or those you discuss with a spouse

Sexuality, Government, Debt





r/Topic Analysis | Positive vs Negative

*rights russia media gender care
ukraine government student*

Topic 1:
Ukraine
Crisis

Topic 2:
Social Welfare

Topic 3:
Gender Politics

*society person exist problems
useless state work important*

*trans reason argument elon gender
valid america wrong*

Positive

*men companies world anything school
religion home change*

Topic 1:
Rightwing
Politics

Topic 2:
American
Culture

Topic 3:
Polarized Media

*society american wrong sports
women ukraine wild problem*

*death champion twitter gender penalty
worst date white*

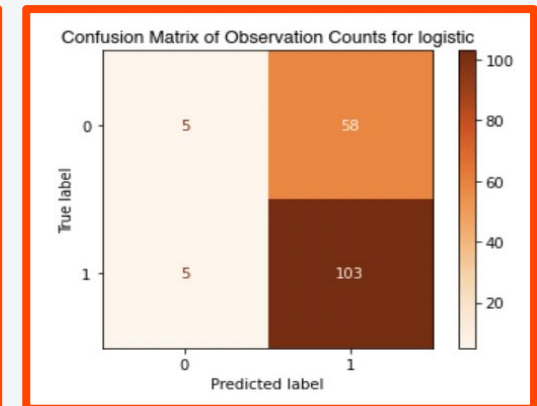
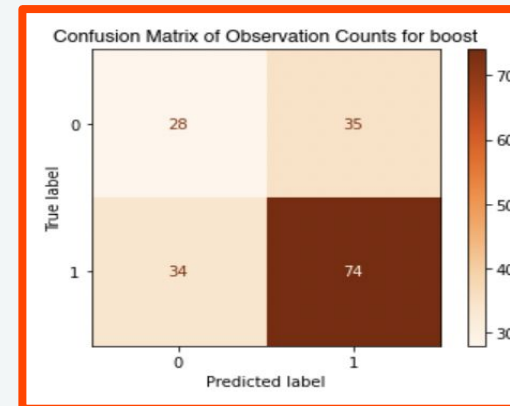
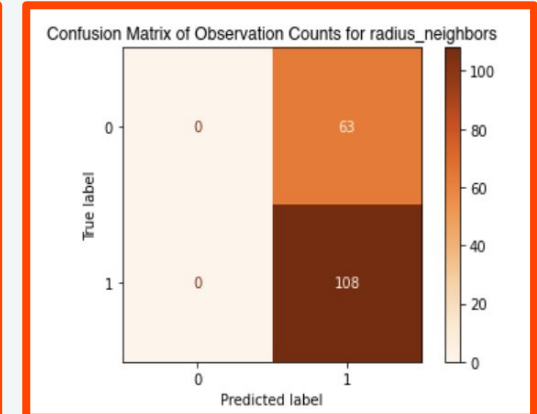
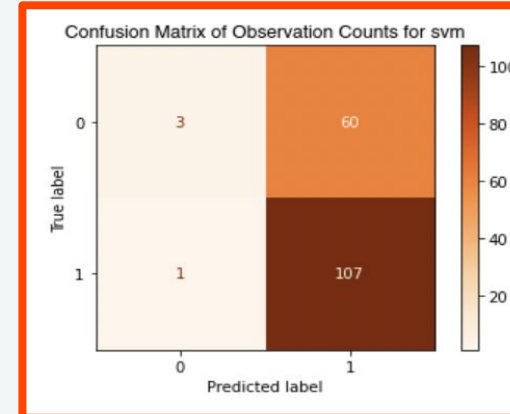
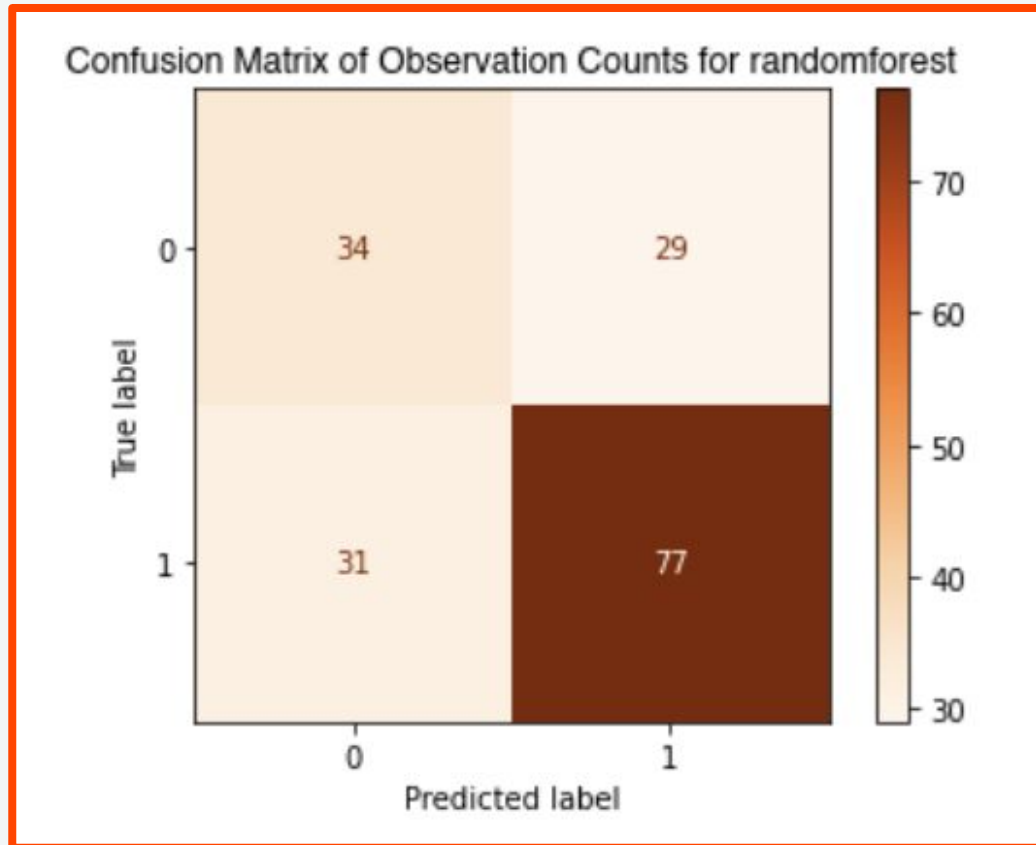
Negative





r/Model 1

Predicting Delta Using Sentiments and Engagement



Accuracy: 56% - 65%





r/Logistic Regression

```

=====
Dep. Variable:          flair    No. Observations:          567
Model:                Logit     Df Residuals:              564
Method:               MLE       Df Model:                  2
Date:                 Tue, 03 May 2022    Pseudo R-squ.:            0.03189
Time:                 12:10:39    Log-Likelihood:           -369.90
converged:            True       LL-Null:                  -382.08
Covariance Type:      nonrobust    LLR p-value:              5.115e-06
=====

```

	coef	std err	z	P> z	[0.025	0.975]
x1 score	0.1426	1.493	0.096	0.924	-2.784	3.069
x2 num_comments	8.5578	2.661	3.216	0.001 ***	3.341	13.774
x3 compound	0.1381	0.145	0.951	0.342	-0.147	0.423

```

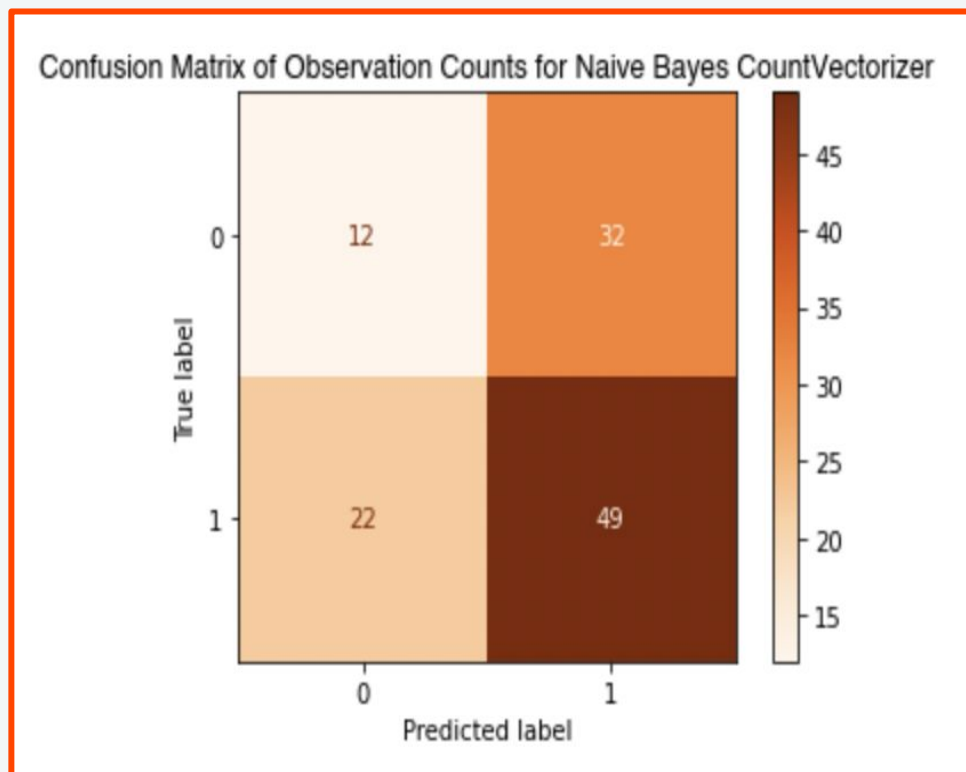
=====

```

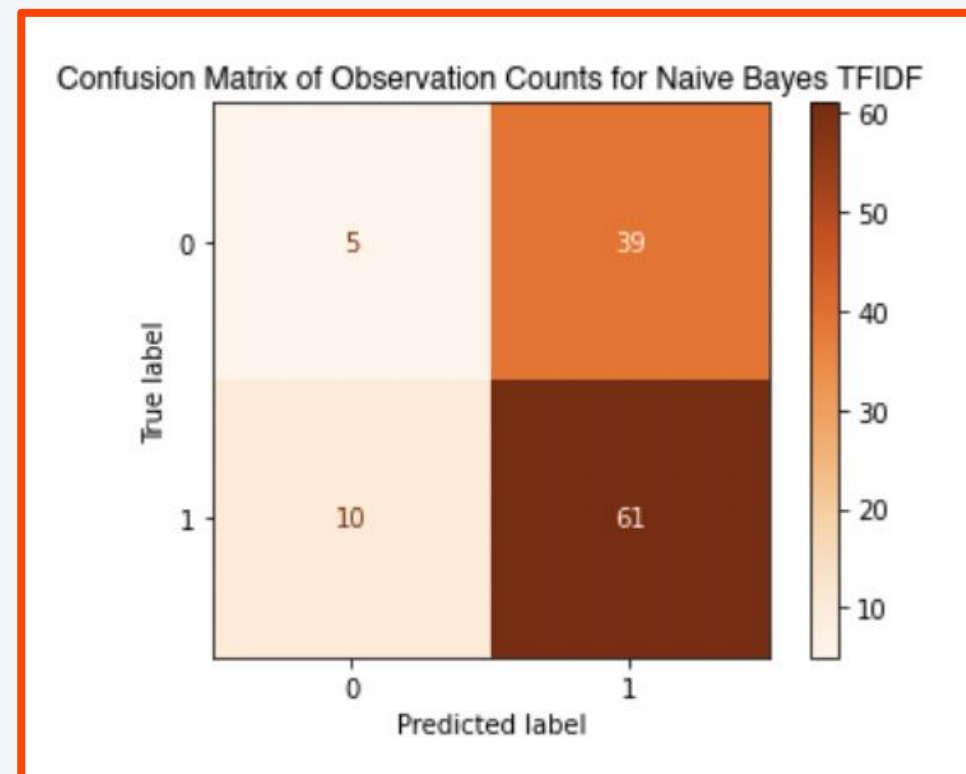
Takeaway: Number of comments is associated with delta awarding.
Hence, there could be a normative component at play that needs to be explored.



Predicting Delta Using Vectorized Text



Accuracy: 53%



Accuracy: 57%





r/Testing Model 2

Using Domain Knowledge to Test Model 2

```
delta_post = "I do not feel very strongly about veganism, so please try to convince me"
nondelta_post = "I will never change my view about veganism, but come at me"

delta_post_vectorized = tfidf_vectorizer.transform([delta_post])
nondelta_post_vectorized = tfidf_vectorizer.transform([nondelta_post])

delta_post_pred = tfidf_svc.predict(delta_post_vectorized)
nondelta_post_pred = tfidf_svc.predict(nondelta_post_vectorized)

print('Predicted Post that Led to Changed View', delta_post_pred)
print('Predicted Post that Did Not Lead to View Change', nondelta_post_pred)
```

```
Predicted Post that Led to Changed View [1]
Predicted Post that Did Not Lead to View Change [0]
```

Our model gives us an important insight - maybe topics don't matter at all for prediction.





r/Testing Model 2

Using Domain Knowledge to Test Model 2

```
delta_post = "I do not feel very strongly about Alex winning Eurovision, so please try to convince me"
nondelta_post = "I will never change my view about Alex winning Eurovision, but come at me"

delta_post_vectorized = tfidf_vectorizer.transform([delta_post])
nondelta_post_vectorized = tfidf_vectorizer.transform([nondelta_post])

delta_post_pred = tfidf_svc.predict(delta_post_vectorized)
nondelta_post_pred = tfidf_svc.predict(nondelta_post_vectorized)

print('Predicted Post that Led to Changed View', delta_post_pred)
print('Predicted Post that Did Not Lead to View Change', nondelta_post_pred)

Predicted Post that Led to Changed View [1]
Predicted Post that Did Not Lead to View Change [0]
```

Yup, they don't seem to matter.





r/Key Takeaways

What did we learn?

- Predicting view change is **HARD**
 - Some idea from prediction model that language used by OP can influence delta awarding
 - Can we even predict view change without understanding characteristics of OP?
- **Limitations:** Using only May 1st data, delta awarding as proxy for view change

Future Research

- Exploration of whether higher engagement on posts causes view change
- Further analysis on the comments of the posts to identify the types of comments that lead to view change

Thank you!



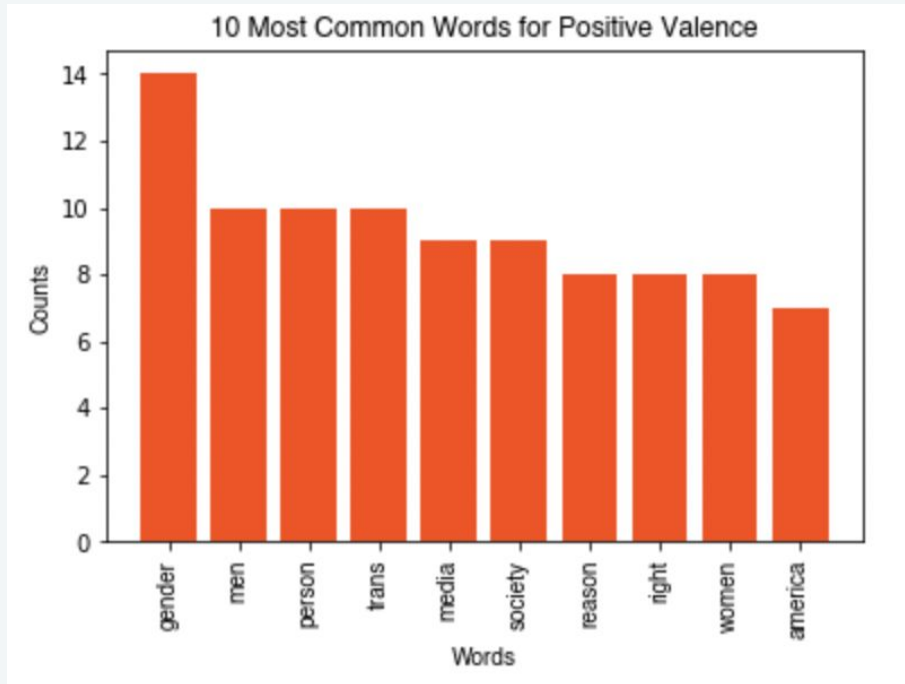
Appendix

Extra Analysis

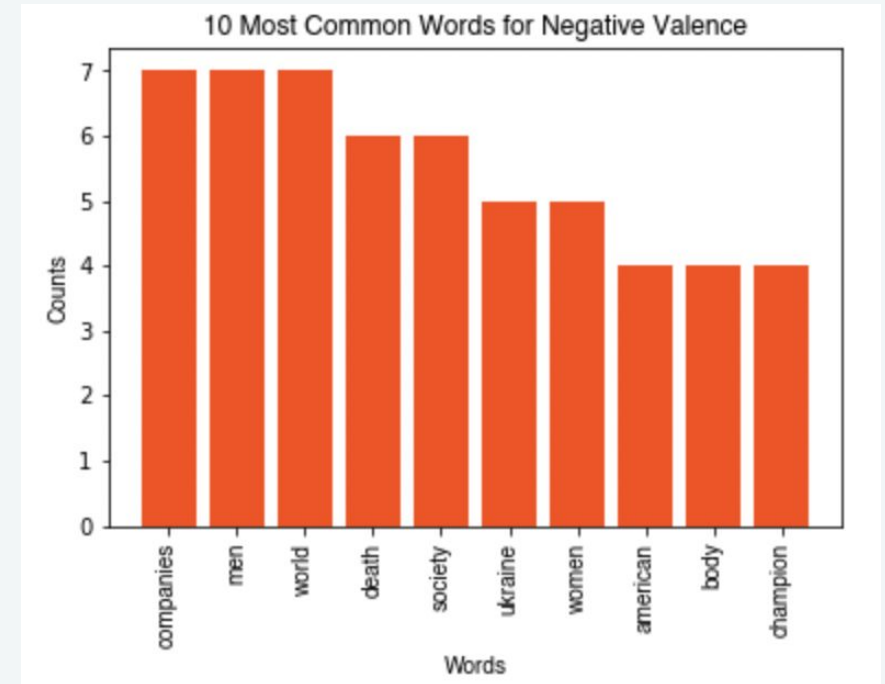




r/Common Words | Positive vs Negative



Positive



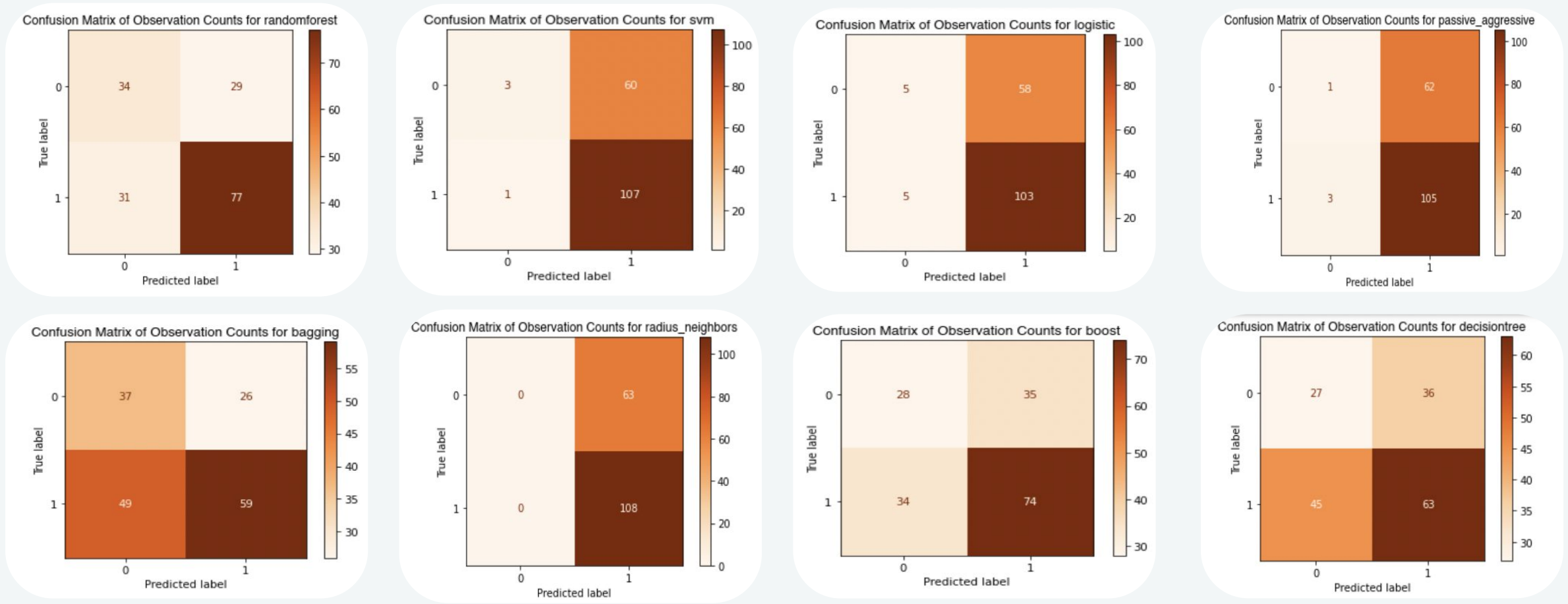
Negative





r/Model 1

Predicting Delta Using Sentiments and Engagement



Accuracy: 53% - 65%

