**RQ: *How does emoji usage affect sentiment intensity in meme stock Reddit communities?***

***Sub-Q:*** *Can the presence or absence of certain emojis in post titles be used to predict sentiment?*

A range of affective functions that (could) complicate lexicon-based sentiment scoring (VADER):

* 🚀 - bullish optimism, upward momentum
* 🍌 - refers to primates/ ape identity, playful, absurdist inside joke
* 🦍 - collective identity, positive, solidarity , group resilience
* 💎 - strength and commitment emotional and financial toughness, community value, urging others
* 🔥- chaos, market “burns”, volatility, excitement? ironic?
* 🌕 - “To the moon”, upward movement in stock price, rally enthusiasm
* 💥- explosive, big announcements, can be ironic?
* KWIC - a method where you look at a specific word or emoji and examine how it appears within actual sentences or comments, to understand how its used rather than just how many times it appears - we can see if its used seriously sarcastically, ironically
* Emoji only scattertext?
* Does the same emoji appear in both positive and negative comments?
* Are some emojis used ironically or sarcastically? Can we find this?
* How does the sentiment score match / clash with our reading of the tone?
* Are emojis used alone, repeated, or in sequences (e.g. “🚀🚀🚀”)?
* Is it possible to visualise emoji use vs sentiment score in a Scattertext?
* Emoji use in online finance communities carries layered meanings (irony, community / collective identity\_ that may conflict with traditional sentiment analysis tools - layered meanings if we use sentiments not emotions
* DH: combines computational text analysis with close reading of cultural symbols in online discourse
* Can we extract more examples for specific emojis? Is this the only data in the r/GME report available? Can we use r/ superstonk instead? Is this the same one we used for the sliding score?

Titles across 4 subreddits (no comments)   
Frequency   
Emojis not rated in VADER

We decide balance / polarised /

**APPROACH:** How does emoji frequency correspond to sentiment scores for text (without emojis). Can the presence or absence of certain emojis in post titles be used to predict sentiment?

**STEPS FROM TODAY:**

* **Pawel:** to look at bulk amount of emojis per title; look at finding out what is the difference between positive/negative sentiment depending on emoji volume?; to generate emoji list with numbers; provide balanced dataset (**primarily look at balanced one**)
* **Yuleum:** to collaborate with Pawel on above listed tasks (stats)
* **Rose:** draft presentation slides; start inputting information available
* **Pawel**: make sample of only titles with Emojis ( balanced and full lists)
* **Ellie, Alma, Tomas, Rose**: close reading (for irony, as sentiment analysis tool cannot detect this) in sample from Pawel

**PRESENTATION ELEMENTS:**

* Introduction and brief background to project (context)
* Short explanation of project’s objectives (RQ, thesis statement)
* Workflow
* Description of dataset (incl limitations - can discuss need for qualitative analysis here which can lead into methods description next)
* Description of methodology (incl work packages/division of tasks) and Criticism of digital tool
* Discussion of findings (analysis and results interpretations)
* Description of team (interdisciplinarity relevant here

**r/Superstonk**

* “The more and more they don’t let this thing pop....Added 6 more! 🚀🌙🚀🚀🚀”  
  VADER score: -0.3164  
  > VADER reads as negative, perhaps due to “don't let this thing pop” which linguistically could signal restriction or frustration. But post also expresses optimism: OP “added 6 more!” - shares, as I see with the accompanying screenshot of share purchase on Fidelity - combined with the rocket emojis means bullishness and excitement. In r/superstonk context the tone is a bit sarcastic or defiant - while market manipulators suppress stock, OP doubles down in defiance. Presence of moon and rocket emojis lifts the sentiment, it is positive and defiant, signalling commitment rather than restriction.
* “Please let this be real 🤪🤪🎉”  
  VADER score: 0.3595  
  > VADER reads as positive, but there is also a tentative, doubtful edge with the tongue emoji. Presence of emoji pokes fun at OP for being hopeful, so it perhaps misses some doubt.
* Guess what came in the mail today? Had asked my broker to send it after I voted. A PIECE OF MOTHERFUCKING HISTORY. BUY HODL VOTE 🚀🚀🚀🚀🚀🚀”

VADER score: -0.5267  
> VADER reads negative, perhaps because of profanity? Within subreddit context profanity could be celebratory, empowering. Shareholder activism, voting, suggests collective affect. Misses meaning of HODL.

* VADER often misclassifies sarcastic or ironic tone. Posts can be rated as negative, even though contextually they express defiant optimism. Sarcasm in the subreddit is also frequently a form of collective, affective solidarity or ‘performative’ belief, not actual negativity
* High frequency emojis like 🌙, 🚀, 💎 consistent with positive sentiment, can contradict what VADER extracts from text - perhaps more reliable than the lexical context in expressing a post’s emotional tone?
* VADER only reads text (please correct if wrong), sometimes accompanying screenshots of Fidelity purchases reinforce affective intent that is invisible to VADER
* For sentiment analysis in these subreddits, very particular knowledge of community context is needed- inside jokes, specific emojis, etc. Even just stating you “added more”, or are “still holding” is not just a neutral investment language about what you are doing, but within the meme stock context they become bullish and defiant, collectivist. Resistance against institutions, commitment to shared cause.

**r/GME**

* After market 🔥🔥🔥💎💎🖐🖐

VADER score: 0.0516

> Only occurrence of fire emoji in sample from GME titles - VADER reads as positive however it’s a low positive score which somewhat aligns with the consistent negative scoring for all other occurrences we’ve seen

* 💎🙌 is not a catchphrase. Its a movement or better yet - a team of complete and total retards. My retarded brothers and sisters in this war against bu bu bu billionaire weasels will earn 1 of these jewels when we get to the moon

VADER score: -0.5703

> VADER marked negative; likely due to the use of a negative (‘not’), and the lexical items ‘retards’, ‘retarded’, and ‘war’ which each carry heavily negative sociolinguistic connotations; ‘retards’ and ‘retarded’ here appear to be used by OP in a way that signals solidarity between members of the community which is a message that is strengthened by the use of the 💎 emoji, and ‘war’ is further used to signal the banding together of those members for a common goal

* HOW MANY TIMES DO WE HAVE TO TEACH YOU A LESSON, OLD MAN. LETS FUCKIN GOOOOOOOO 🚀🚀🚀🦍💎💎💎

VADER score: 0.9774

> VADER marks highly positive despite use of profanity (*compare with r/superstonk example of profanity usage and negative scoring*); perhaps the use of ‘goooooooo’ is recognised as positive rallying and the multiplied use of 🚀 and 💎, combined with 🦍, strengthens this expressed sentiment

* Don't forget to set your price alerts retards 💎🙌🚀🚀🚀🌕

VADER score: 0.9563

> Interesting that the 'retards' here doesn't seem to impact VADER’s scoring negatively (*compare with the above listed title)*; 💎, 🙌, 🚀, and🌕emojis appear to boost positive sentiment where the lexical elements don't necessarily communicate an overly positive sentiment; underlying sarcastic tone combined with a sort of endearing reference to community members - solidarity building

* Retarded Ape repost coz barcode was showing. 🚨 🚨 🚨 DRS DRS DRS DRS DRS DRS DRS DRS DRS 🚨 🚨 🚨

VADER score: -0.5719

> VADER marks as negative; while it appears that OP is reposting due to a mistake in the original and this could be read as negative, the presence of the 🚨 emoji and the repeated ‘DRS’ seem to be intended to communicate urgency with respect to a calling for registration of shares; repeated phrase serves to rally support and signal commitment to a common goal/cause + reinforce solidarity. VADER seems not to pick up on this positive message and instead assigns a negative score, possibly due to misinterpretation of the context assigned to the alarm emoji by this community

**r/gme\_jungle**

* Good morning my SSJ🦍s!!! The \*\*MOASSIVE\*\* 💎🙌 \*\*BOMB\*\* is growing!!! Keep it up! 🙌🙌🙌
  + Vader score: -0.3614
  + Yet it seems clearly positive, because it means that it is growing incredibly and that they should continue
  + Is it the word bomb? Can it not understand the MOASSIVE meaning?
    - Problem with community specific internet lingo?
  + Chatgpt analysis:
    - “Extremely positive, energizing, and community-focused, with strong undertones of confidence in a major financial event” +0.900
* A glimpse at our future home 🌚 (does anyone else see confirmation bias everywhere)
  + Vader score: -0.1027
  + Confirmation bias I understand it means for others and not for them?
  + Emoji to the moon and saying it is their future home implies they are going to succeed in a way?

Don’t worry, everything is fine 🔥🔥🔥🔥🔥🔥 Just $1.35 billion in RRP 🔥🔥🔥🔥🔥🔥

* Vader: -0.9774
* Chatgpt: “Negative, expressed through sarcastic irony” -0.450
* What is sentiment? We are working with different baselines
* Improper tool for the task = finding
* VADER can identify sentiment but we are trying to identify context(?)
* VADER needs to be adapted to community context
* Assign new VADER scores based on our close reading

Findings:  
- Limitation of the dataset : abstract for dataset, post sentiments are improper because they have limitations, does not measure sentiment of what is posted

* Limitation: removed comments, body
* Data-driven initial just to figure out what is there
* Suggest how VADER can be improved - what Pawel wants to do with the scores
* VADER needs to be adapted to community context
* Critical of the datatset reuse potential
* Do not list in methodology what sentiment score was done on, post, comment, title
* 565k posts most deleted

Use examples with fire emoji ! combination between qual and quant

While it might be different in other field, in this field/community fire is negative

We are not challenging sentiment analysis as a whole but this VADER method

Can we try with chatgpt to assign a sentiment

New pipeline in assessing the emoji's sentiment score

Github: make sure we fully document the whole process, even things that seem too trivial; e.g. the limitation -> what we decided to do

Explain clearly

DIGITAL TOOLS:

Before concluding, we want to include a reflection on the digital tools that we have engaged with. Although they have guided our research, we have found several limitations in them that, due to time reasons, we have not been able to improve ourselves. However, we would like to also point out certain possible directions for future research in our database, and in similar ones. VADER has been proved ineffective at rating the emojis, or comments ridden with emojis, in our dataset. Very particular emojis and lexicon created by the community, like the one that we have discussed here, needs a more nuanced approach. We suggest two options. The first one we have quickly mentioned it. It would be a human approach in which a team of researchers would manually calibrate VADER with general guidelines in the emojis’ scores. Our second approach would be to explore other digital tools that might be more recent and therefore more advanced in sentiment analysis. To have a quick comparison, we used CHATGPT, which is easily accessible, and asked it to perform sentiment analysis in the following posts. The VADER score is above and underneath is the CHATGPT one. The first post has been scored negatively by VADER. However, as we can see from the diamond hands emojis, and how it says that they need to “keep it up,” the sentiment is actually quite positive. They are saying they are succeeding and need to keep holding the stock. CHATGPT in this case has a more suitable score. The second one is the most negatively scored comment from our sample. It includes the fire emoji, whose meaning and problematic we have discussed. Despite the comment having a positive sentiment, saying they made some winnings and that “everything is fine,” both VADER and CHATGPT give it a negative score. This would perhaps hint again at how a close reader that understands the context of the community might be needed to calibrate the tools.

Meme stock:

* Shares of a company that have gained viral popularity due to heightened social sentiment
* Usually due to activity online and particularly on social media platforms
* These online communities can dedicate heavy research and resources toward a particular stock

**Data**

* Reddit dataset comprised of four subreddits
* Approximately 1.6M posts
* Produced as part of a research project to study communication patterns of subreddit communities around meme stocks (*and their belief in using meme stocks to facilitate a global wealth transfer movement*)
* Key components of interest for our project
  + Titles (column C)
  + VADER sentiment scores (pos/neu/neg + compound scores (column D))

**Data Limitations**

* Original posts unavailable - titles & comments only
* Not clear whether VADER sentiment score is attached to titles, posts, comments, or all of the above
* Comment dates are unavailable (only titles) - this restricts our ability to accurately identify correlations between shifts in community sentiment and market changes

**Reflection Upon Digital Tools**

* Good guidance
* Several limitations that we could not mitigate (with the time/skills/resources at our disposal)

**CHATGPT Experiment**

* Out of interest - asked CHATGPT to perform sentiment analysis on two titles
* Easily accessible; not conclusive
* In both titles we see positive sentiment reflected but negative VADER score
* CHATGPT score improves sentiment reading
* 🔥x12
  + Most negatively scored title from our sample
  + CHATGPT improves score but still does not bring it up to a positive score = still reading 🔥 as negative
  + Further indicates the need for community context and manual annotation of emojis to calibrate tools for accurate scoring
* Title 1 sentiment: Communicates success and the need to keep holding the stock
* Title 2 sentiment: Made some winnings and that “everything is fine”