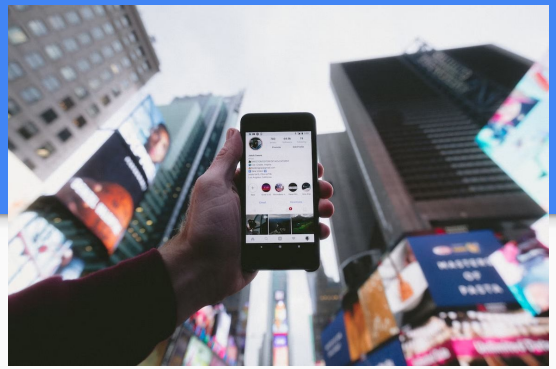
Using hashtag and word graph networks to analyze trending social media topics and communication

Andrew Curtis

Presentation Overview

1. Goal and motivation of research
2. Background: Twitter, useful terms, example topics
3. Data collection
4. Methods employed
5. Results
6. Analysis
7. Future work

Motivation



Social media usage is ubiquitous in current culture

It acts as a stage for communication and discourse

Often leading to policy and societal changes outside of the platform

Trends and social commentary arise and die rapidly

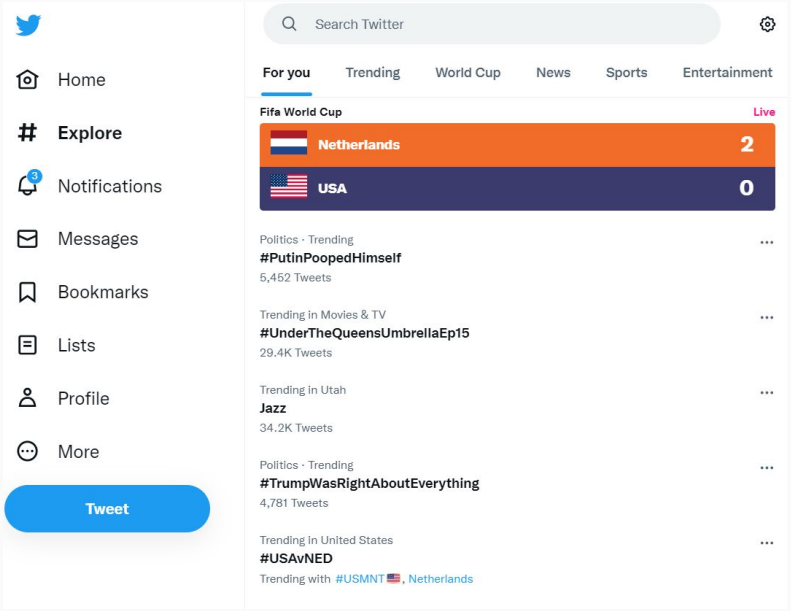
The goal of this research is to study the use of graph network analysis as a tool for analyzing engagement on a social media platform

Hashtags and trending Twitter

Hashtags are a way to index and group tweets

These can be more useful in grouping ideas across the platform than traditional liking, commenting and sharing

In this analysis, hashtags and words appearing in tweets were used to construct network graphs



A Co-Creational Perspective

A useful term, Co-Creational Perspective (Xiong, Cho, Boatright 2017)

Shared meaning is created from words and interpretation

Entities on social media platforms, including the public and organizations rapidly agree on these terms

They arise and die, but can elucidate key themes within viral movements

Different communities of interest are involved in this process depending on topic selected

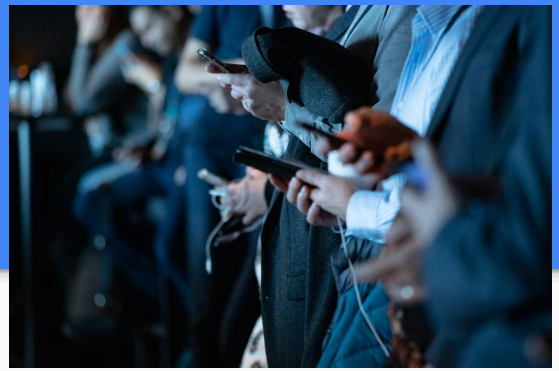
Chosen Topics

Topics were chosen to be representative of trending and culturally significant events and ideas with a variety of subject matter

- NFL and Concussions
- Elon Musk and Twitter Takeover
- Kanye West and Anti-Semitism

In practice, the subject matter and different communities interacting with each topics greatly influenced the results of the analysis methods applied

Data Collection



For each topic,

- Search terms of hashtags and words appearing in tweets
- 3000 tweets pulled per test data set
- Taken between the period of 1st-3rd December 2022
- Retweets filtered out
- Attempts to remove stopwords and similar words

Graph Network Creation

Two methods were used for selecting nodes when creating the graph networks

- Nodes using hashtags or words appearing in tweet
- Nodes using bigrams

Edges created for hashtags, words or bigrams that appear together in other tweets

Analysis Methods

Several qualitative and quantitative measures were employed including:

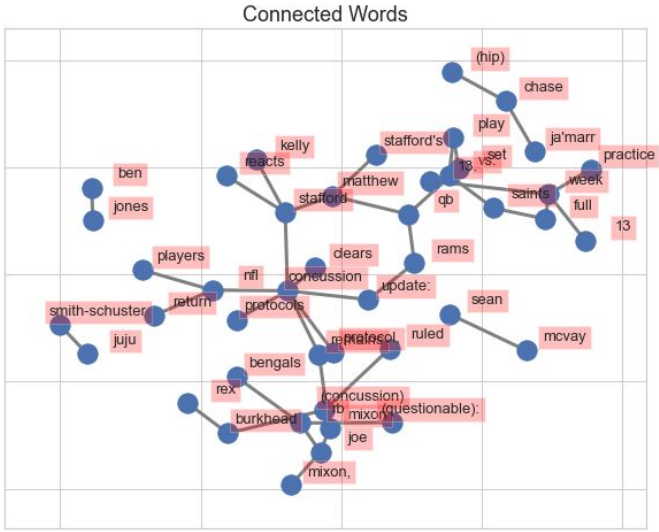
- Graph visualizations of connected words, hashtags and bigrams
- Word counts for subject analysis
- Normalized degree centralities of created network words and hashtags
- Sentiment analysis histograms, total and polarized

NFL and Concussions: Bigram Graphical Representation

Search terms used were “NFL” and “concussions”

Most content of the tweets was informational about player injuries

Often discussion being led by sports news organizations



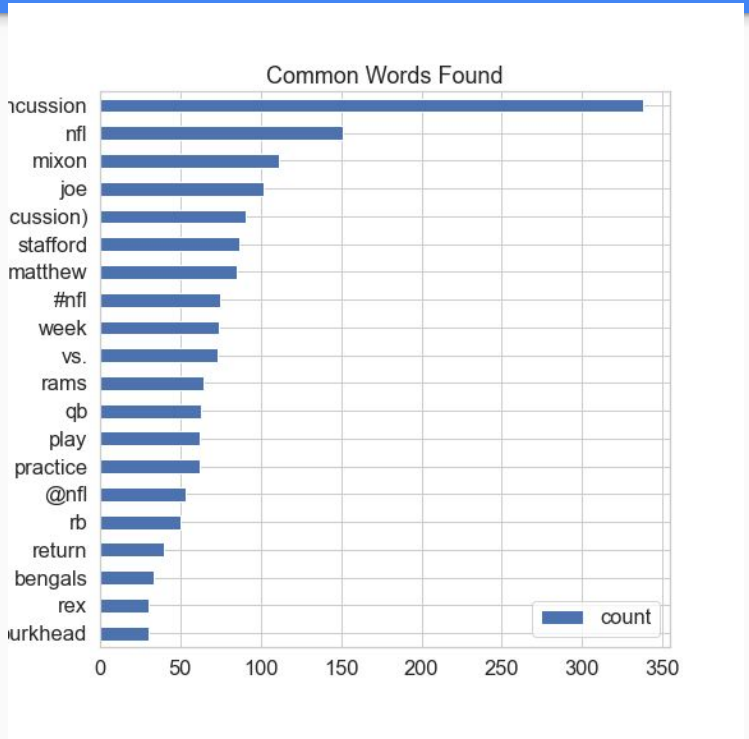
40 most common bigrams and connections

NFL and Concussions: Common Words and Degree Centrality

Informational terms

Mostly player and team related

Tweets tend to come from teams and sports journalism organizations

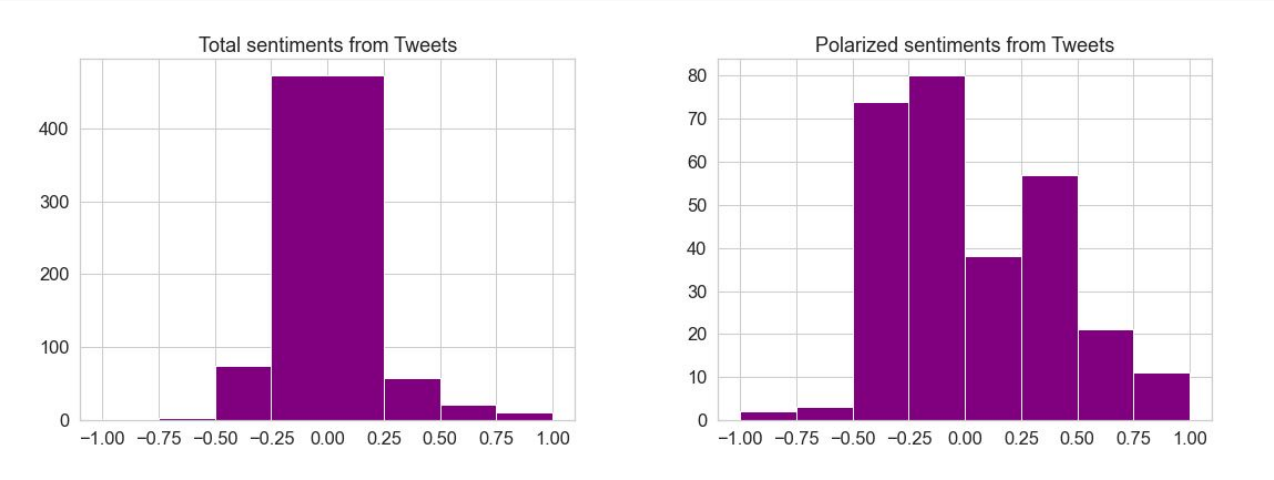


	Words	Normalized Degree Centrality
0	concussion	0.162791
1	stafford	0.093023
2	week	0.093023
3	joe	0.069767
4	mixon	0.069767
5	(concussion)	0.069767
6	matthew	0.069767
7	qb	0.069767
8	nfl	0.069767
9	rb	0.069767
10	rams	0.046512
11	play	0.046512
12	vs.	0.046512
13	burkhead	0.046512
14	set	0.046512
15	update:	0.046512
16	saints	0.046512
17	full	0.046512
18	remains	0.046512
19	chase	0.046512
20	rex	0.023256
21	practice	0.023256
22	kelly	0.023256
23	reacts	0.023256
24	return	0.023256
25	13	0.023256
26	protocol,	0.023256
27	stafford's	0.023256
28	ben	0.023256
29	jones	0.023256
30	ja'marr	0.023256
31	(hip)	0.023256
32	bengals	0.023256
33	clears	0.023256
34	mixon,	0.023256
35	ruled	0.023256
36	juju	0.023256
37	smith-schuster	0.023256
38	protocols	0.023256
39	players	0.023256
40	(questionable):	0.023256
41	13.	0.023256
42	sean	0.023256
43	mcvay	0.023256

NFL and Concussions: Sentiment Analysis

Mostly unpolarized

Few extreme negative polarized tweets

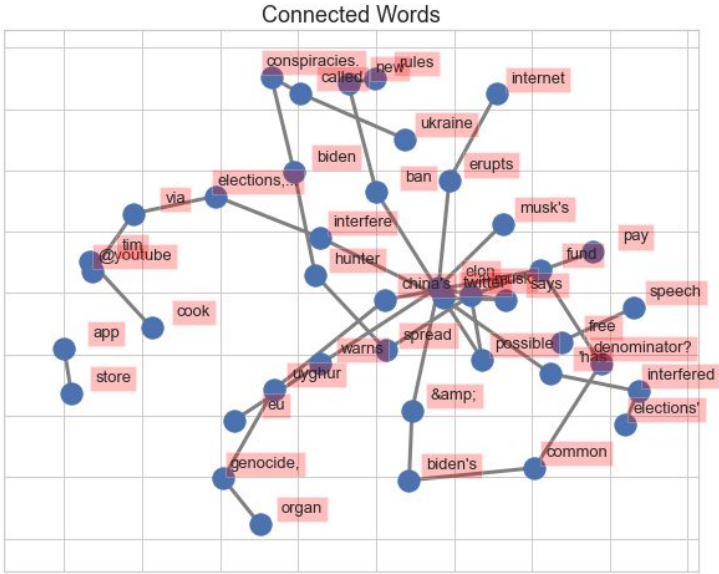


Elon and Twitter: Connected Words

Search terms used: “Elon” and “twitter”

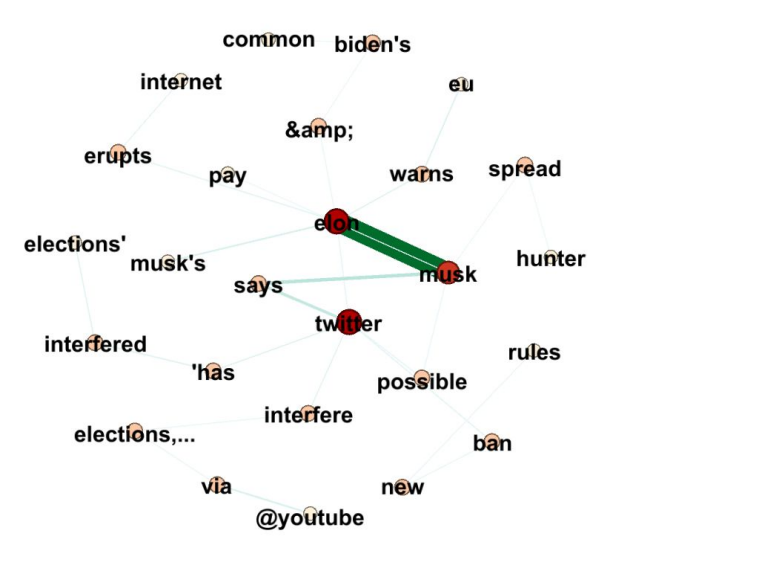
The content of tweets in these connected spaces are often very temporally-related

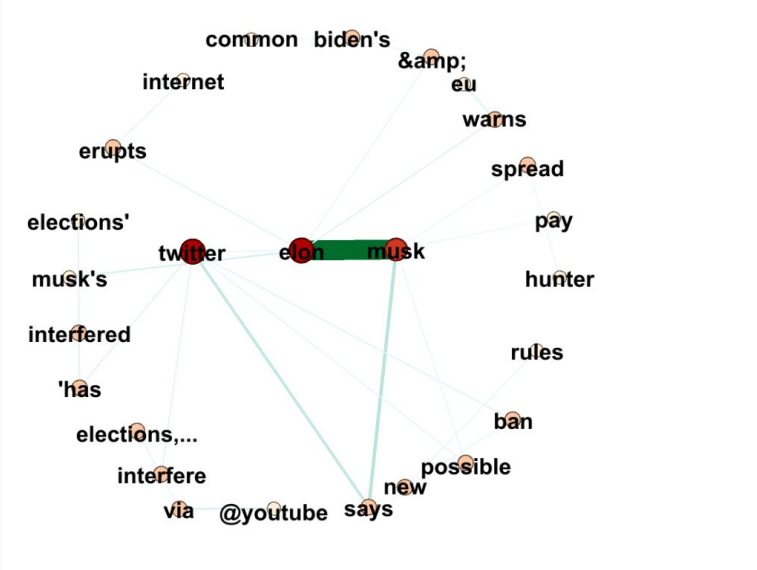
For example, most tweets relating to Elon when this was captured discussed rumors of Twitter’s possible ban from the Apple app store



40 most common bigrams and connections

Elon and Twitter: Ego-Centric Representation



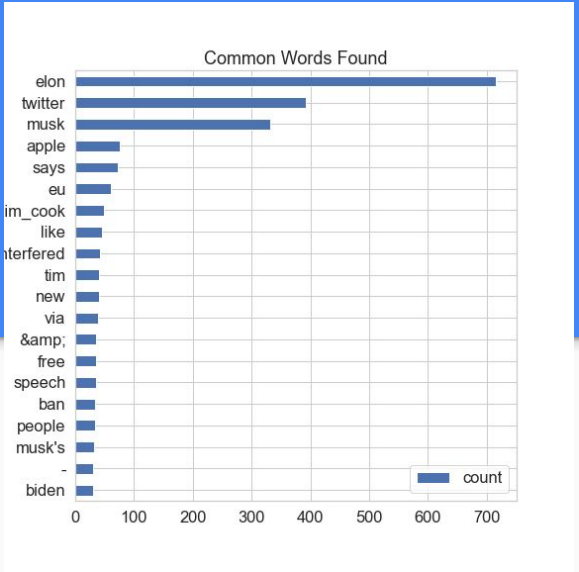


Elon and Twitter

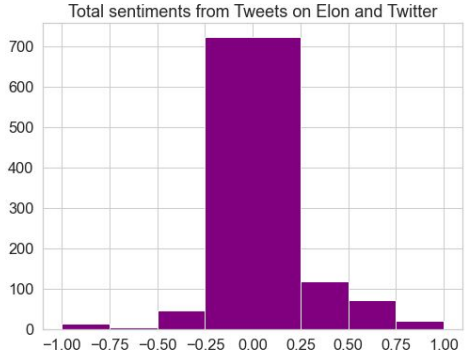
Not many co-creational terms present

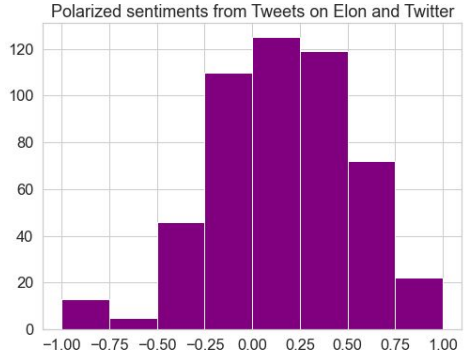
However topics of discussion were kept on track by common themes

Such as free speech and the potential app store ban



	Words	Normalized Degree Centrality
0	twitter	0.250000
1	elon	0.208333
2	musk	0.083333
3	says	0.083333
4	warns	0.083333
5	'has	0.083333
6	interfered	0.083333
7	erupts	0.083333
8	via	0.041667
9	@youtube	0.041667
10	musk's	0.041667
11	free	0.041667
12	speech	0.041667
13	tim	0.041667
14	cook	0.041667
15	eu	0.041667
16	elections'	0.041667
17	ban	0.041667
18	interfere	0.041667
19	internet	0.041667
20	possible	0.041667
21	app	0.041667
22	store	0.041667
23	hunter	0.041667
24	biden	0.041667





Kanye West: Topic Results

The Kanye West data was collected in a slightly different fashion than the other topics

Two pulls were made four hours apart for comparison

This topic featured many more tweets from accounts representing individuals

Public engagement was much higher as well as the presence of co-creational terms

These terms rapidly came to use and then died out

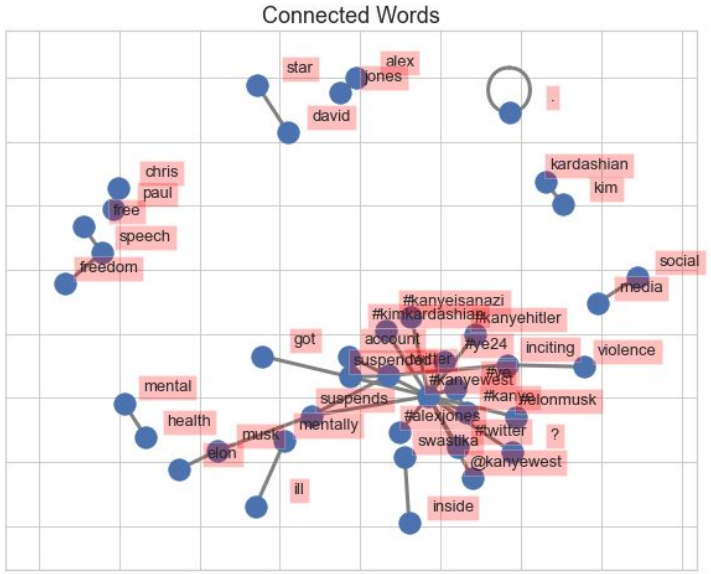
Kanye West: Connected Words

Co-created hashtags and words used are more informal than other topics

Examples can be seen like: #kanyeisanazi, #kanyehitler

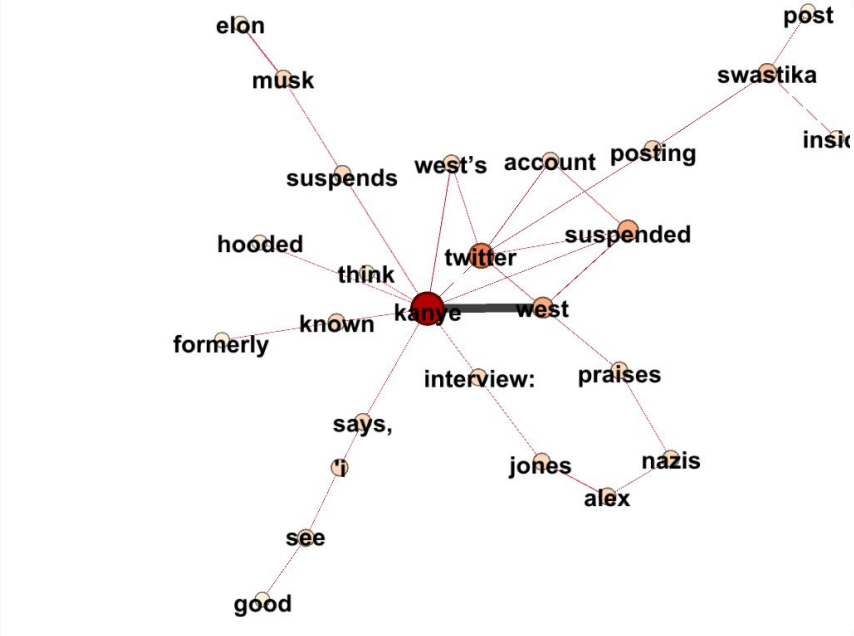
Verbs are more prevalent

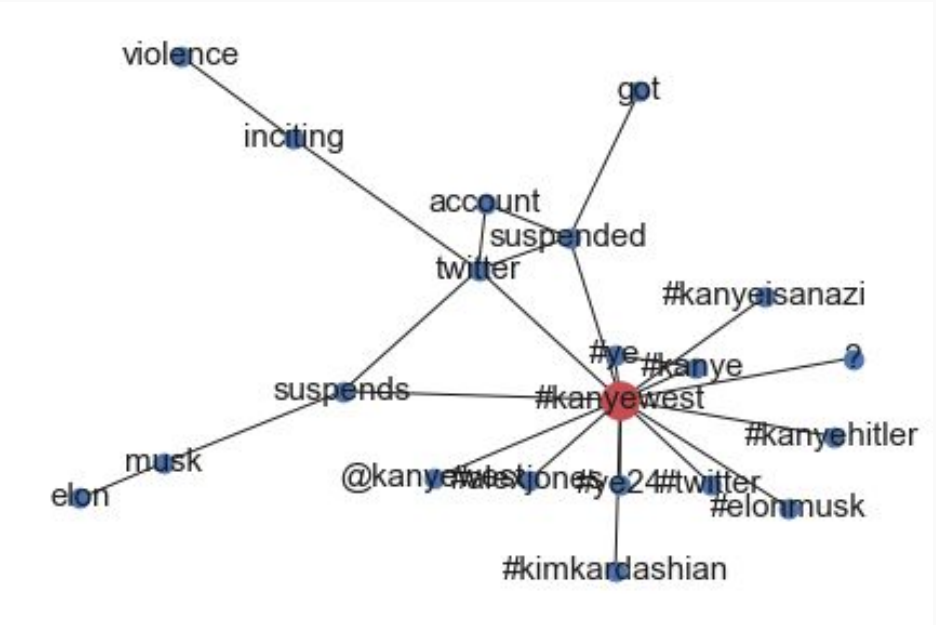
Topics of conversation such as mental health and free speech are prominent



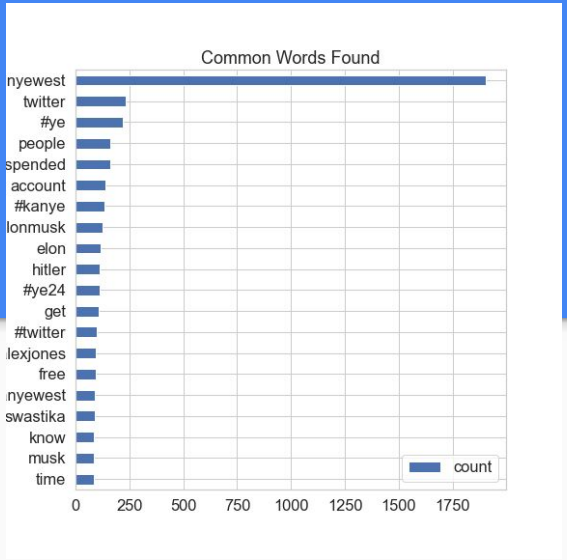
40 most common bigrams and connections

Kanye West: Ego-Centric Representations



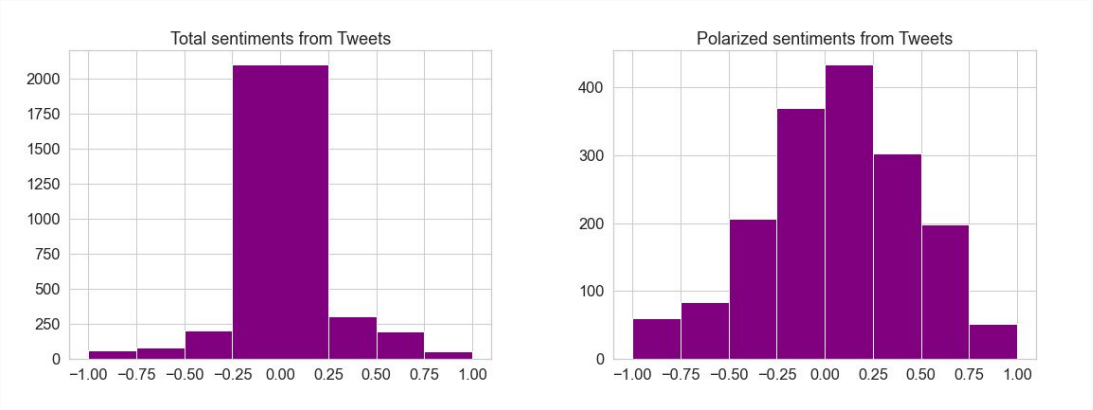


Kanye West



	Words	Normalized Degree Centrality
0	kanye	0.256410
1	twitter	0.153846
2	west	0.102564
3	suspended	0.102564
4	swastika	0.076923
5	musk	0.051282
6	alex	0.051282
7	jones	0.051282
8	west's	0.051282
9	account	0.051282
10	posting	0.051282
11	suspends	0.051282
12	praises	0.051282
13	good	0.051282
14	see	0.051282
15	known	0.051282
16	10	0.051282
17	nazis	0.051282
18	interview:	0.051282
19	'i	0.051282
20	says,	0.051282
21	elon	0.025641
22	kim	0.025641
23	kardashian	0.025641
24	chris	0.025641
25	paul	0.025641
26	things	0.025641
27	think	0.025641
28	free	0.025641
29	speech	0.025641
30	formerly	0.025641
31	top	0.025641
32	inside	0.025641
33	searches	0.025641
34	hooded	0.025641
35	nick	0.025641
36	fuentes	0.025641
37	post	0.025641
38	star	0.025641
39	david	0.025641

	Words	Normalized Degree Centrality
0	#kanyewest	0.350
1	twitter	0.125
2	suspended	0.100
3	suspends	0.075
4	musk	0.050
5	#ye	0.050
6	#kanye	0.050
7	account	0.050
8	speech	0.050
9	inciting	0.050
10	.	0.050
11	elon	0.025
12	alex	0.025
13	jones	0.025
14	chris	0.025
15	paul	0.025
16	#kanyeisanazi	0.025
17	free	0.025
18	#ye24	0.025
19	#alexjones	0.025
20	kim	0.025
21	kardashian	0.025
22	#twitter	0.025
23	#elonmusk	0.025
24	violence	0.025
25	mental	0.025
26	health	0.025
27	#kanyehitler	0.025
28	freedom	0.025
29	star	0.025
30	david	0.025
31	swastika	0.025
32	inside	0.025
33	social	0.025
34	media	0.025
35	?	0.025
36	got	0.025
37	mentally	0.025
38	ill	0.025
39	@kanyewest	0.025
40	#kimkardashian	0.025



Analysis: Trends among Trending

The co-creation of terms and hashtags is dependent on whether they are being primarily driven in posts by organizations or the public

Using words and hashtags as filters for creating graph networks for analysis can have very different effects

Cannot conclude that choosing words or hashtags is more effective than the other, depends on co-creational development within the network

Trends and topics are very dynamic, development of consistent co-creational terms is essential for long-term movements and discussion

Analysis: Difficulty of Social Network Analysis

Analysis of social media networks can be approached from many different perspectives

Fairly easy to generate networks and quantitative analysis measures,

Can be difficult to interpret and extract actionable meaning

Social media networks and trends are extremely dynamic

Snapshots in time can be interesting to analyze but information diffusion seems like it have more actionable results

Analysis: Recommended Methods and Improvements

Focusing on Hashtags increases topic relevance

Focusing on words increases connections in the networks analysis

Contextualize the networks by giving background on topics and trends

Suggested methodology:

1. Identify topic of interest
2. Network exploration
3. Identify community of interest
4. Revise analysis
5. Contextualize and present results

Future Work

Introduce dynamic networks to capture information diffusion and better elucidate trending topics

Improve visualization and leverage to convey information more effectively

Select and limit work on more specific communities of interest

Select a more narrow goal or motivation for research