

Hussain Wasif Latif  
UID: 3035834577

# Memes in Motion

Analyzing effects of Memes & Virality



# Data Science Question to Answer

What are the financial impacts of memes going viral?

# Why does it matter?



# Why does it matter?

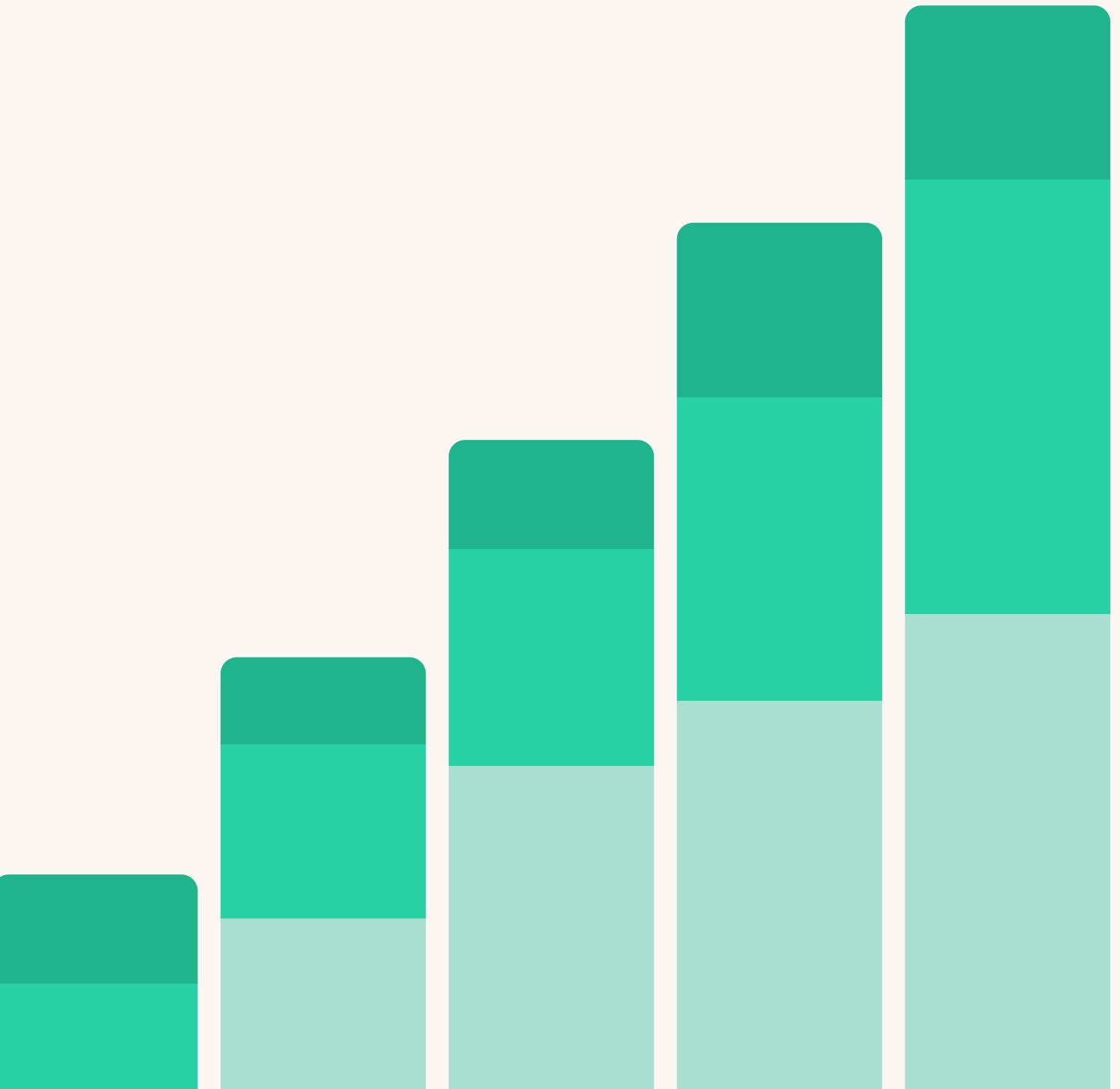


## Significant Financial Impact

- GameStop shares (also called Memestock)
  - Price \$20 -> \$483 intraday and \$90 closing price
  - \$20 Billion+ in losses
  - Millions of individuals hyped up the stock via memes and viral content

# Overview

- Data Sources & Preprocessing
- Difficulties I Faced
- Phenomenon & Detailed Analysis of Trends
- Reflecting on whether it actually had an effect.
- Recommendation



# Data Collection Methods



## Stock Data

Yahoo Finance API (<https://developer.yahoo.com/api/>)



## Reddit Data

**Pushshift API & Online Dataset** ([https://www.researchgate.net/publication/362904749\\_Rddit\\_Dataset\\_on\\_Meme\\_Stock\\_GameStop](https://www.researchgate.net/publication/362904749_Rddit_Dataset_on_Meme_Stock_GameStop))

**Scraped r/WallStreetBets under “GME”/“gamestop”**

## Google Trends

Search regarding WSB, \$GME, Roaring Kitty, etc:

<https://trends.google.com/trends/>



## Twitter Data

**Apify Twitter Scraper - Fetched tweets with \$GME, #GameStop** during the relevant period

(<https://apify.com/quacker/twitter-scraper>)

**Kaggle Dataset for Tweet Sentiments** (<https://www.kaggle.com/datasets/rahilrathod/gamestop-gme-reddit-sentiment-analysis>)



# Data Processing and Cleaning

- Handling missing values

```
reddit.dropna(subset=['text'], inplace=True) # Remove posts with no text  
tweets['text'] = tweets['title'].fillna('') + ' ' + tweets['body'].fillna('')
```

- Data Standardization

```
tweets['date'] = pd.to_datetime(tweets['timestamp'], errors='coerce').dt.date  
gme['Date'] = pd.to_datetime(gme['Date']).dt.date
```

print(tweets_daily.head())				
✓ 0.0s				
date	tweets_per_day	avg_polarity	avg_subjectivity	
0 2021-01-28	249	0.035378	0.264680	
1 2021-01-29	2940	0.024758	0.249673	
2 2021-01-30	290	0.017414	0.273029	
3 2021-01-31	236	0.042245	0.321905	
4 2021-02-01	202	0.025666	0.328620	

print(combined_dataset.head(6))				
✓ 0.1s				
date	Close	return	posts_per_day	avg_upvotes
2021-01-04	0.3125	NaN	2.0	1.000000
2021-01-05	0.3425	0.006957	2.0	1.000000
2021-01-06	0.5900	0.056995	2.0	1.000000
2021-01-07	0.5200	-0.015251	0.0	0.000000
2021-01-08	0.4225	-0.021571	0.0	0.000000
2021-01-11	0.9850	0.127190	4.0	1.250000
2021-01-12	0.9875	0.000502	2.0	1.500000
2021-01-13	0.8500	0.573935	14.0	1.285714
2021-01-14	0.9775	0.271019	9.0	1.111111
2021-01-15	0.8750	-0.110499	14.0	1.642857



Unnamed: 0	id	title	url	score	author
0	kqfajb	You NEED to see this about GME 🚨🚨🚨 https://www.		1	TitsDownC
1	kqjh2t	Short squeeze incoming 🚨🚨🚨 wallstreetb		1	zoomermon
2	kqvp7l	Y'all ready for this? 🚨🚨🚨 wallstreetb		1	TitsDownC
3	krcwch	ICR conference (11th Jan)	https://www.	1	dontforget
4	krnths	Hey guys! We have a free discord channel t	https://www.	1	nicky94
5	kryizd	GME is FINALLY going to the moon, this tec	https://www.	1	thehelper9
6	kuo3w1	Ryan Conen appointed to board!!!!?	https://news.	1	nicky94
7	kv1t51	BUCKLE YOUR SEATBELTS OR YOU'RE GOING	https://www.	1	username-
8	kv1w9e			1	jonastiron
9	kv3vrm				

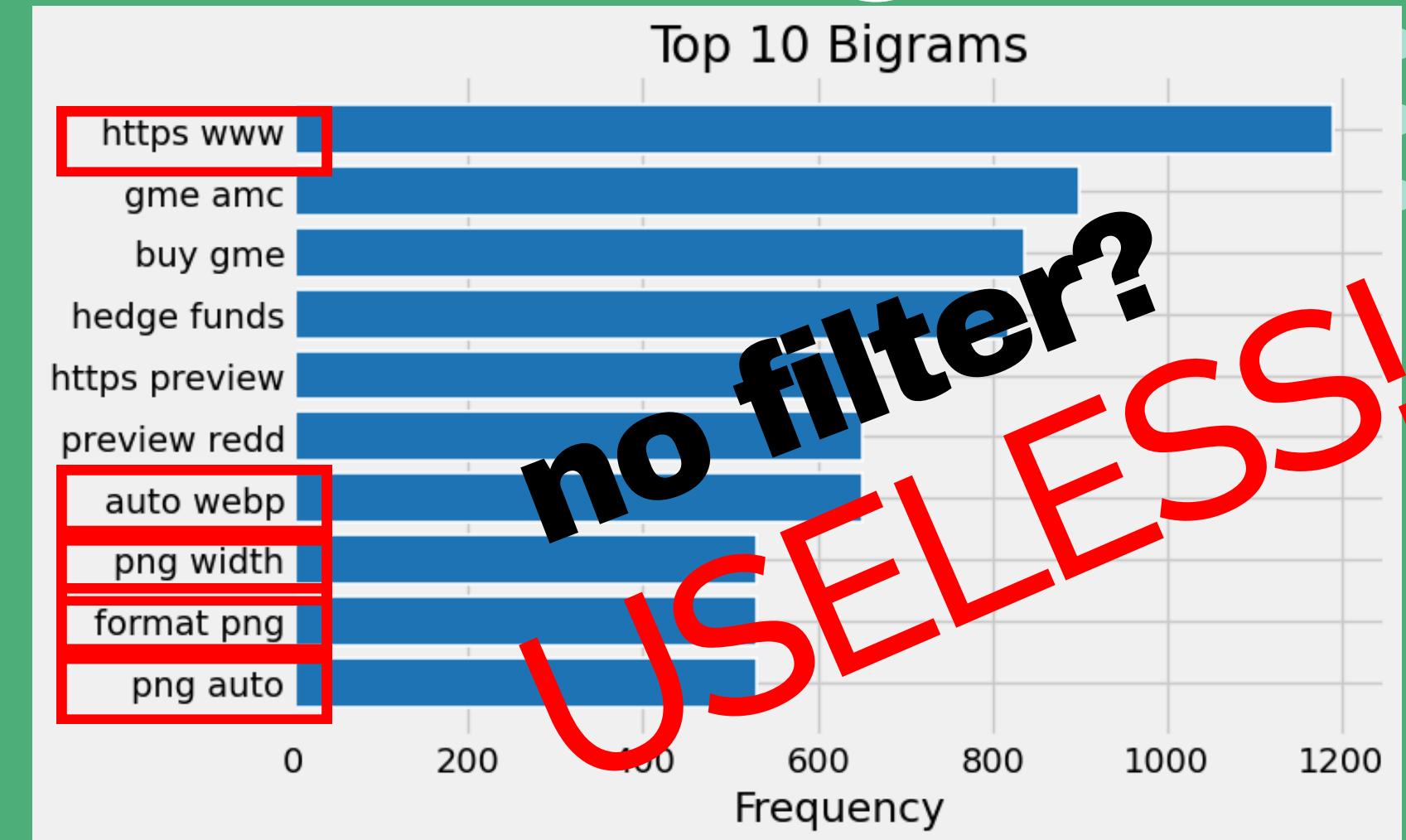
# Data Processing and Cleaning

- Filter Junk Tokens from tweets

```
junk_tokens = ['https', 'www', 'png', 'webp', 'auto', 'preview', 'format', 'com']
clean_bi = bi_df[~bi_df['term'].str.lower().str.contains('|'.join(junk_tokens))]
```

- Aggregate Daily Metrics

```
reddit_daily = (
    reddit
        .groupby('date')
        .agg(
            posts_per_day = ('id', 'size'),
            avg_upvotes = ('score', 'mean'),
            avg_comments = ('num_comments', 'mean'),
            avg_sentiment = ('compound', 'mean'),
            avg_length = ('word_count', 'mean'),
        )
        .reset_index()
)
```



Simplified Analysis!

WHERE



# Difficulties and Limitations

## Twitter and Elon

api limits, paywalls  
post-Musk takeover



Elon Musk @elonmusk

Honestly, I did it for the memes.

1:22 PM - 6 Feb 2018

8,701 Retweets 83,278 Likes



2.4K 8.7K 83K

## Reddit Noise

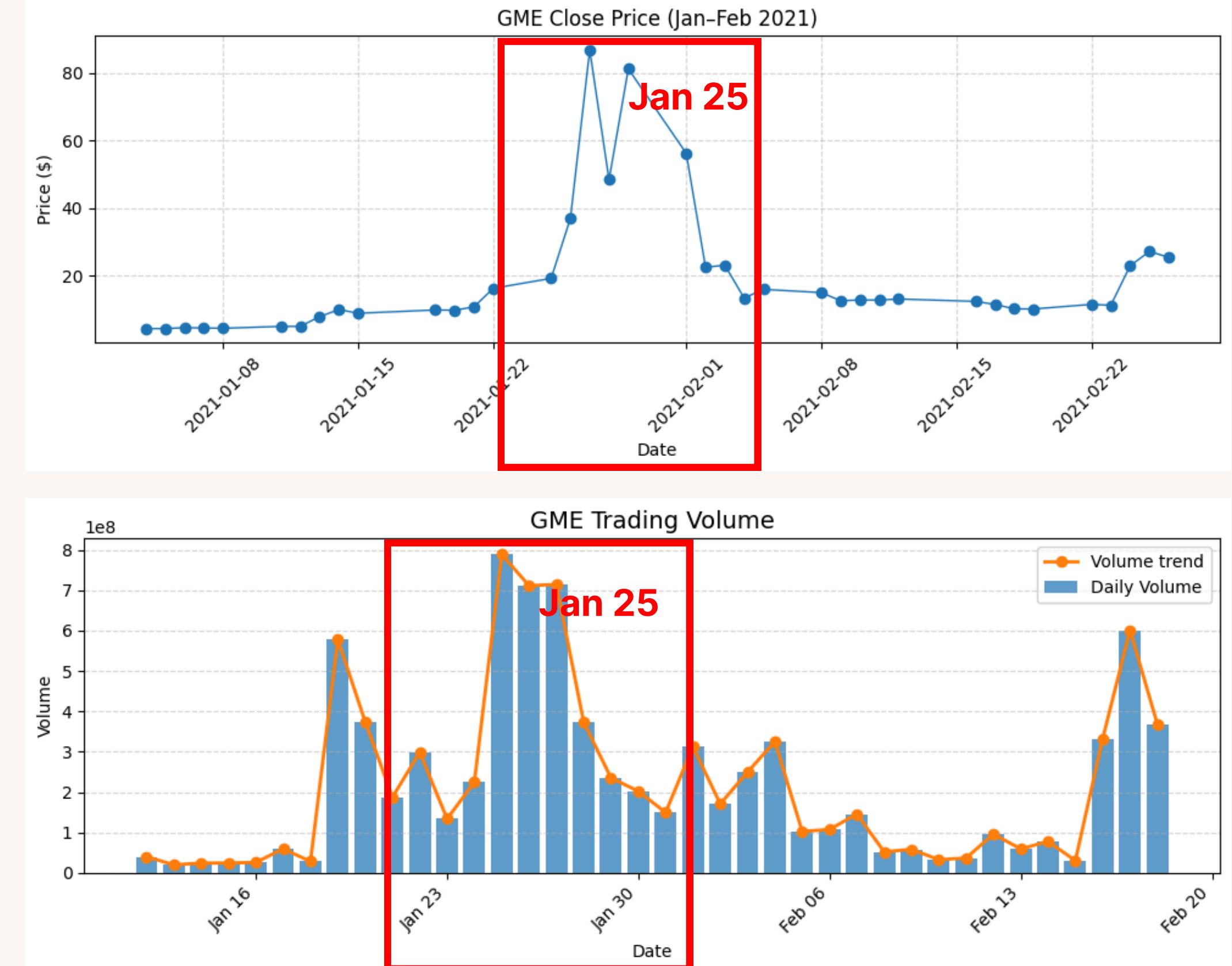
thousands of posts  
→ manual filtering

## Variation in Data

need to analyze  
different data from  
several platforms

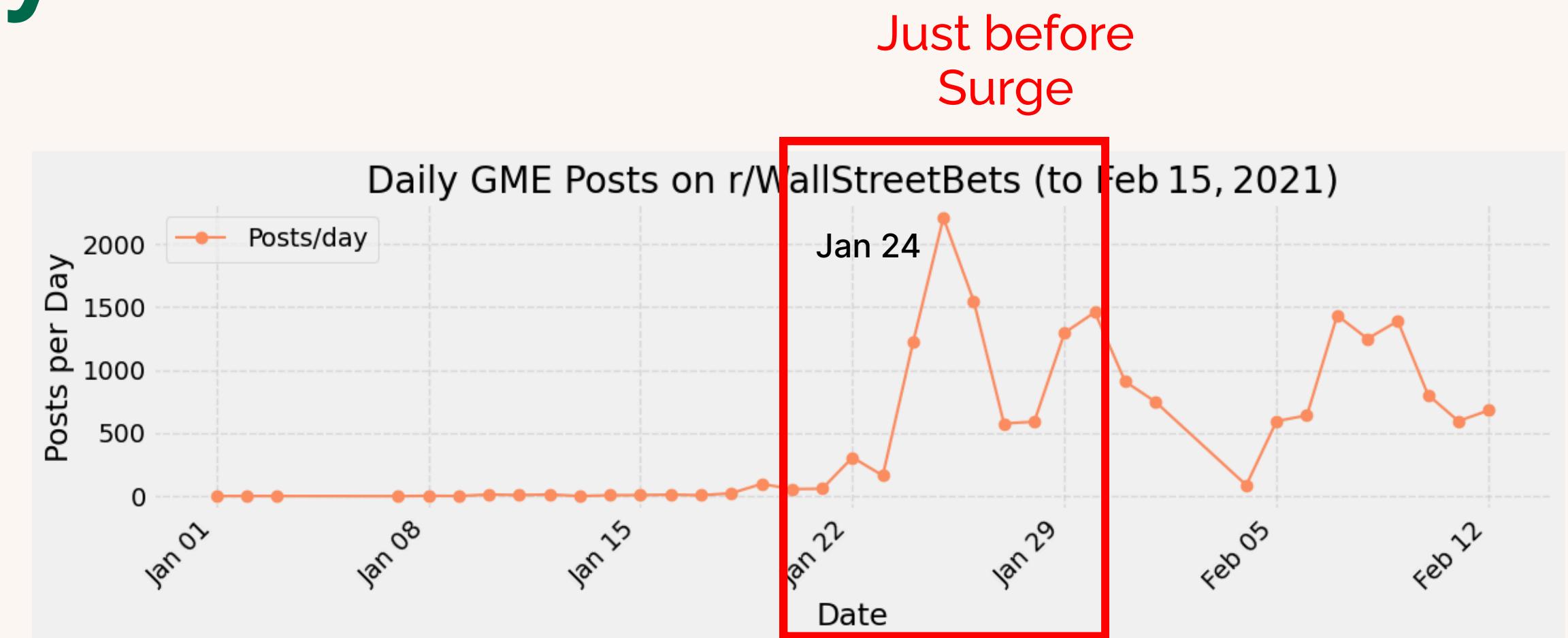
# Phenomenon

\$GME price surged from \$20 to \$90 between Jan 25-Jan 28), with trading volume exploding 26x



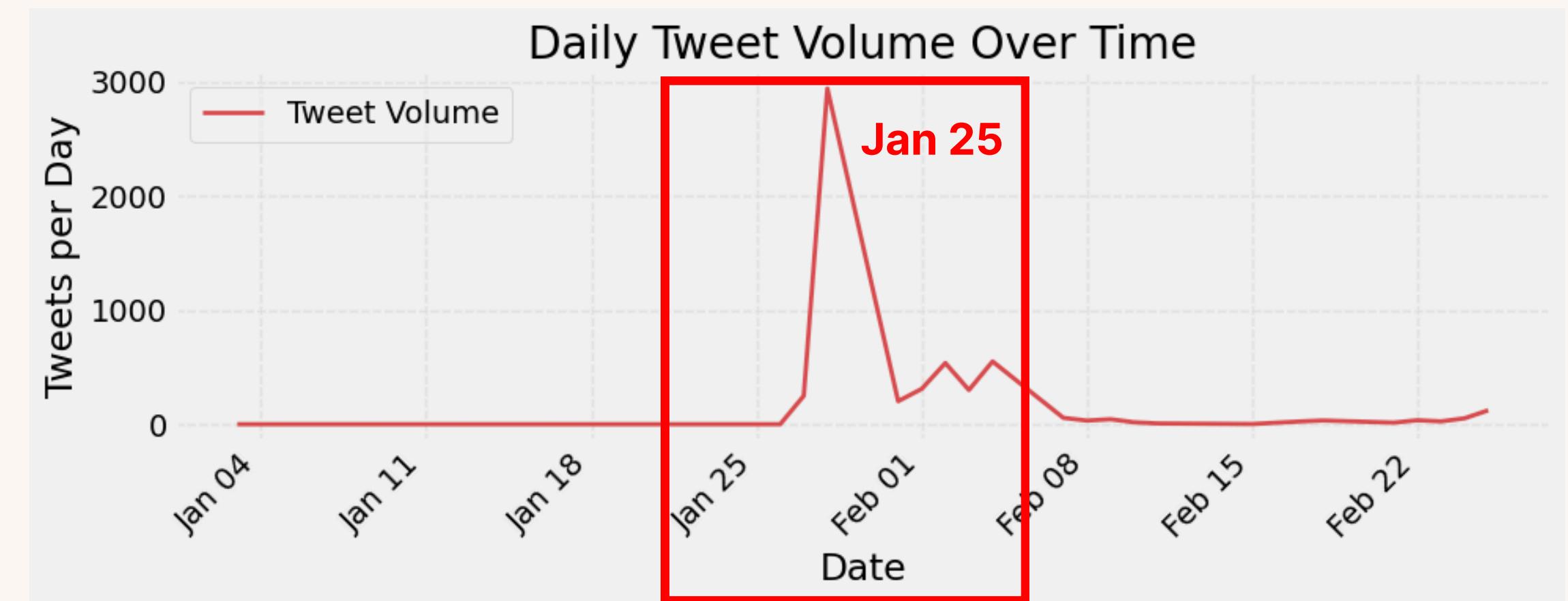
# Meme Activity as the Driver

**Reddit Surge:**  
Posts on the subreddit  
[r/WallStreetBets](#) spiked  
from 200/day to 2,200/  
day before the rally."



# Meme Activity as the Driver

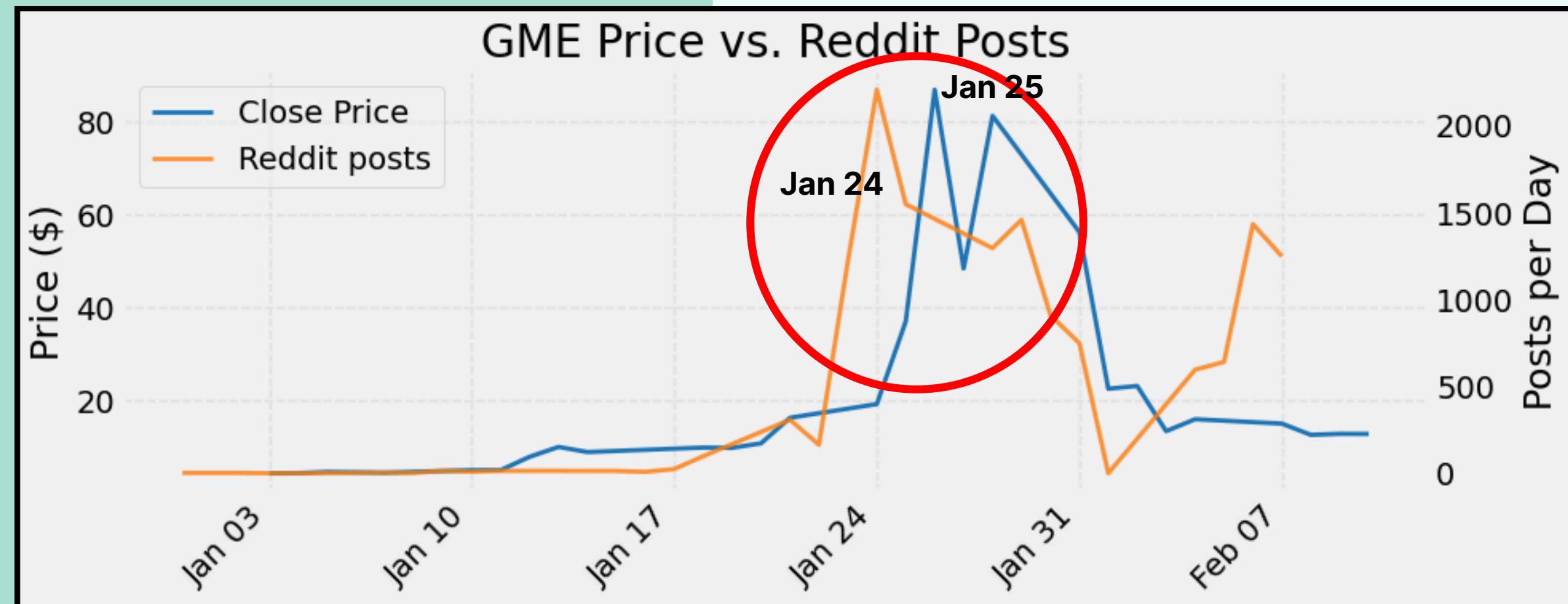
**Twitter Surge:**  
Tweets with \$GME,  
GameStop and similar  
keywords spiked during  
this time the same time.



# Combine and Correlate

After merging and standardizing the data,  
clearer patterns emerge

Chatter spikes come before  
price moves. Notice how,  
Reddit posts for GME  
increased on Jan 24, and GME  
price spiked on Jan 25.



**Calculation of correlation  
between yesterday's posts  
and today's return showed  
an approximate 0.62  
correlation.**

```
# correlation between yesterday's posts and today's return
r = df['posts_per_day_lag1'].corr(df['return'])
print(f"Pearson r = {r:.2f}")
```

**Strong Correlation!**

# Roaring Kitty

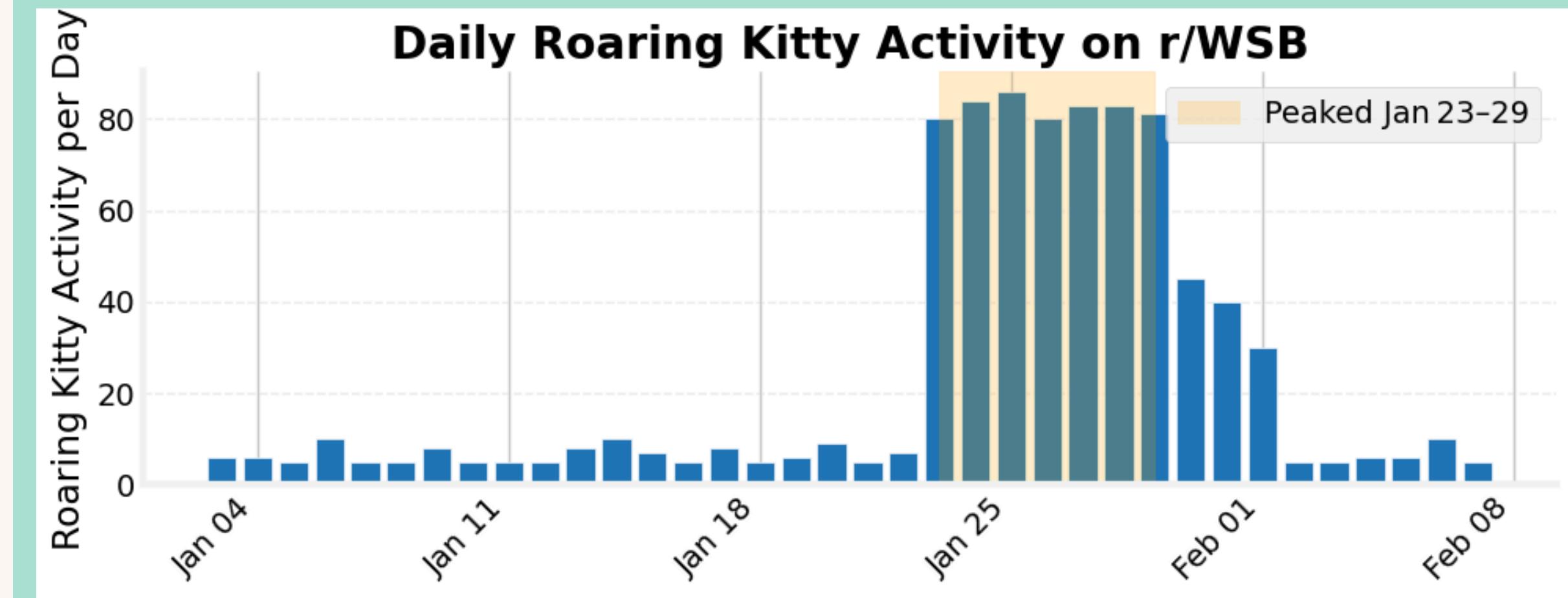
- Real Name: Keith Gill
- Online Aliases:
  - Reddit: DeepF\*\*\*\*\*Value
  - YouTube: Roaring Kitty
- Background:
  - Had been analysing GameStop \$GME for prolonged time.
- Role in Jan 2021:
  - People trusted his analysis.
  - Sparked the initial surge in chatter and buying activity



# The Roaring Timeline

- Baseline activity hovered around 5-10 interactions/day (Jan 3-22)
- Peak surge to roughly 85 interactions/day during Jan 23-29
- Relatively sharp drop to from Jan 30 onward

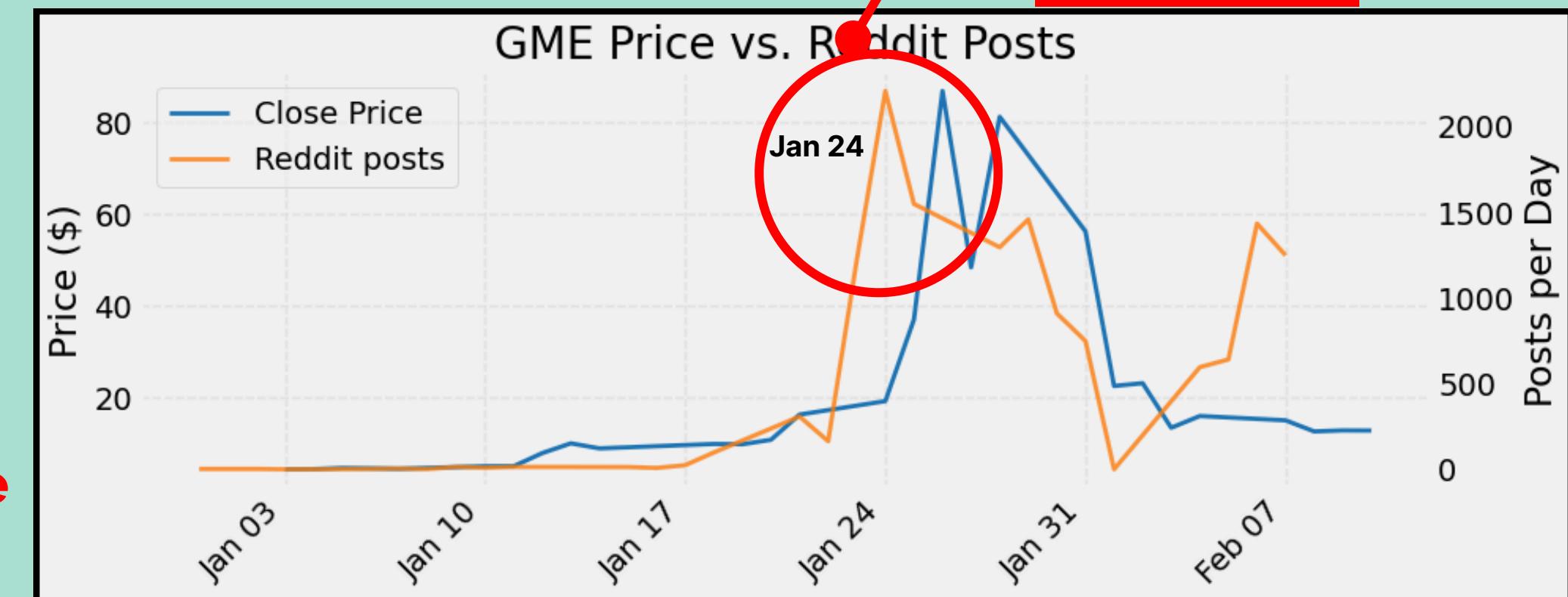
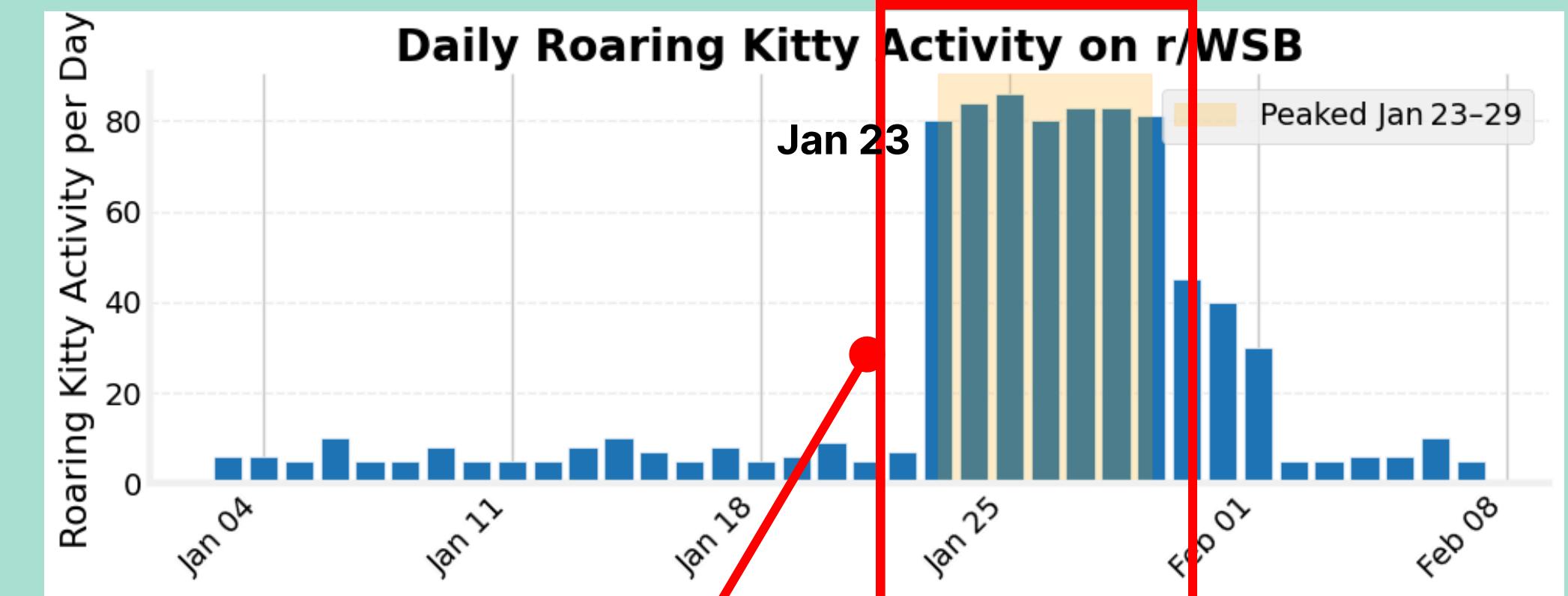
**Daily Roaring Kitty Activity on r/WSB**



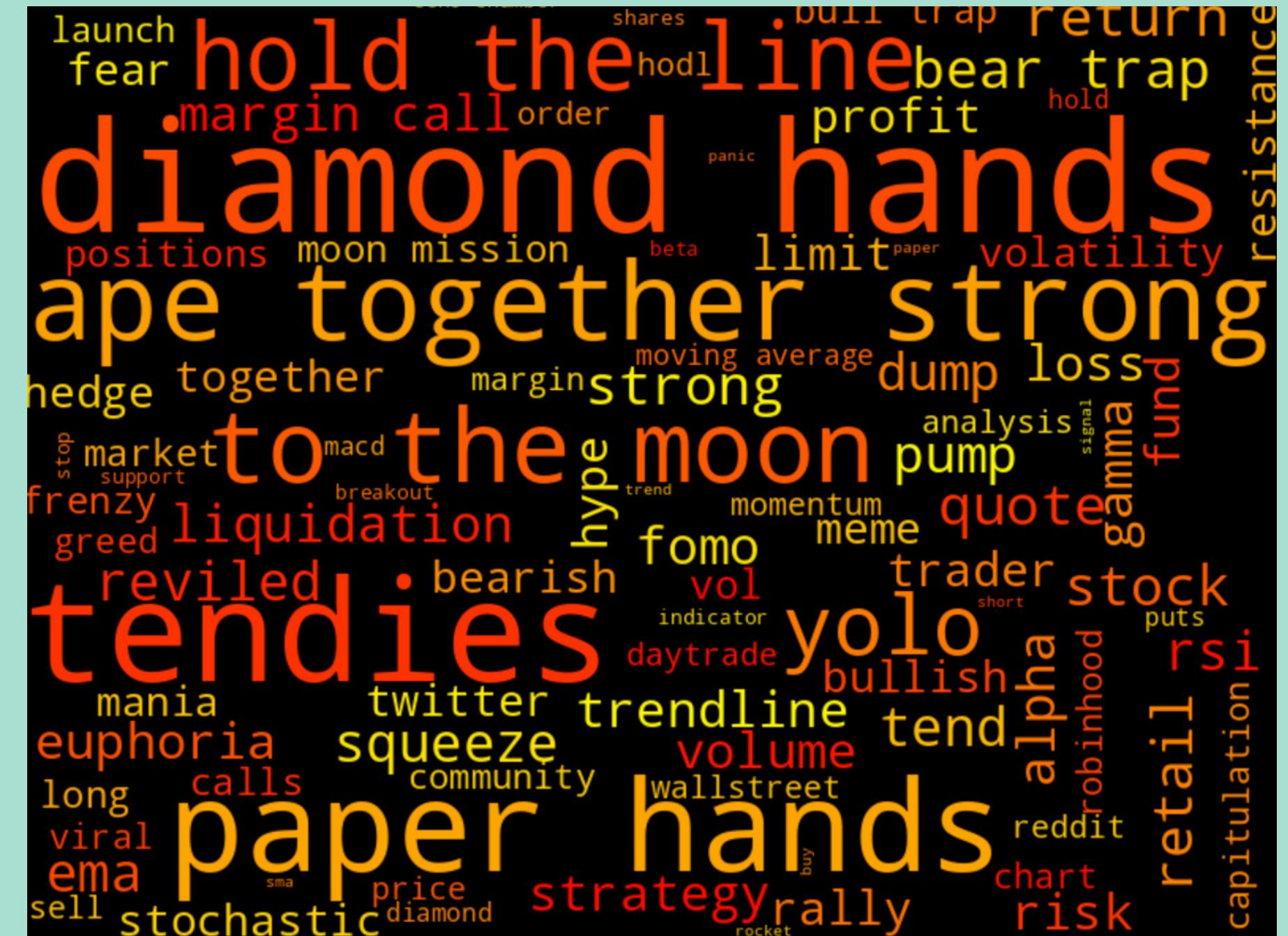
# The Roaring Timeline

- RK's posts spiked on Jan 23rd
- Reddit chatter spike on Jan 24th
- \$GME Price also exploded from \$20 to \$80+ (closing) during that time.

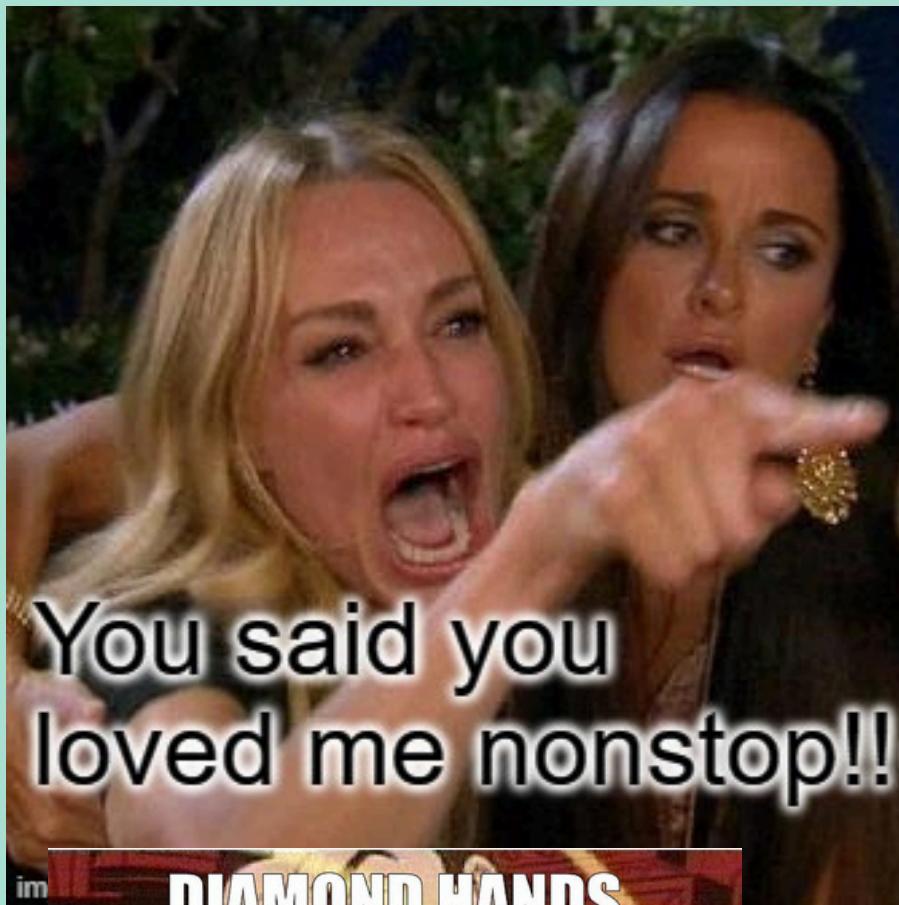
**Roaring Kitty -> Reddit hype  
-> price surge**



# Key Themes in the Chatter

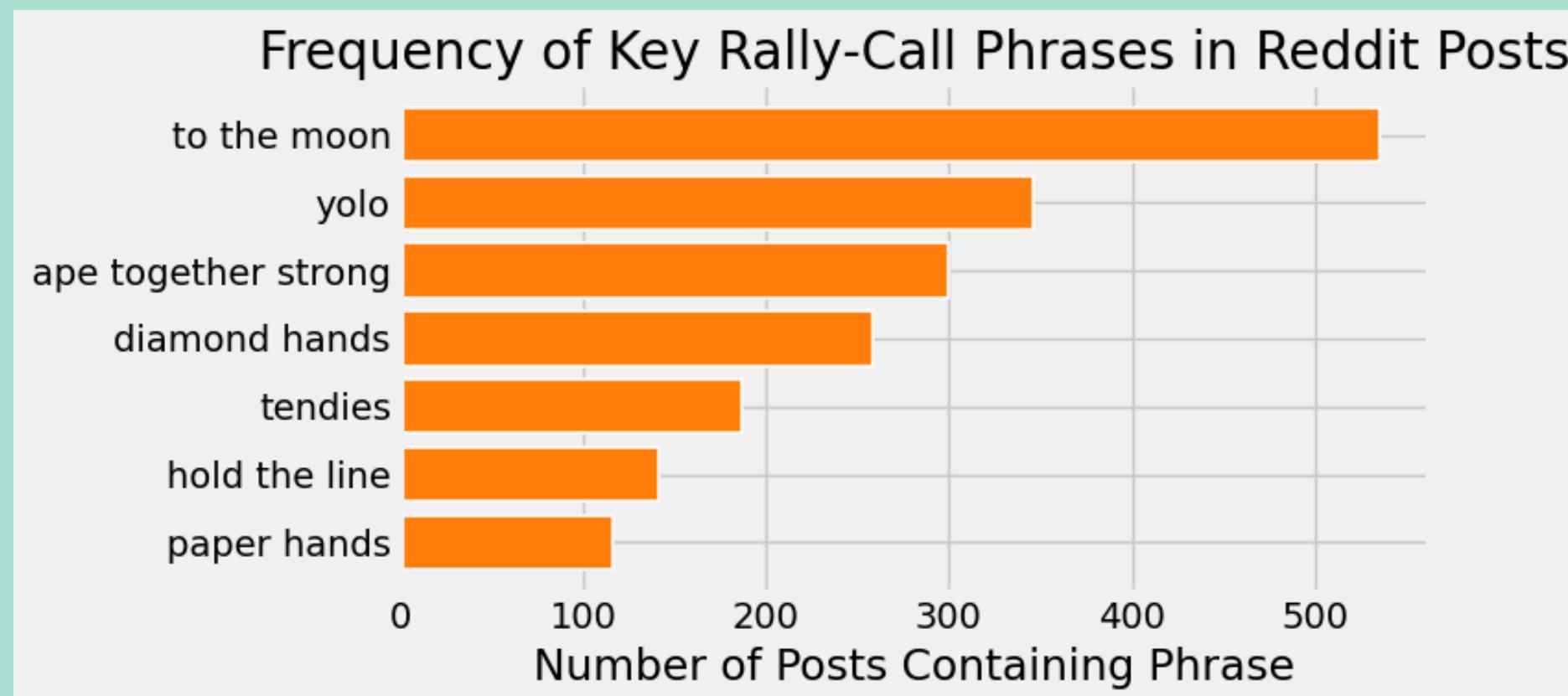


# Key Themes in the Chatter



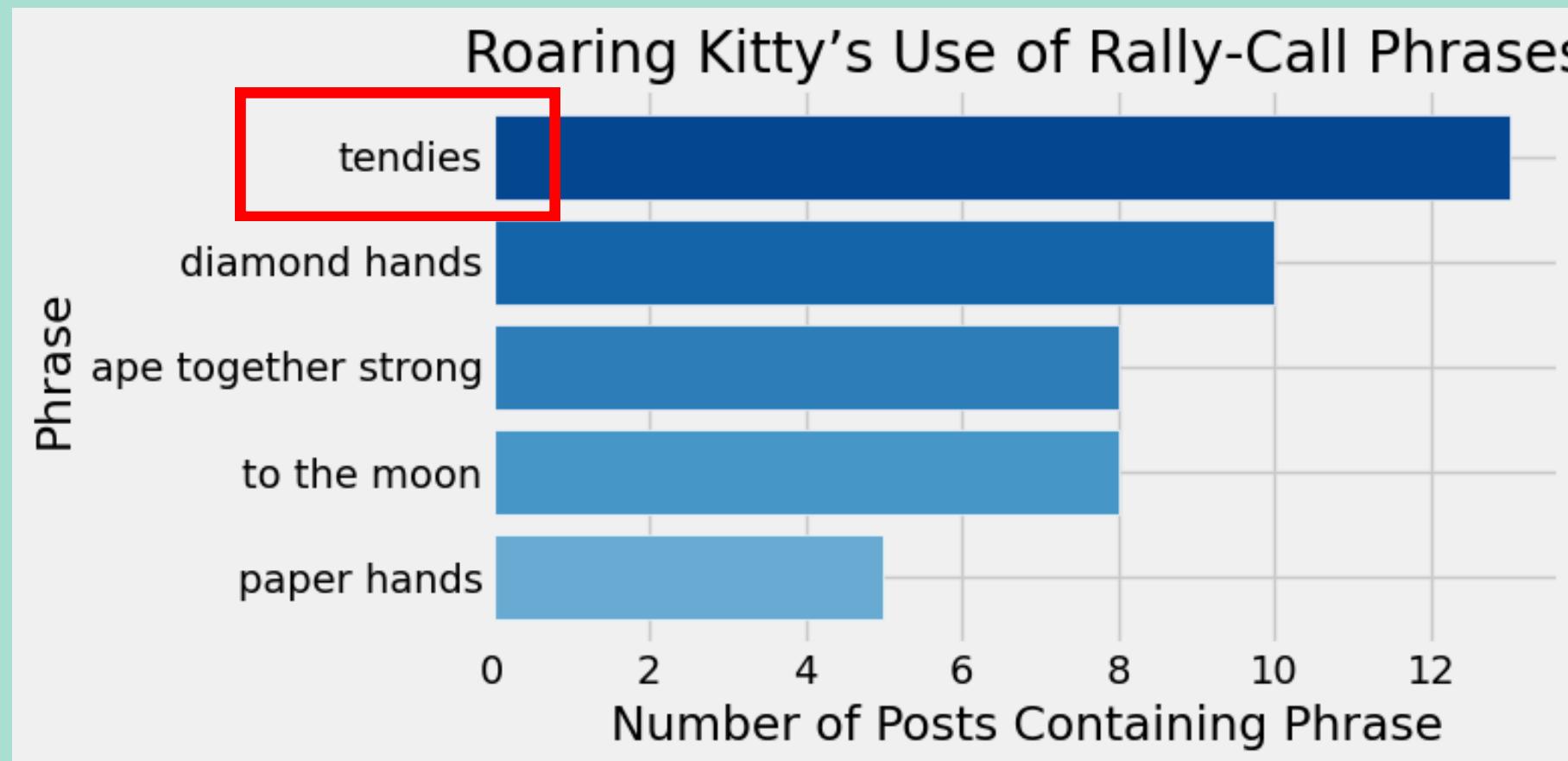
Widespread sharing of the memes and virality of the stock led to its price surge.

# Key Themes in the Chatter

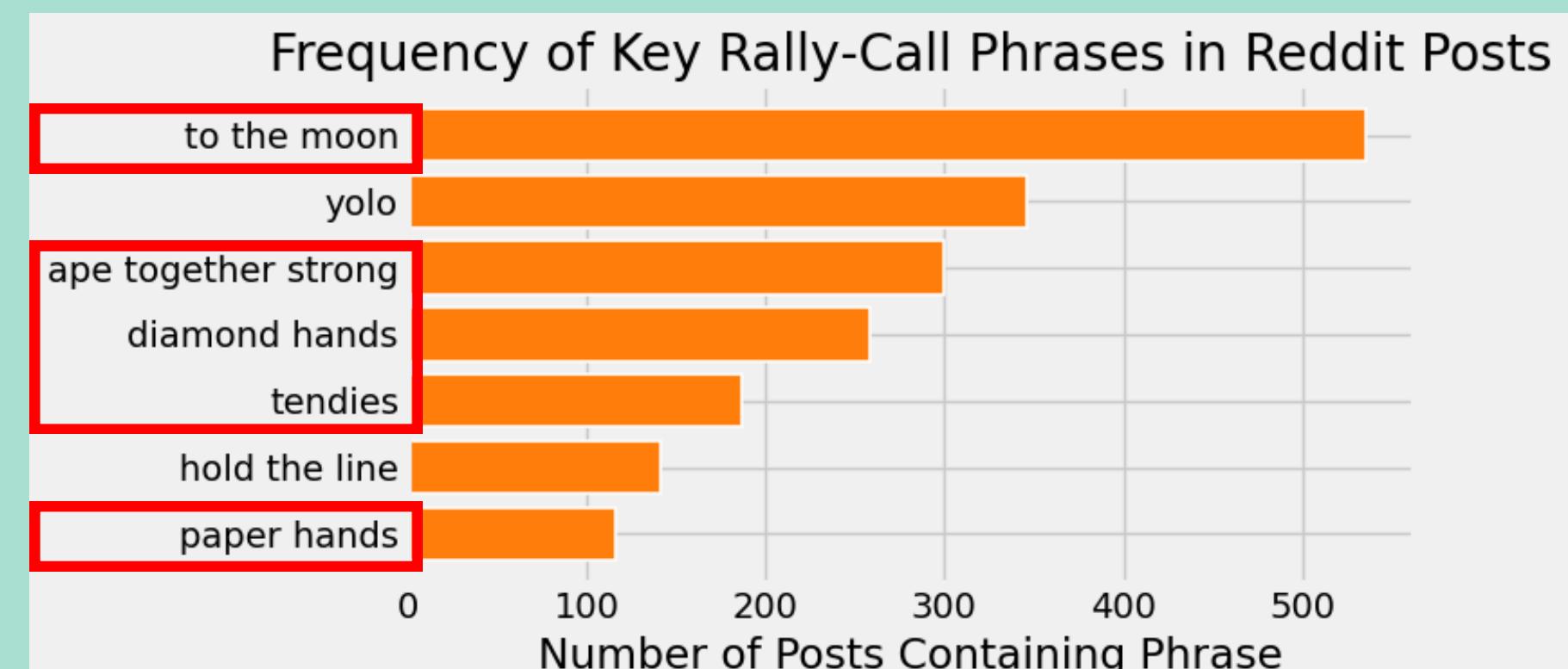


Reddit users used these rally calls incorporated with memes to hype the stock and lead to its surge.

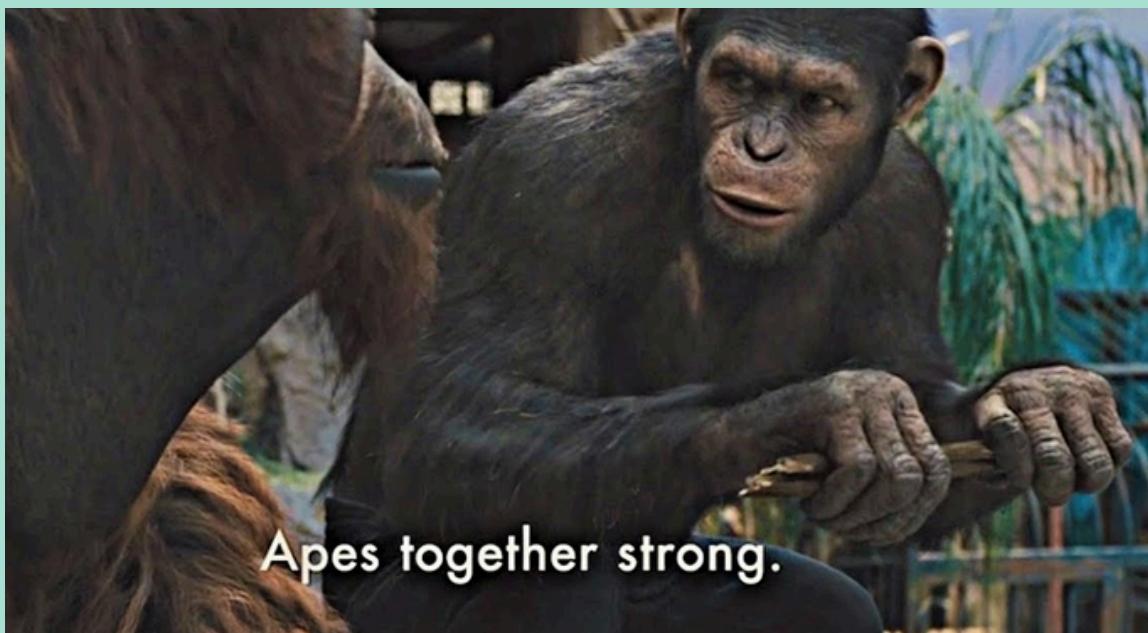
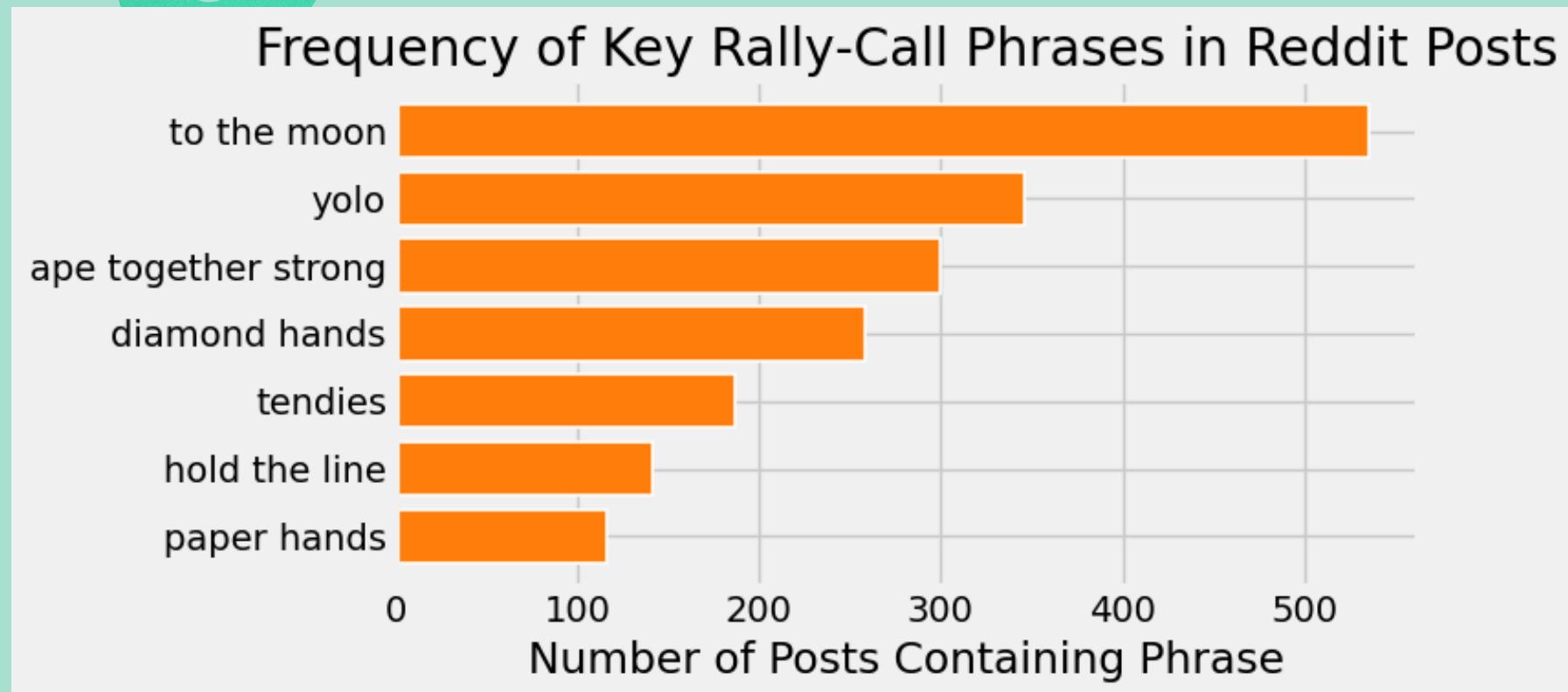
# Key Themes in the Chatter



Roaring Kitty's Phrase frequency chart shows 5 of the highest frequency meme phrases. He also had a YouTube video title: "Show me the tendies"



# Mob Mentality



Risk Aversion -> "**Paper Hands**"

- People were cautious and scared at first.

Greed & FOMO -> "**To the Moon**" & "**Tendies**"

- Profit dreams and FOMO = more aggressive buying.

Solidarity -> "**Ape Together Strong**"

- Shared identity and mutual encouragement transformed individuals into a unified group.

Herd/Mob Mentality -> Collective Action

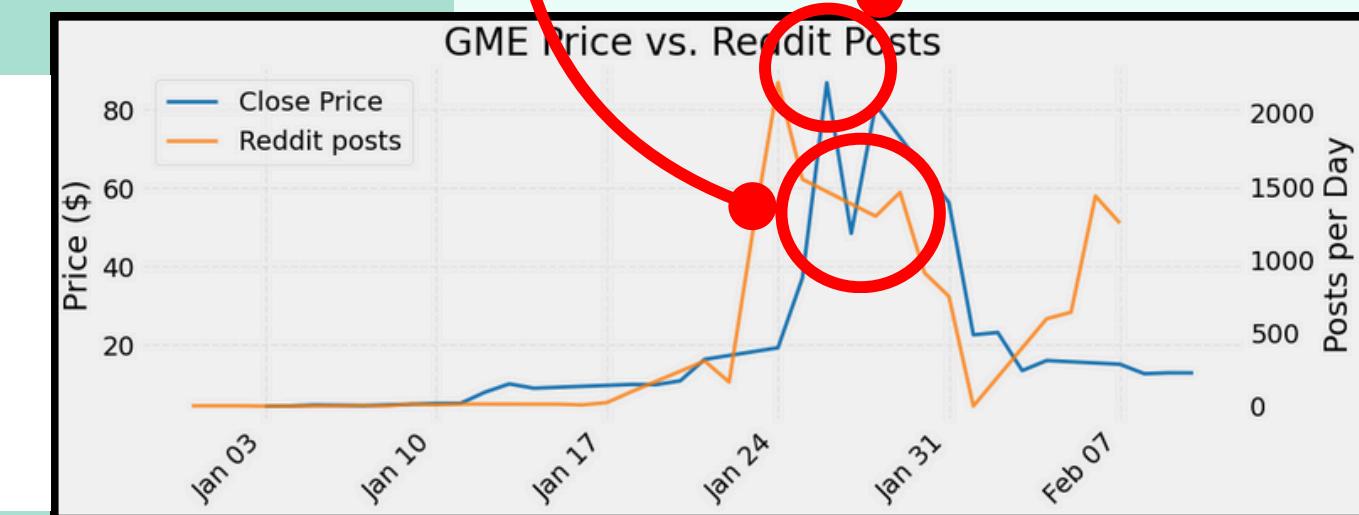
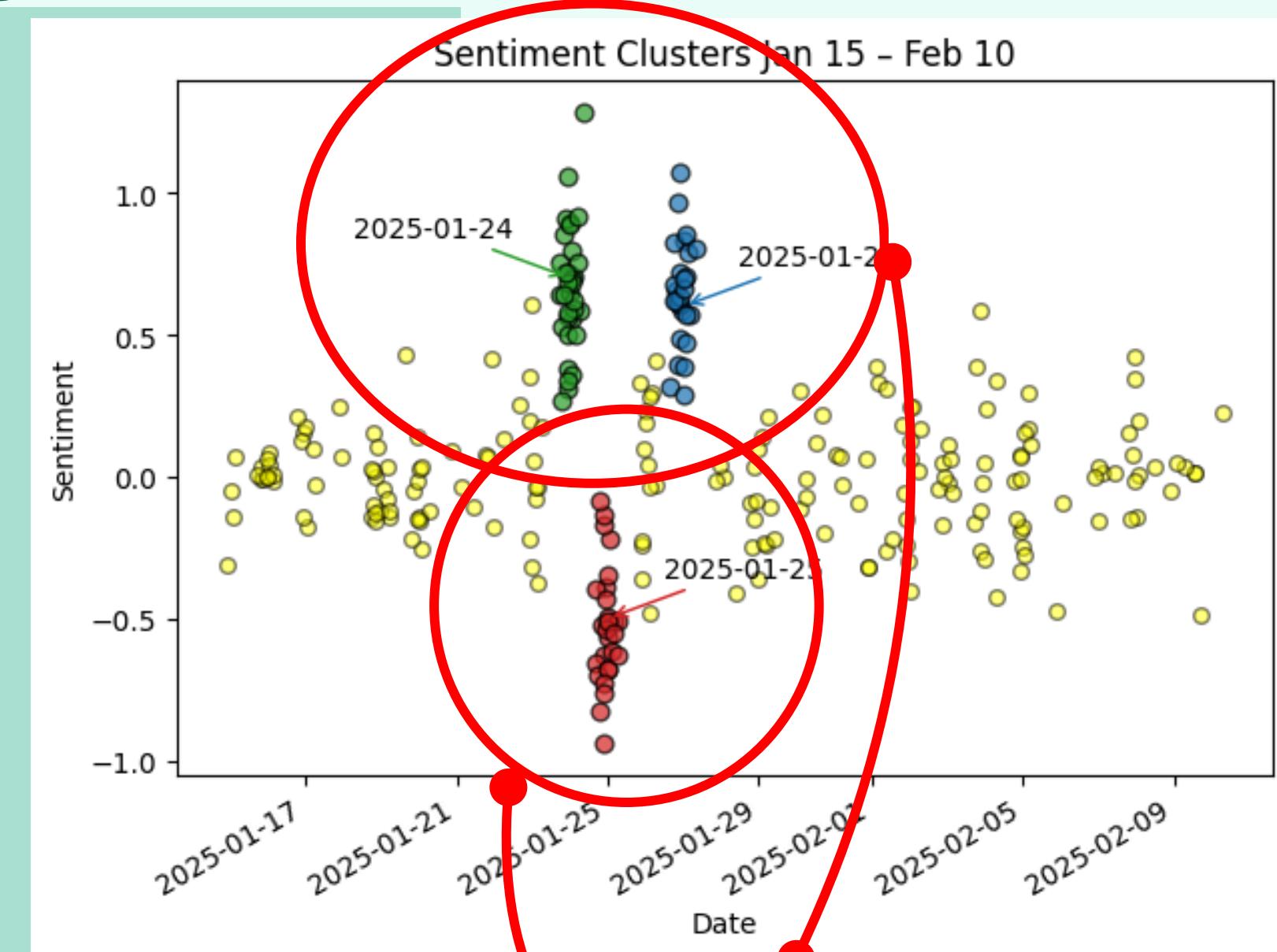
# But did it actually affect?

After merging and standardizing the data, clearer patterns emerge

Analysis of the sentiments dataset showed that:

- Before a spike in price, the sentiment of the Reddit posts was positively-high (at 24<sup>nd</sup> Jan and 27<sup>th</sup> Jan).
- Sentiment was severely negative before a purge in price.

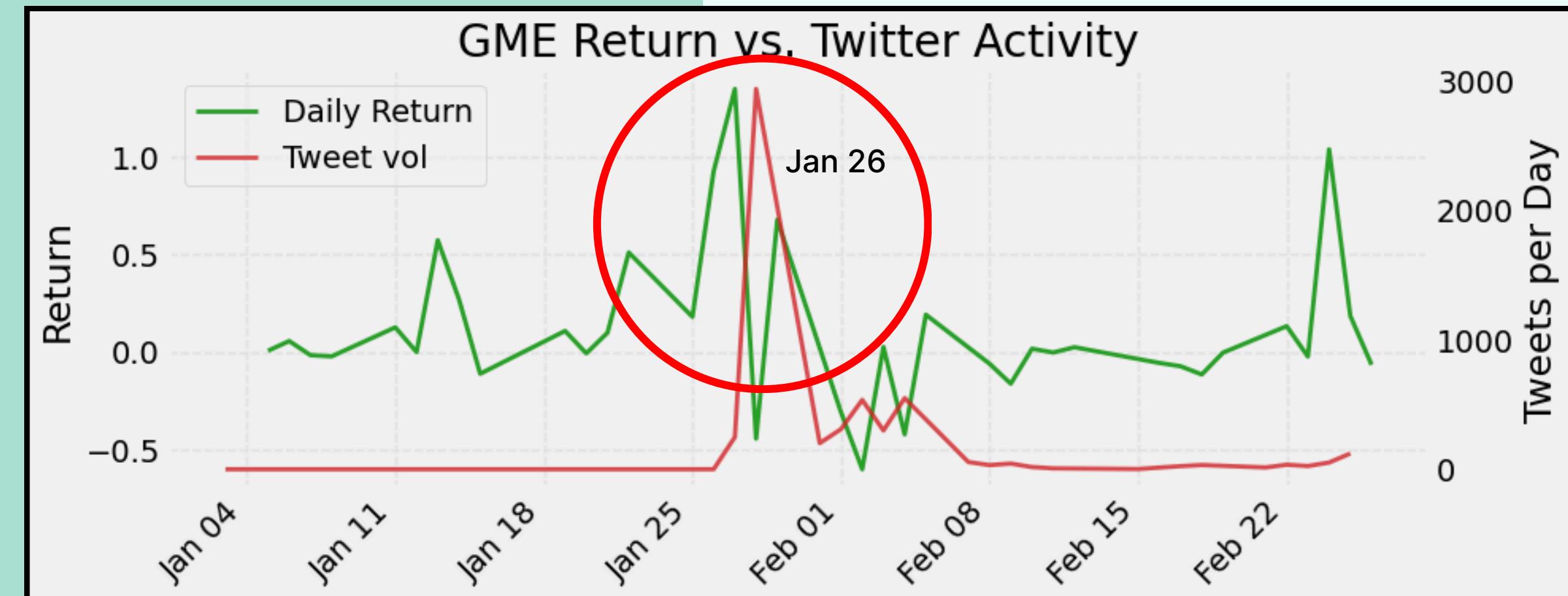
**Correlation Coefficient  $\rho = 0.68$**  => New members unsure and getting upto speed, while the herd was making teamed up moves.



# An Interesting Observation

Notice there was a return plunge after the initial Jan 25 spike.

But here, **the elevated Tweet volume on Jan 26<sup>th</sup>, Jan 27<sup>th</sup> led to price and return rebound again around Jan 28th.**



# Answering: Impacts Summary?

- Meme Activity Predicts Market Moves: Chatter spiked just before market surges.
- Roaring Kitty and herd mentality of people turned memes into coordinated action.
- Sentiment Drives Action: Positive sentiment correlated strongly with price increases ( $p \approx 0.68$ ).
- Virality Sustains Momentum: Meme waves moved from Reddit to Twitter, extending impact across days.



# Recommendations for the Future

Real-time tracking of forums and post volume and sentiment to predict market spikes/drops.

Community Education Initiatives Provide clear risk-management tutorials to incoming retail users to temper herd-driven losses.

Work with forums and exchanges to flag extreme meme spikes and consider tempered margining during high-volatility social events.

Hussain Wasif Latif  
UID: 3035834577

Thank  
you very  
much!

