



America's Greatest Mystery

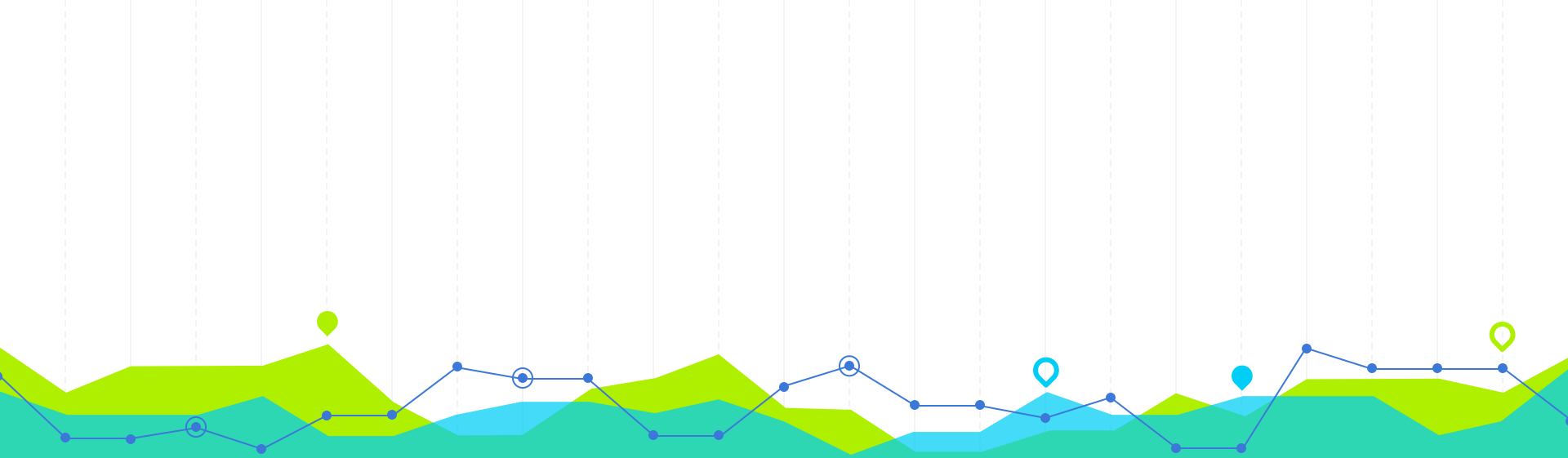
The Stock Market and Reddit

Carlos Jimenez, Endi Caushi, Owen Moreau
Dr. Salem Othman

r/wallstreetbets

**You have probably heard of this over
the past year.**





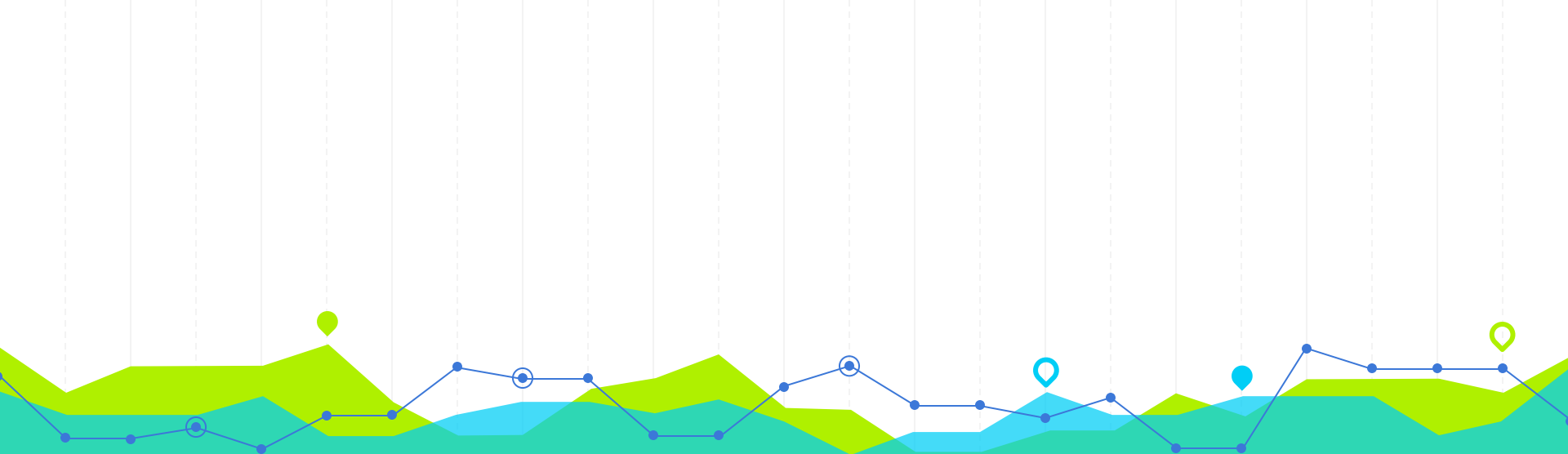
Background

1

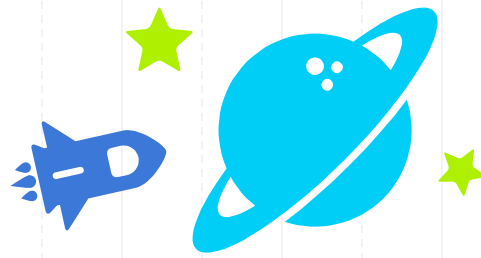
BACKGROUND

During the past year, we have seen a lot of hype around r/wallstreetbets. Hedge funds and market makers held short positions against stocks like GME, AMC, and TSLA hoping their value would fall. Retail investors on Reddit came together to perform the best short squeeze of history by buying these stocks in mass - shooting the value through the roof.





Objective 2

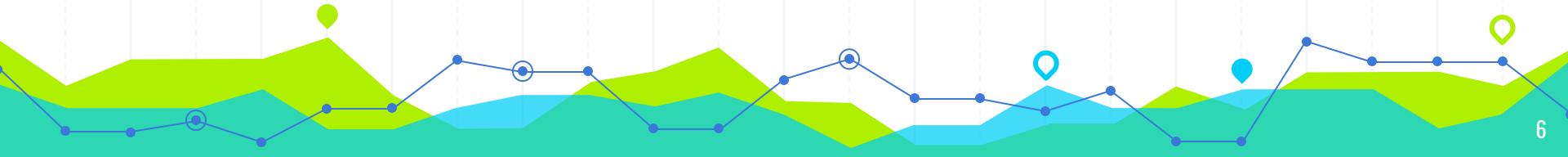


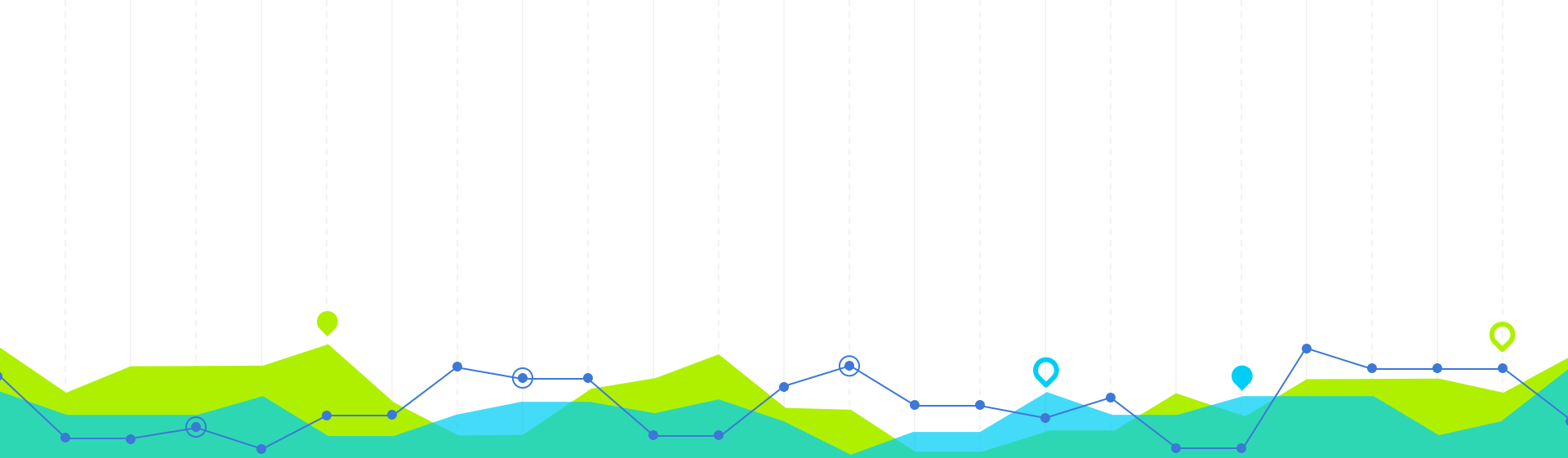
We aimed to solve the following:

Does an increase in discussion on Reddit lead to an increase in stock value?

Is the stock trading volume influenced by retail investors generating hype on Reddit?

How impactful is Reddit in assisting retail investors in beating powerful market leaders like trading firms and hedge funds?





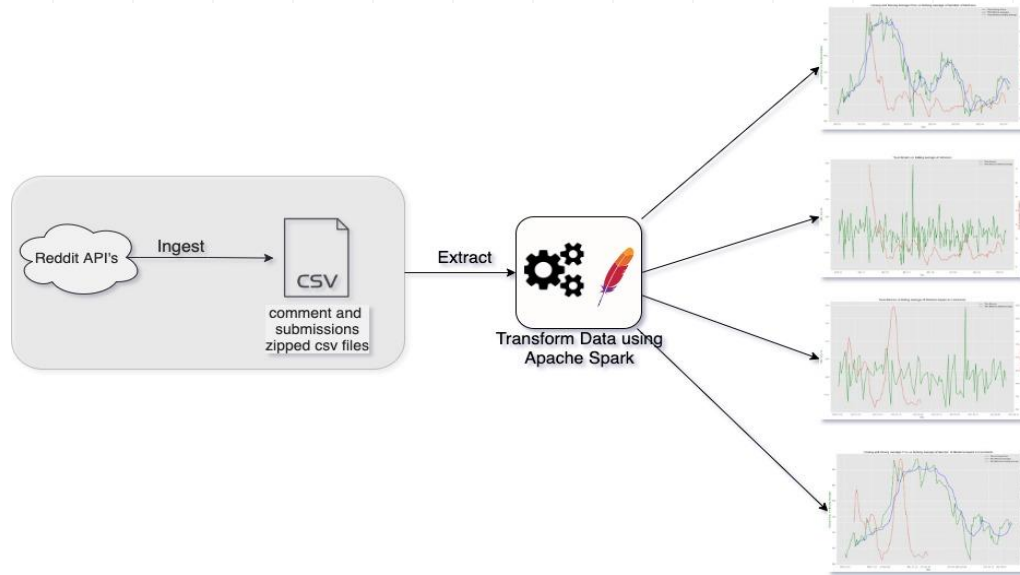
Implementations and Results



3

IMPLEMENTATION

Reddit submission and comment data was pulled from the Reddit API.



Spark was used to manipulate and analyze this data to see if it solved our objective questions.

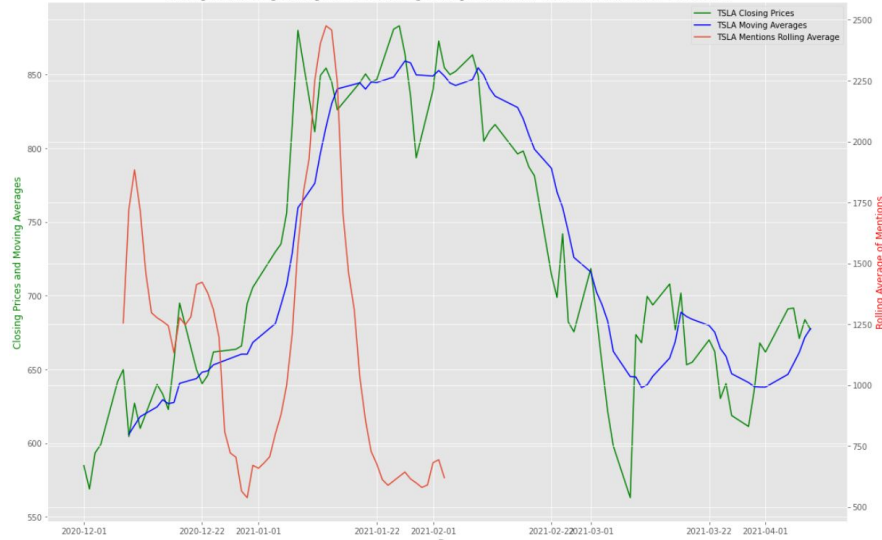
IMPLEMENTATION PART 2

- Gather Reddit submission and comment data through Reddit API - stored in csv.
- Use Spark to efficiently determine the number of mentions per day for each stock (GME, AMC, & TSLA).
- Convert to Pandas Dataframe and graph alongside stock price data from the same time and analyze results.

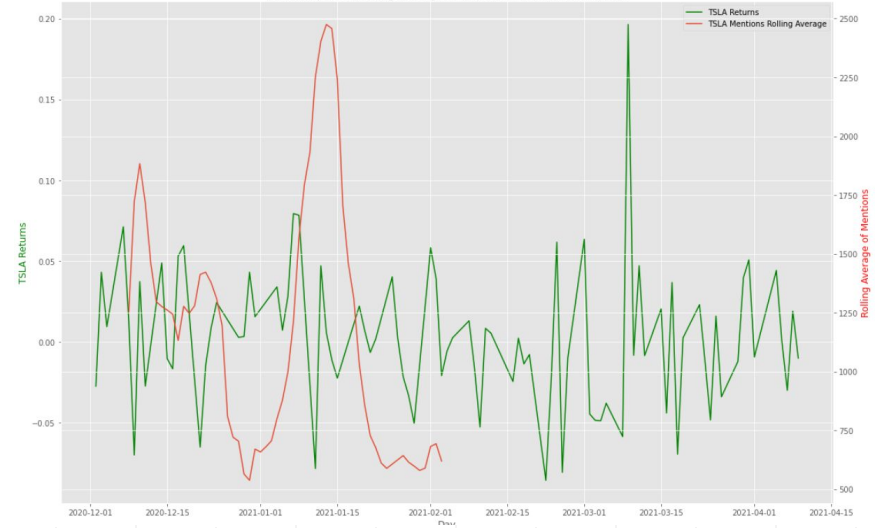


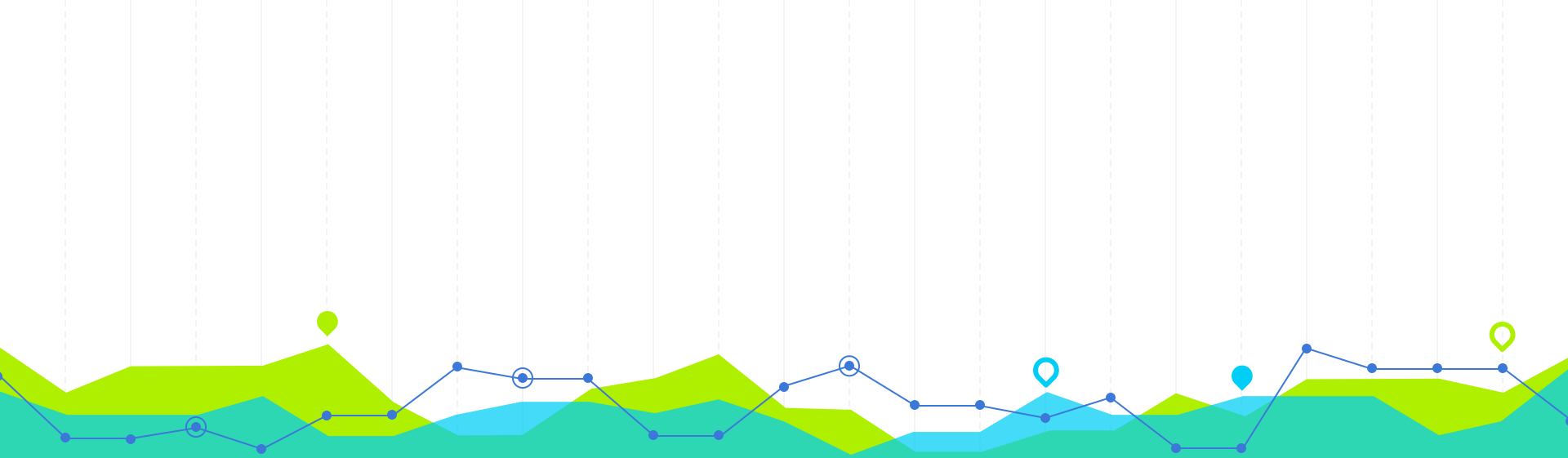
RESULTS

Closing and Moving Average Price vs Rolling Average of Mentions based on Comments



TSLA Returns vs Rolling Average of Mentions based on Comments





Conclusion

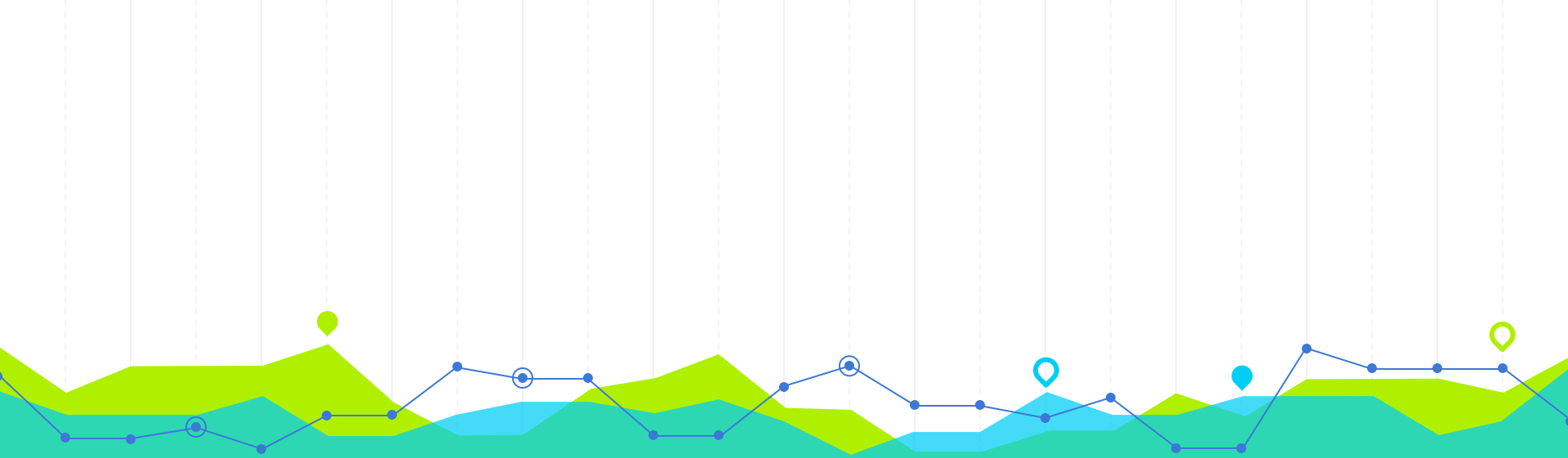
4

CONCLUSION

In conclusion, it is clear that there is a correlation between the number of mentions of a stock per day and its future value. Platforms like Reddit can let retail investors influence the market in ways never before seen.

It is valuable information about the market and investors can use this to their advantage, but it is important to be wary of the high risk involved trading against major market makers.





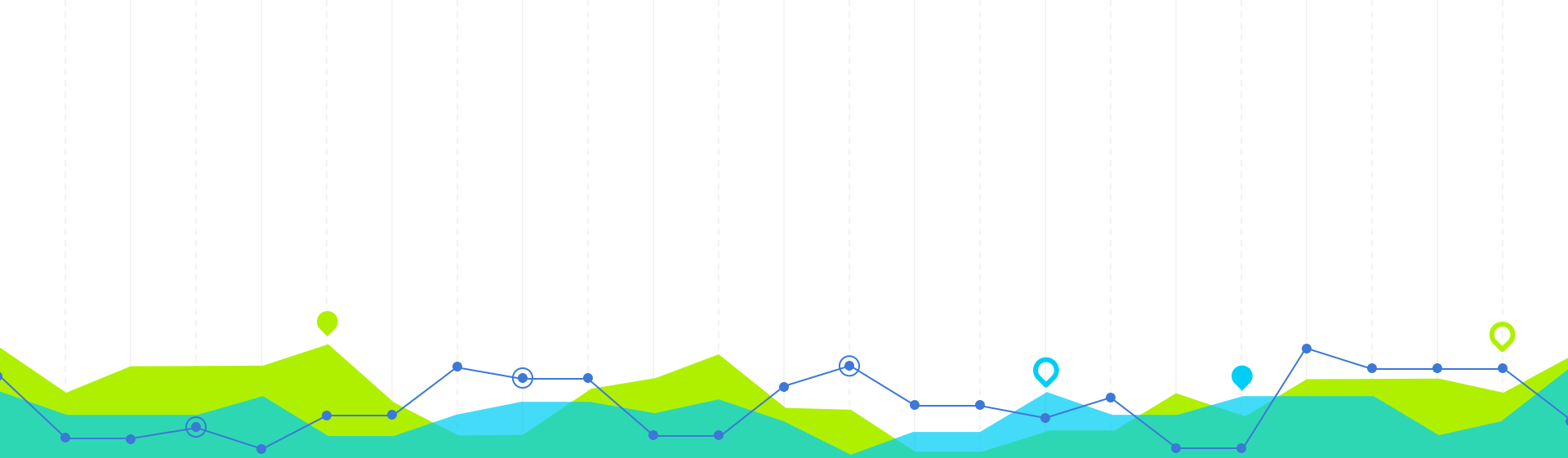
Future Work

5

FUTURE WORK

We want to introduce the Expected Reddit Enthusiasm Score which could more accurately take into account the increased discussion, sentiment, and timeliness of Reddit activity. This can allow us to more accurately determine the effectiveness of Reddit activity around a certain stock and how it relates to the change in value.





References

6



REFERENCES

Benjamin D. Horne, Sibel Adali, and Sujoy Sikdar. “Identifying the social signals that drive online discussions: A case study of Reddit communities”. May 2017. <https://arxiv.org/pdf/1705.02673.pdf>

William Sattelberg. “The Demographics Of Reddit: Who Uses The Site?”. April 6, 2021. <https://www.alphr.com/demographics-reddit>