

# Meme stonks, Apes, Tendies, to the moon....!

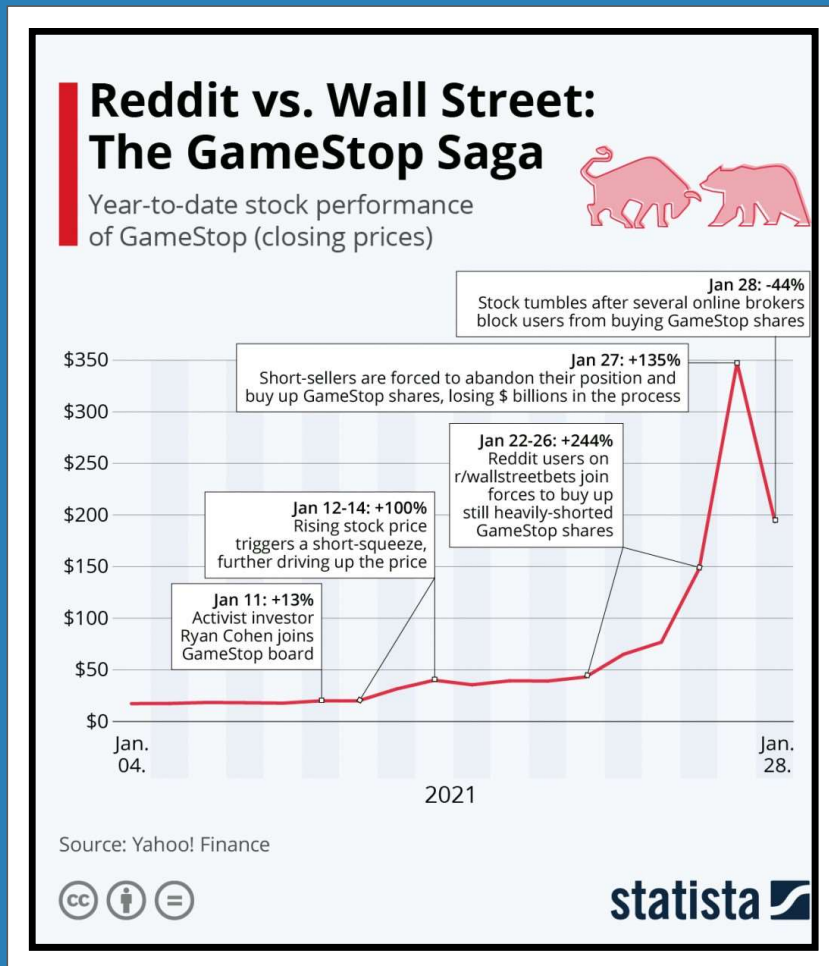
---

Using Text Analytics to Identify  
Reddit's Influence on Wallstreet  
Investments

Sowmiya Muruganandam

# Background Information

- In the past few months, the stock valuation of GameStop (GME) has fluctuated significantly. This unexpected phenomenon was presumably spurred by the massive purchasing power and hype generated from subscribers to a subreddit community called "WallStreetBets".
- Similar concepts of buying into hype has happened before:
  - Optimism in Certain Markets (i.e. Housing) generated through lax policies and opportune rates.
  - Cryptocurrencies (Bitcoin (BTC) and Dogecoin (DOGE)) generated through continued discussion in news and other mediums.
- However, the massive collection action of bulk investment into one stock such as what's seen with GME appears to be unprecedented.



# Business Question

“Is there a correlation between the amount of online dialogue around a particular stock and the stock's price?”

- Does social media (in particular, Reddit) play an impact on a stock's price?
- If there is a correlation between online communication and stock prices, does sentiment of a given stock also have an impact on the stock's price?





# Agenda

- Methodology
- Data Identification
- Topic Modeling Output vs Stock Price
- Stock Mentions vs Price Fluctuations
- Project Metrics

# Methodology

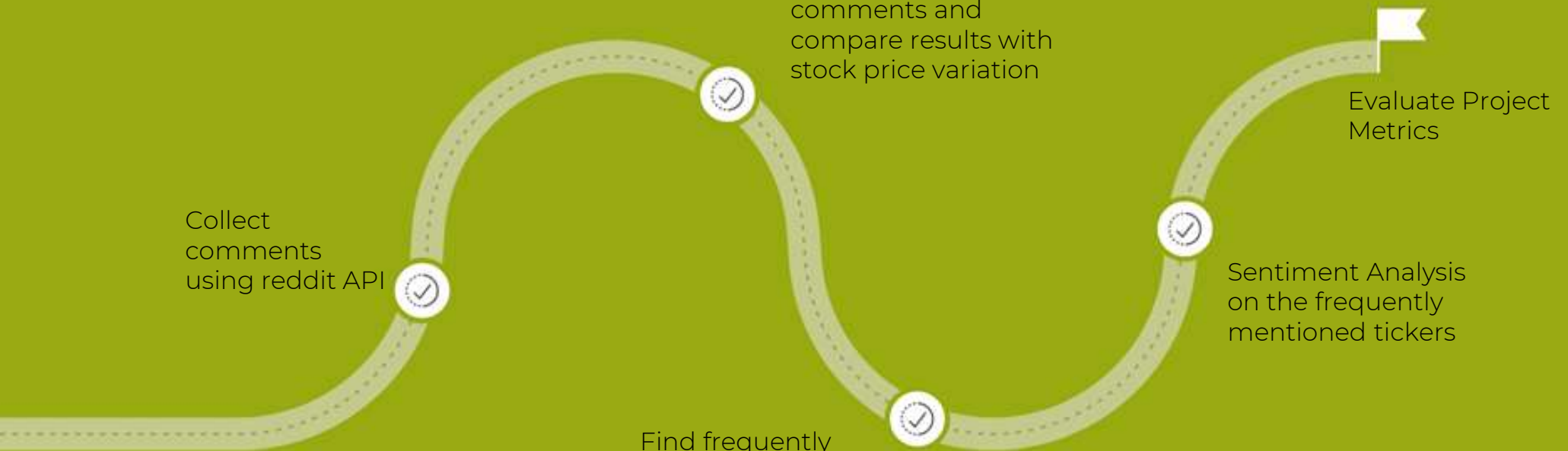
Collect  
comments  
using reddit API

Perform topic  
Modelling on user  
comments and  
compare results with  
stock price variation

Find frequently  
mentioned tickers and  
check for correlation  
with the stock price

Sentiment Analysis  
on the frequently  
mentioned tickers

Evaluate Project  
Metrics





# Data Identification

## Reddit Comments

- Subreddits used:
  - wallstreetbets
  - wallstreetbetsnew
  - GME
  - AMC
  - AMC Stock
- No. of Comments : 450+K
- Text Blob output:
  - No of Words: 7555797
  - No. of sentences: 725392

## Stock Information

- Stock Tickers
  - Companies traded in NASDAQ, NYSE
  - Tickers collected : 7272
- Stock Price List
  - List of stock prices collected from finance.yahoo.co
  - No. of stocks used : 10



# Topic Modeling

## Purpose in the project

- To discover the mentions of stocks to identify the meme stock in trend
- To discover topics to understand the trend for buy/hold activities among the group
- To discover correlation of stock movement and most discussed topics

## Process



# Topics Extracted by month vs Stock Price of GME

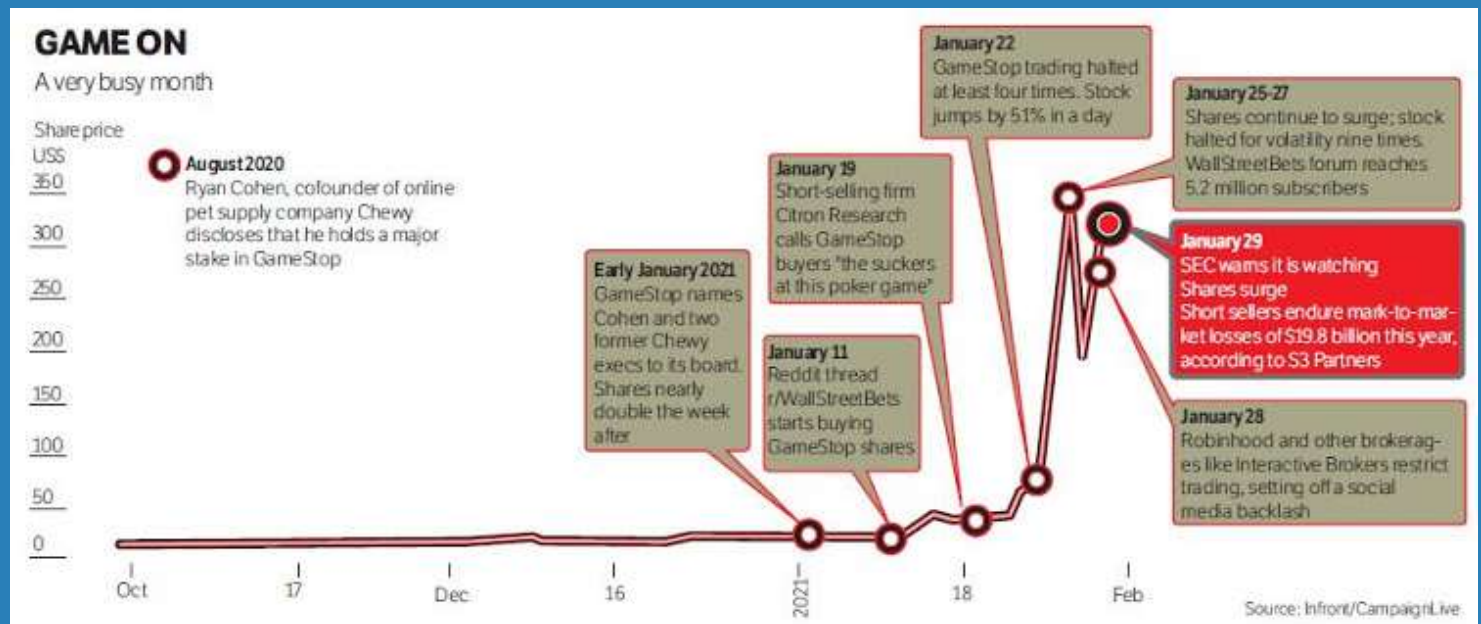
Jan	Feb	Mar
love	gme	gme
gme	amc	your
im	apes	people
time	diamond	dont
hedge	hands	just
buying	stock	big
stock	hold	money
just	like	like
shares	moon	time





# Topics Extracted (Jan 21 – Feb 7)

just
moon
loss
stop
yes
dont
selling
hold
sell
gme



# Stock Mentions on Reddit vs Stock Price

## Our Process

- 1 Identify frequent tickers
- 2 List top 10 tickers from comments
- 3 Check for variation in the stock price

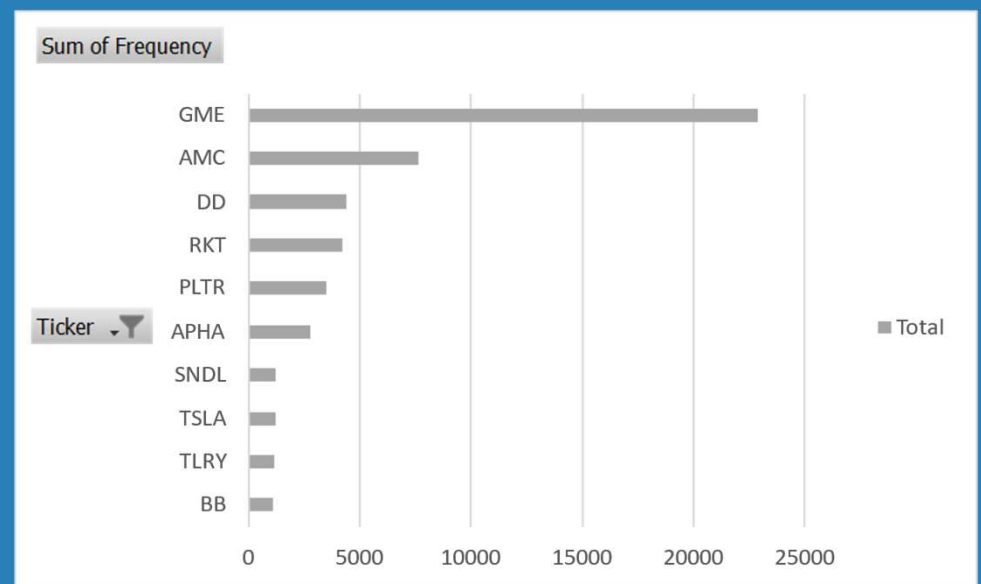
Using the tickers from a list of 7k companies, the words in the reddit data were analysed and the most frequent tickers were identified

Selected top ten tickers out of 21, for further analysis.

Checked the price variation of the top ten stocks with the yahoo stock data for the months Jan-April 2021

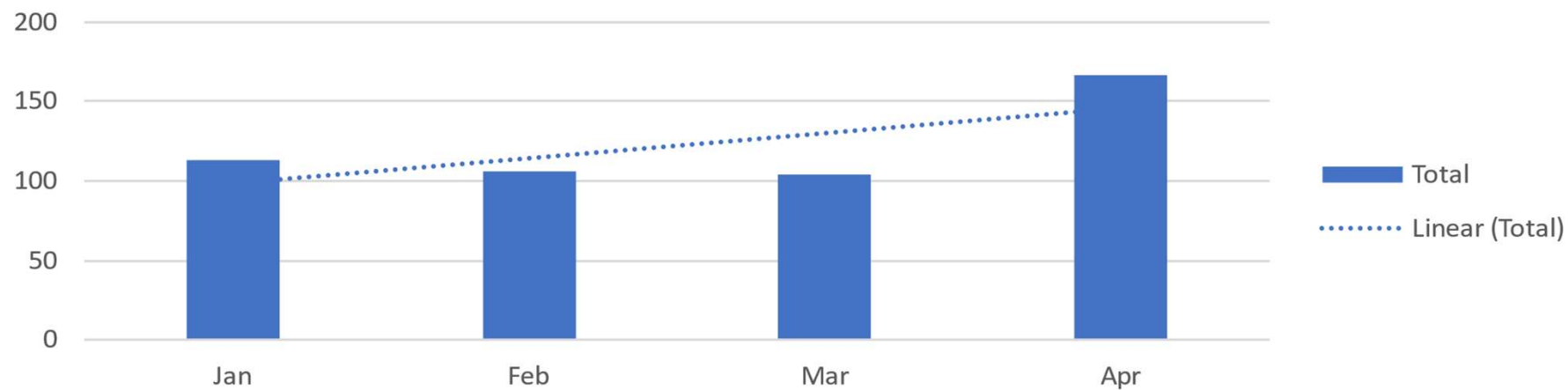
# List of Top 10 Stock Tickers identified from Reddit data

Stock Ticker	Company	Frequency
GME	GameStop Corporation Common Stock	22890
AMC	AMC Entertainment Holdings Inc. Class A Common Stock	7649
DD	DuPont de Nemours Inc. Common Stock	4408
RKT	Rocket Companies Inc. Class A Common Stock	4202
PLTR	Palantir Technologies Inc. Class A Common Stock	3494
APHA	Aphria Inc. Common Shares	2757
SNDL	Sundial Growers Inc. Common Shares	1208
TSLA	Tesla Inc. Common Stock	1205
TLRY	Tilray Inc. Class 2 Common Stock	1158
BB	BlackBerry Limited Common Stock	1105



Average of Adj Close

## Top 10 Stocks and their monthly average closing price (Jan to April 2021)

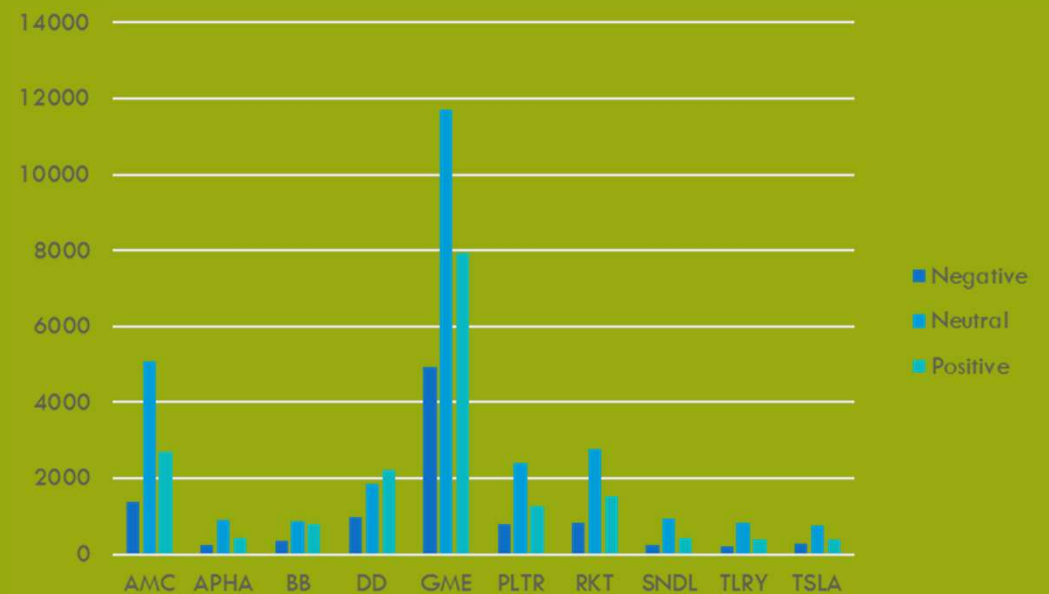


Months ▼ Ticker ▼ Date ▼

+ -

# Sentiment Analysis for the Top 10 Stocks

Stock Ticker	Positive	Negative	Neutral
AMC	1381	5064	2713
APHA	248	887	438
BB	357	870	800
DD	963	1854	2228
GME	4935	11698	7918
PLTR	810	2415	1265
RKT	819	2752	1511
SNDL	251	954	438
TLRY	215	815	388
TSLA	264	750	402



# Metrics to Determine Project Objective

## Project Objective: Influence of Reddit on Stock Market

- Topic modeling – The results of the analysis were in correlation with the stock price movement
- Frequency of mentions of stocks – We were able to deduce the most mentioned stock tickers and were able to see their increase in price for the months specified
- Sentiment Analysis : Sentiment analysis did not provide any accurate correlation. Negative, neutral, or positive yielded the same result ,(i.e. stock price increase). More research needs to be done to understand this event



Thank You!