

# Stock Movement Analysis Based on Social Media Sentiment

To build a comprehensive sentiment analysis of stock movements using the Reddit API and visualize the findings, we'll break the task into multiple steps. Here's a structured approach to achieve this:

## 1. Data Collection

I collect Reddit posts (from relevant subreddits from 'wallstreetbets') and stock price data for the same period. Sentiment scores will be calculated for the posts, and stock price changes will be analyzed in relation to Reddit discussions.

## 2. Tools and Libraries

Firstly installed necessary libraries like

1. Praw for Reddit API wrapper
2. Nltk/ VaderSentiment for Sentiment analysis
3. Pandas for Data Manipulation
4. Yfinance for fetching stock price data
5. Matplotlib/seaborn for data visualization

## 3. Data Scrapping from Reddit API

- By using Praw Library fetch posts from relevant subreddits and compute sentiment scores. Here I fetched 1000 posts and for sentiment score I Initialize Vader sentiment analyzer and also connect Reddit API
- Performed sentiment analysis on the scraped data to gauge the overall mood

Extract key features such as:

- Sentiment polarity (e.g., whether the discussions around a stock are generally positive, negative and neutral).
- Frequency of mentions for specific stocks or market trends.
- Explore potential correlations between the features (sentiment, mentions) and actual stock price movements over time.
- By using Pandas Manipulate data and convert it into csv

## 4. Stock Price Data Collection

Use 'yfinance' to get the stock price data corresponding to the same period. This will fetch stock price data for a particular ticker (e.g. Microsoft) for the time range defined in start\_date and end\_date.

## 5. Data Preprocessing for Visualization

- Next, we need to preprocess the data to aggregate sentiment scores and stock price changes by date.
- Aggregate the Reddit sentiment scores by datetime
- calculate daily stock price change in percentage
- Merge Reddit sentiment with stock price data on datetime using inner join

## 6. Visualization

now create visualizations for:

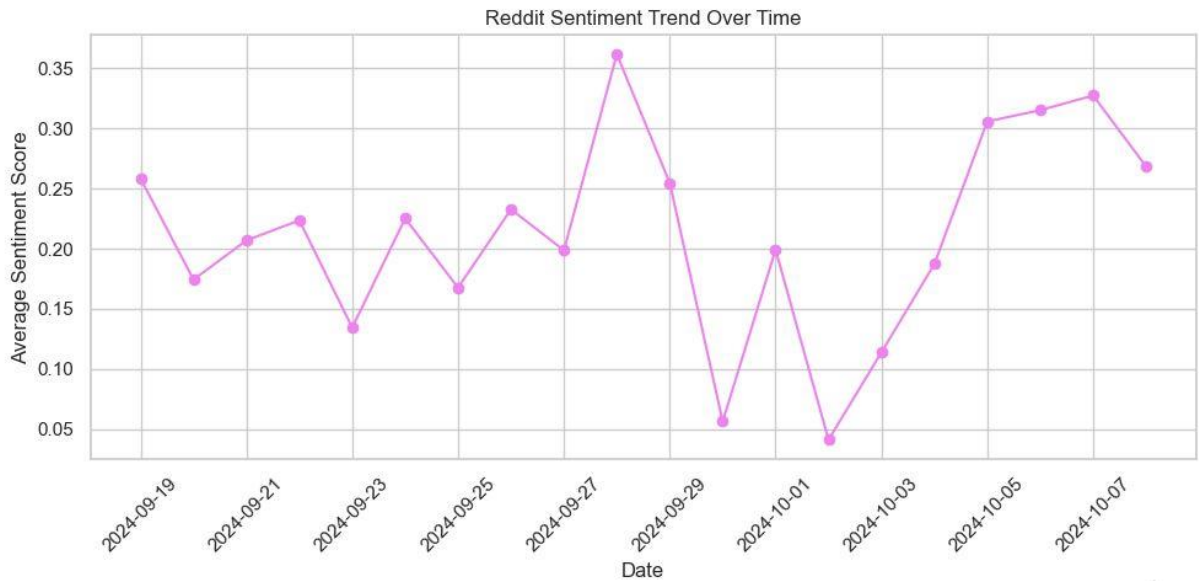
- Trends in Reddit sentiment over time.
- Correlations between Reddit sentiment and stock price movement

## 7. Report and Insights

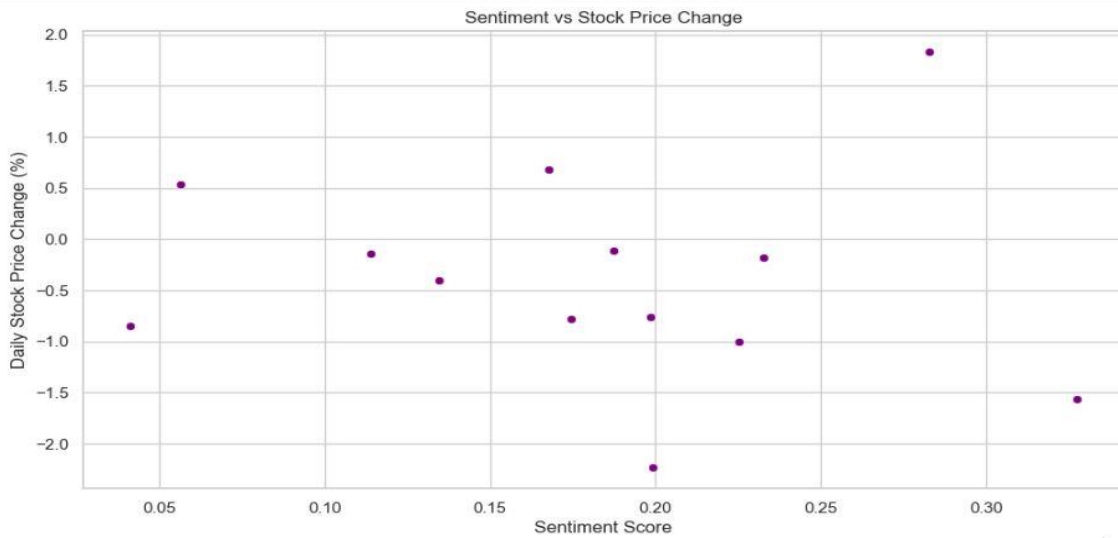
Based on the visualizations and the data, provide insights such as:

- **Sentiment Trends:**
  - If sentiment was increasing and the stock price was rising, it could indicate that positive Reddit discussions influenced the price upward.
  - Conversely, if negative sentiment was rising and the stock price was falling, it might suggest that negative discussions contributed to the decline.
- **Correlation Analysis:**
  - A strong positive correlation between sentiment and stock price movements may indicate that Reddit sentiment has a tangible impact on stock behavior.
- **Buy/Sell Signal:**
  - Buy Signal: If sentiment is positive and rising, and stock price movements show upward trends after sentiment spikes, it could indicate a potential buy opportunity.
  - Sell Signal: If sentiment becomes overly positive and starts to drop (or if negative sentiment spikes), and the stock price starts to decline, it might be a sell signal.

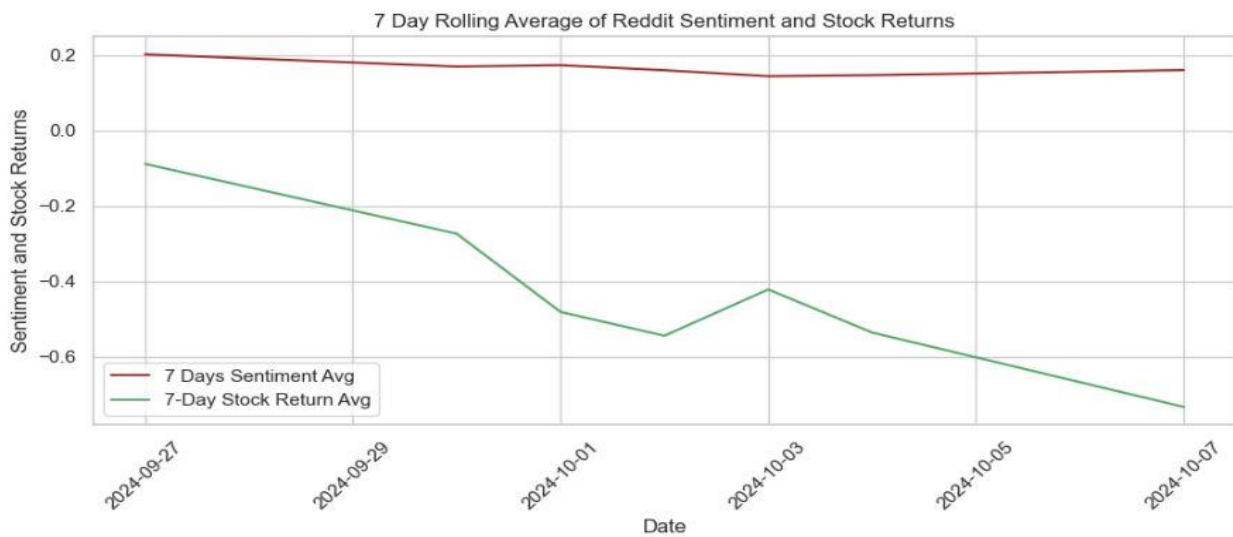
## 8. Output



- A line graph showing sentiment scores (positive, negative) from 19<sup>th</sup> Sept to 7<sup>th</sup> Oct.
- The x-axis will represent the date, and the y-axis will represent the avg sentiment score.
- Positive sentiment values indicate optimism, while negative values indicate pessimism.



- A scatter plot showing the relationship between sentiment and stock price change.
- Each point on the plot represents a single day's sentiment score and stock price return.
- Positive sentiment scores and price movements will cluster in certain regions of the plot, providing insight into correlation patterns.



- A line graph showing two rolling averages:
  - One for sentiment (in red).
  - One for stock returns (in green).
- The 7-day average smooths out the fluctuations and reveals broader trends in sentiment and price changes.
- Peaks in sentiment may correlate with rises in stock returns, and vice versa.

## 9. Actional Insight



## Conclusion

With the scatter plot and correlation coefficient in hand, you can provide insights like:

- "For MSFT, negative Reddit sentiment has a **moderate negative correlation** with stock price movements, suggesting that negative posts on Reddit may have an influence on price decreases."
- "Based on the past data, when sentiment was consistently negative, the stock price tended to fall, making it a **'Sell signal'**"