

War on Wall Street:
Info 247 Final Project Writeup

Alex Popescu and Kayvon Tadj

Project Goals

In early 2021, a large number of retail investors, connected through social media, primarily Reddit, were able to take advantage of some hedge funds and execute a short squeeze, which yielded them significant gains at the expense of those hedge funds. This marked the beginning of a new era for trading in the stock market: retail investors demonstrated the power to rival hedge funds. Over the course of one week, GameStop stock, which was the target of the short squeeze, soared from \$20 to \$480. Since the events happened so suddenly and over such a short period of time, it was incredibly difficult for anyone, let alone those new to retail investing, to keep track of it all.

The overarching goal of this project is to explain the events surrounding GameStop stock in early 2021. To do so, the infographic educates new retail investors on the basics of stock trading (how stocks are traded, where the money flows, and the institutions involved). Once the foundation has been laid, data, such as stock price and short interest, is visualized to give a detailed understanding of what exactly happened to cause the short squeeze. Correlations are presented among short interest, stock prices, and the activity surrounding GameStop on Reddit, specifically r/WallStreetBets. As the reader gains a better understanding of what exactly a short squeeze is, they are presented with the next piece of the story: halting trades. The infographic aims to show all of the connections among the relevant financial institutions (market makers, brokers, and hedge funds) to suggest a potential conflict of interest. Furthermore, this is used to show new retail investors the inherent disadvantage they take on when making trades. Finally, the infographics present a summary of the aftermath of the events.

Tasks:

1. Education on basic stock knowledge
2. Presentation of important financial data
3. Correlations between financial data and r/WallStreetBets posts
4. Suggest conflict of interest
5. Explain aftermath

After viewing the visualization, the reader should be able to:

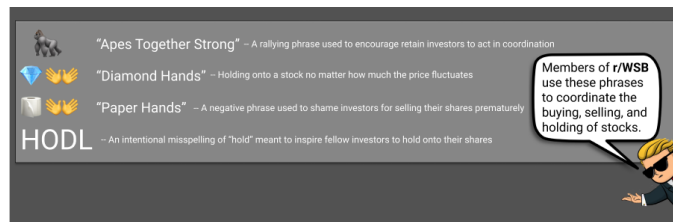
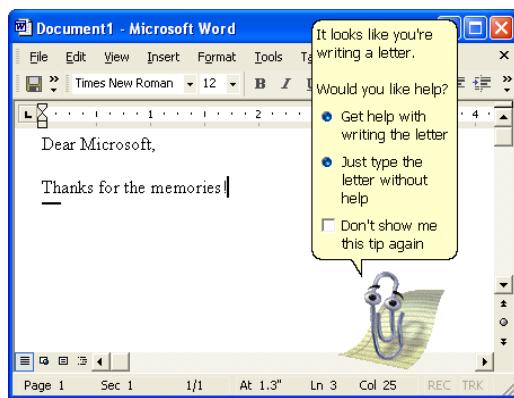
1. Distinguish between long/short trading
2. Define all the institutions involved with trading

3. Describe what happened to GameStop (short squeeze and conflict of interest)
4. Identify correlations when looking at the graphs in the visualization
5. Describe the aftermath

Discussion of Related Work

1. Microsoft Clippit (Clippy)

Microsoft Clippit was introduced to Microsoft Office products in the late 90's as an intelligent assistant for users. While not well received and eventually dropped by Microsoft, Clippy survived as a meme for many years. When we were trying to find a vehicle to introduce definitions and additional context throughout our narrative infographic, Clippy made its way to the forefronts of our minds. By merging the concept of Clippy with the mascot for r/WallStreetBets, we were able to achieve our goal without distracting from the core content, all while maintaining the theme of the infographic as a whole.



On the left, we see Microsoft Clippy offering help in writing a letter. In the same way, the r/WallStreetBets mascot appears from the border of the visualization to offer support in better understanding concepts. The use of a speech bubble was inspired by Clippy as well; the speech bubble helps humanize the assistants and, hopefully, transfer knowledge better.

2. [Yahoo Finance SlideShow/Animation](#)

A month after the dust settled, Yahoo Finance reporter Brian Cheung presented a slideshow describing the flow of money when a retail investor places an order. This described the

interactions among retail investors and financial institutions, how those institutions make profits, and how the structure was a contributing factor towards brokerages halting trades.



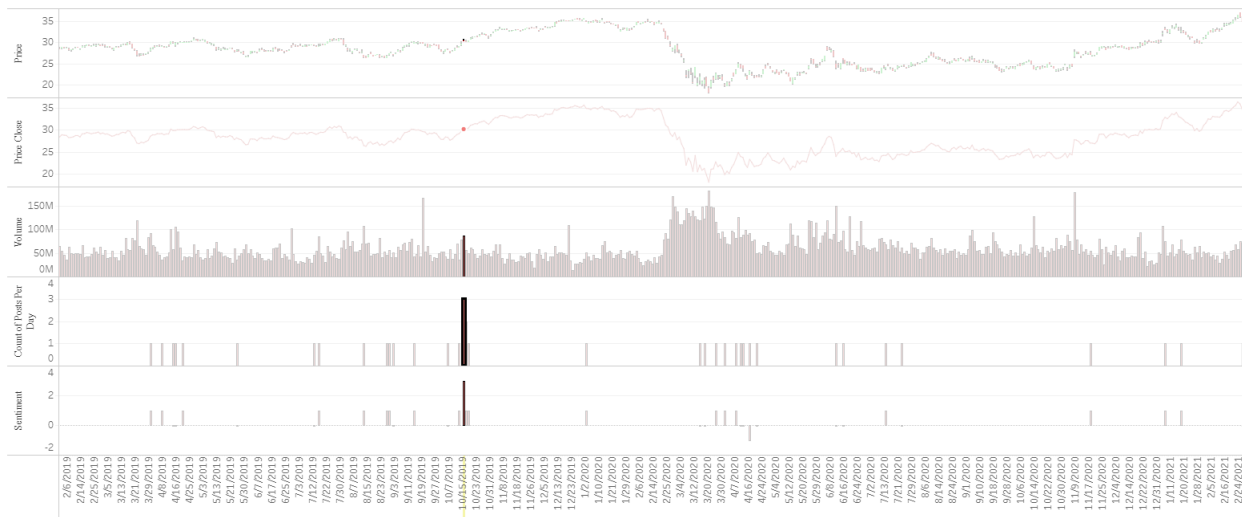
On the left, we see the Yahoo Finance slideshow explaining the relationships among brokerages, money makers, and the stock exchange. The use of logos and arrows to simplify the flow of money inspired our “Flow of Money” visualization on the right. The Yahoo Finance presentation used captions to walk the viewer through an example; however, our target audience, unlike Yahoo Finance’s, are investors with little to no background in finance. In order to simplify the process, we chose to provide definitions of “market makers” and “brokerages.”

3. [ABC News Timeline](#)

In mid-February, ABC News released an article presenting a timeline of events relating to the game stop short squeeze and all the institutions and people involved. While the article presented a lot of data surrounding hedge fund investments/losses and stock prices, it didn’t focus on these as its audience is composed of the general public. The narrative was presented through a series of tweets and quotes by research firms, hedge funds, brokerages, celebrities, and public officials. This inspired the “Aftermath” section of the infographic, which aimed to depict how the events challenged the status quo of the financial system. In the same way as the article, a series of tweets were presented with captions to contextualize them and explain their influence.

4. Related Visualization of Stock Price and r/WallStreetBets Posts

r/WallStreetBets Posts and Stock Prices



We found this visualization through a colleague that was working on a visualization project that was focused on a similar topic. The visualization shows price faceted with several other variables like close price, transaction volume, posts per day, and sentiment. This graph inspired us to make a similar faceted graph that showed how GME and AMC price fluctuated with mentions in posts on r/WallStreetBets.

5. New York Times GameStop reader's guide:

<https://www.nytimes.com/2021/02/02/business/gamestop-stock-trading.html>

This New York Times article is essentially a compilation of NYT articles that are related to the GameStop short squeeze. It splits up the articles into two sections: one that explains the background and one that gives the reader advice on investing in GameStop. We liked the comprehensiveness of this list, but felt that it is not organized enough for the reader to understand the situation without doing a lot of work. It is similar to our project since it discusses the context, r/WallStreetBets, Robinhood, hedge funds, lawmakers, and their involvement in the situation. However, we feel like ours improves on this significantly by presenting the information to the reader in a narrative sequence, accompanied by images that explain key concepts and the interconnectedness of all the parties involved.

6. WallStreetBets Due Diligence Post:

https://www.reddit.com/r/wallstreetbets/comments/laikd0/i_feel_like_clarification_is_needed_about_today/

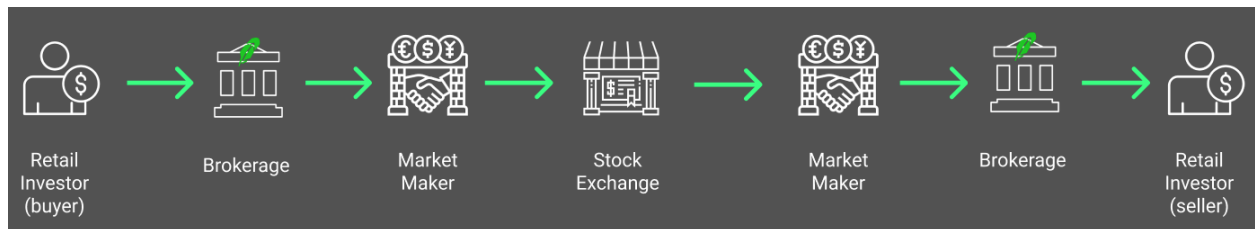
This post on r/WallStreetBets was made to calm down the community in the middle of this event, attempting to explain why the price was not increasing. They explained to the community some of the more complicated dynamics of the stock market, describing hedge fund tactics and the role the media was playing during this time. It provides links to other posts that explain concepts like short ladder attacks and show examples of media manipulation. Although this is a much more specific explanation of the GameStop short squeeze than our project, this post uses similar methods to explain complicated concepts to less knowledgeable investors. It segments the situation into different sections, each of which explains a different angle. This is similar to our approach, where we explain different perspectives and concepts.

Visualization Description

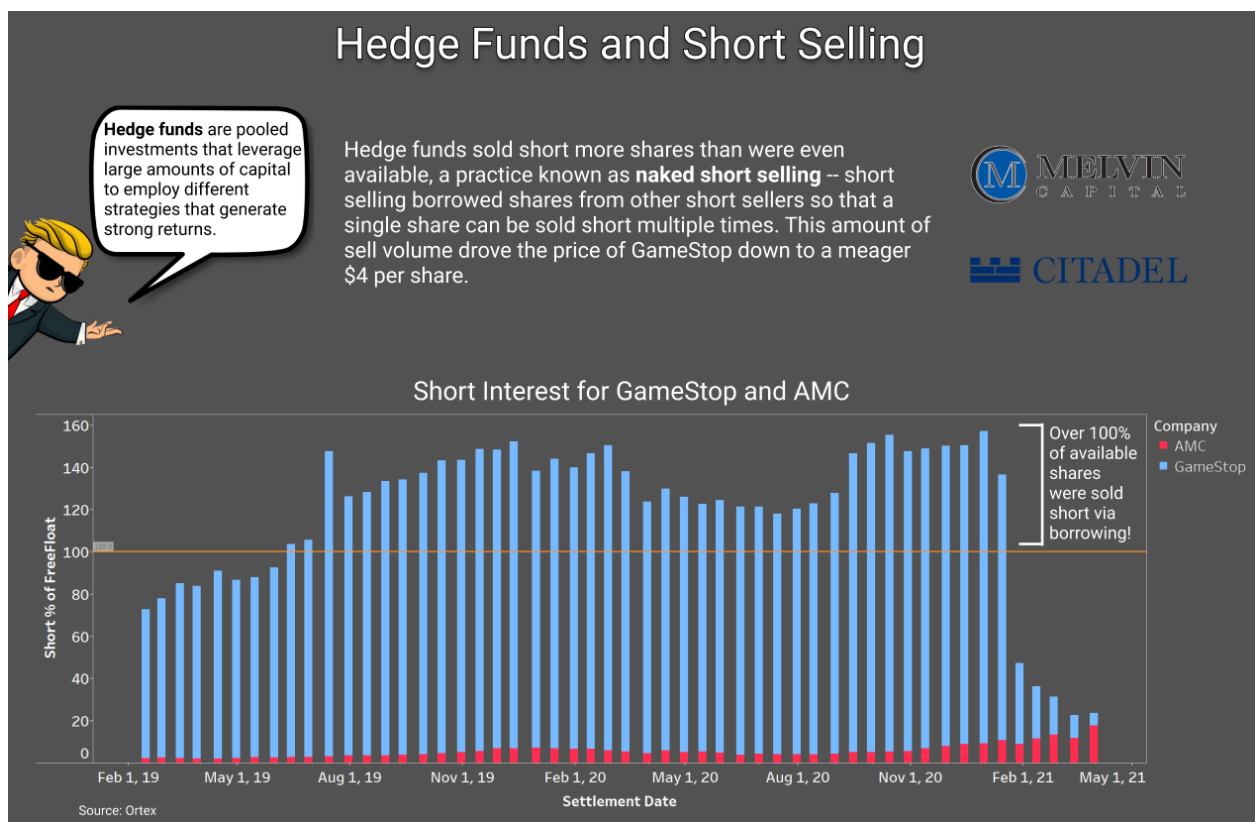


Here the basics of stock trading are introduced to the audience. Green and red are used to indicate increases and decreases in price, respectively. Long trading is shown to be the process of buying at a cheap price and selling at a higher price. While many people may not be familiar with the term, or its context in stock trading, a vast majority of people are familiar with the concept of “buy low, sell high,” regardless of the good being traded for profit. Trading long is primarily introduced here to contrast trading short, which simply flips the order: “sell high, buy low.” Short trading is a concept that many beginners have trouble wrapping their heads around; the most common question is “how can I sell something that I don’t own.” To answer this question, the r/WallStreetBets mascot has been used to describe the borrowing process implicit in

short trading. Short trading is explained here because it is an integral part of a “short squeeze,” which is the central cause of the GameStop debacle.

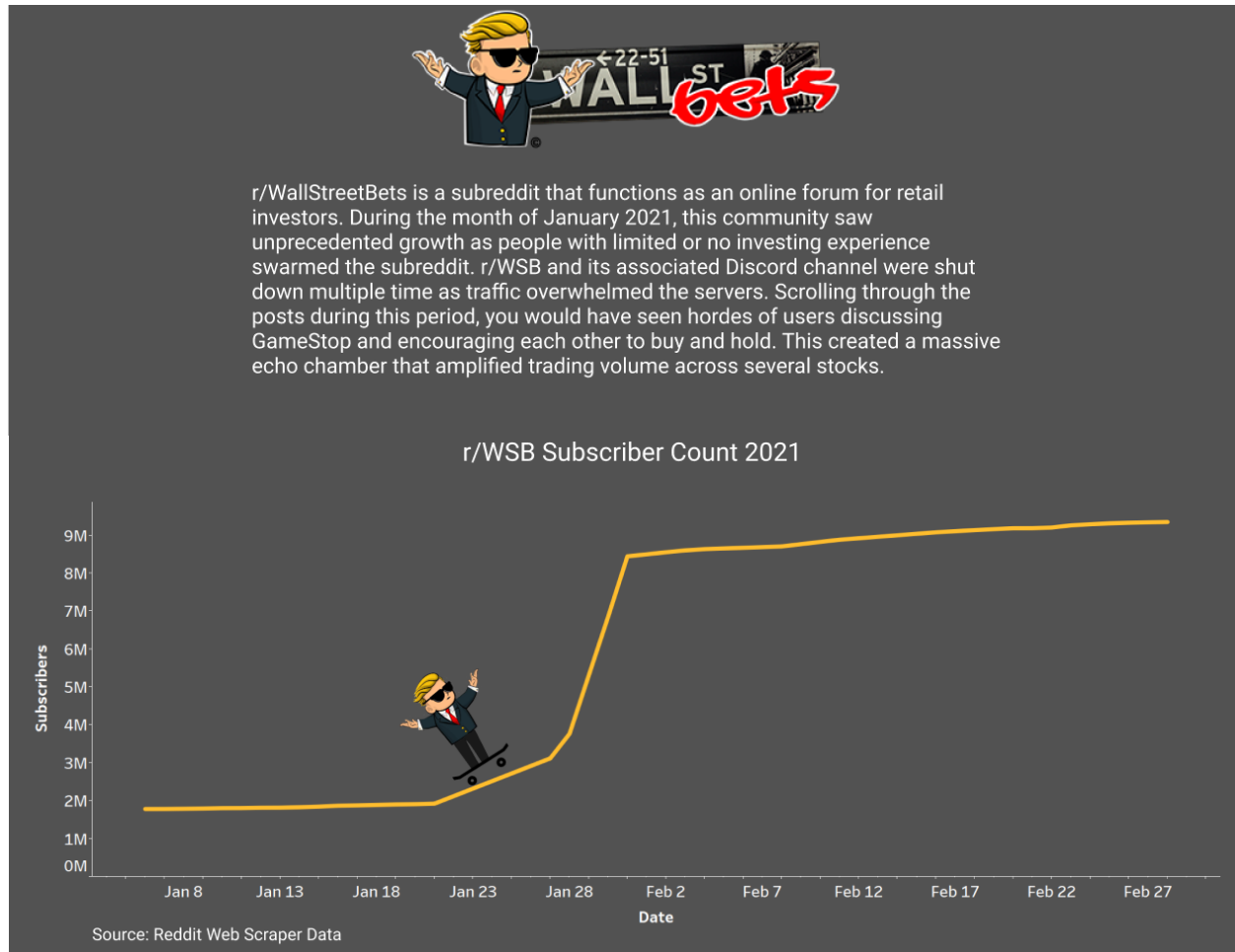


The goal of this visualization is to further build the foundation of stock trading. In this case, the entire financial system from the retail investor to the stock market is depicted. Financial institutions such as brokerages and market makers are included as they all play roles in the short squeeze. Originally one retail investor, broker, and money maker were included, but the count was doubled to clarify that each transaction requires 2 retail investors. The r/WallStreetBets logo is used here to define brokerage and money makers.

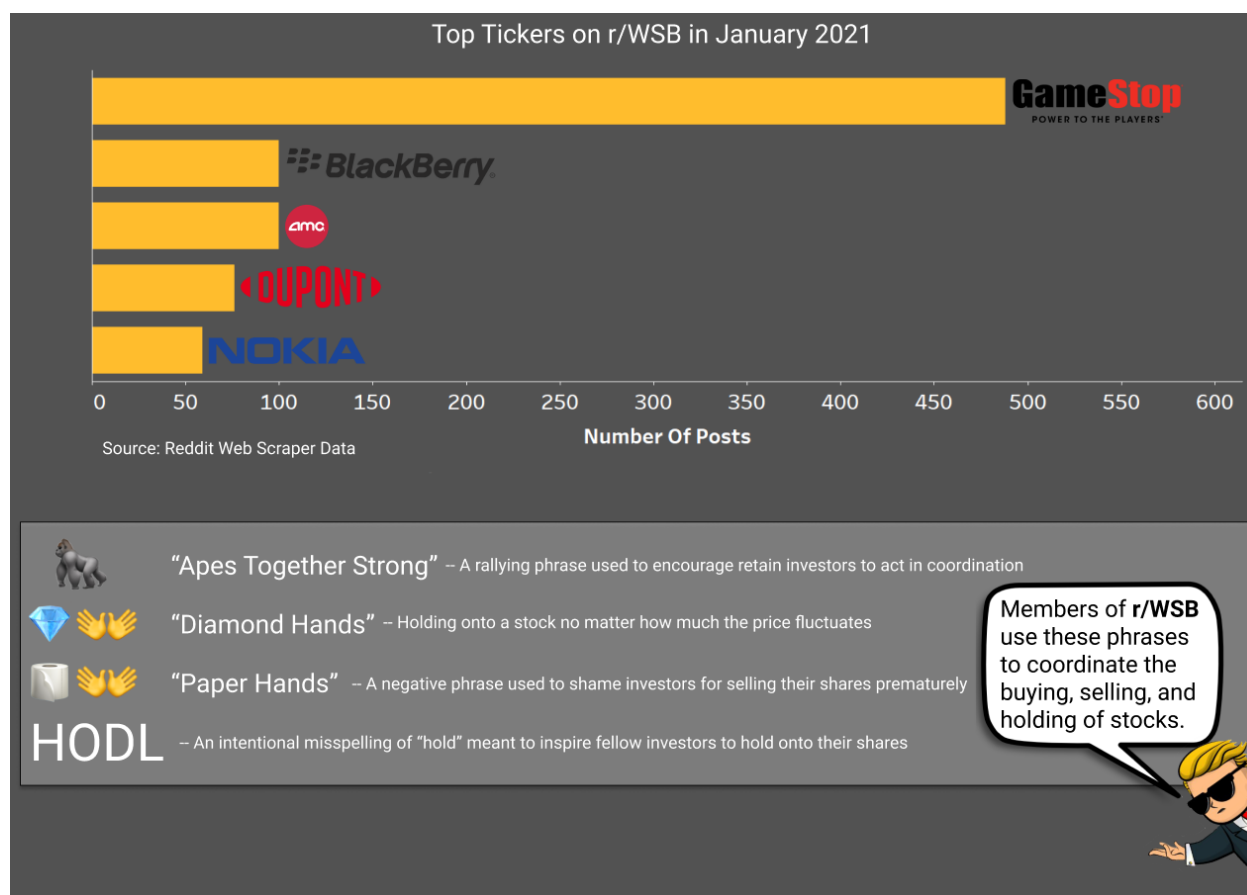


This section introduces the topic of hedge funds and explains the concept of short selling. This is meant to build on the previous section’s explanation of trading short and the stock market money

flow by explaining the practice of naked short selling and how hedge funds used it to drive the price down. The graph is meant to show the short interest of GameStop and AMC, both relative to each other and the total amount of available shares. This graph also illustrates how short interest crashed following the spike in price, which enforces the concept of hedge fund manipulation.



The WallStreetBets section introduces the story of the retail investor, contrasting with the previous section and setting up the central conflict of the narrative. The screenshot above shows the short introduction that gives context and the subscriber count graph, which is supposed to show a slightly comedic representation of the community's rapid increase in popularity.



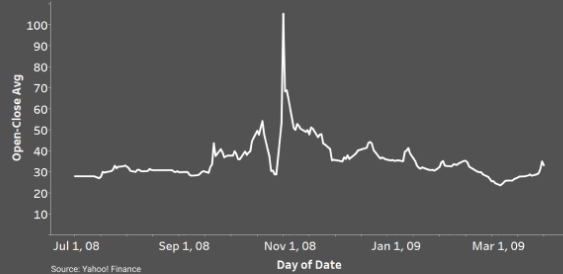
This part demonstrates how actively the community was talking about GameStop with respect to other companies. The viewer can clearly see that the number of GameStop posts dwarfs the number of posts for the next most popular company by 5 times. These two graphs in conjunction really help paint the picture for how popular both the community and GameStop got in January. The section describing the WallStreetBets vernacular complement the previous graphs by giving the viewer insight into the culture of the community and how they used these phrases to praise those who buy GameStop and shame those who sell.

What Happened



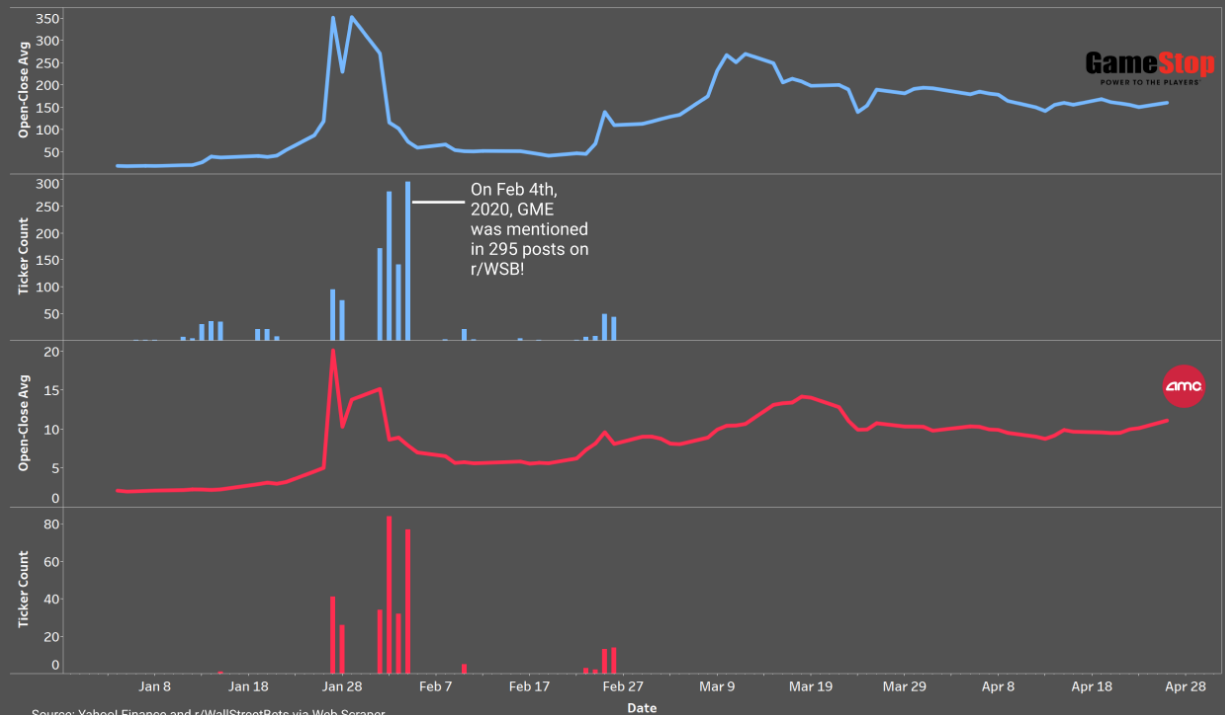
This market situation was perfect for something called a *short squeeze* – a phenomenon that occurs when a heavily shorted stock experiences heavy buy volume and light sell volume, causing an increase in price as short sellers are forced to pay back their borrowed shares in an attempt to stop losses. This was seen before in 2008 with Volkswagen and retail investors noticed this. Suddenly, the internet was buzzing about GameStop as social media sites such as Reddit and Twitter amplified the hype. This caused a legion of retail investors to feverishly buy stocks, skyrocketing the value of GME stock.

Volkswagen Stock Price



Source: Yahoo! Finance

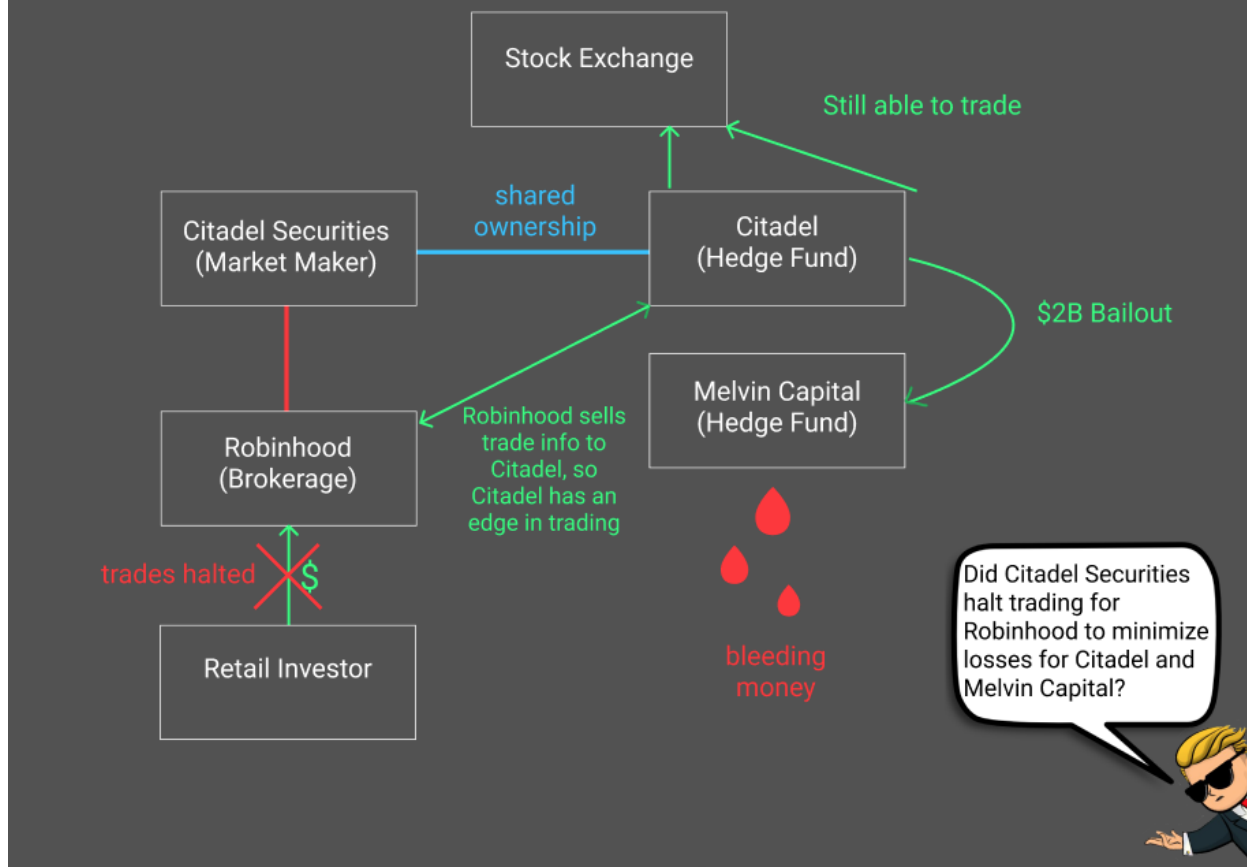
GME and AMC Stock Price and Ticker Mentions



This section is the meat of the narrative, describing the fundamental concept of a short squeeze and showing the relation between stock price and WallStreetBets post activity. It ties together all the concepts mentioned in the previous sections and shows how the involvement from hedge funds and retail investors affected the stock market. It also shows the viewer that this situation is not unique by showing a graph of Volkswagen stock price during its short squeeze.

Conflict of Interest

When GameStop price rose more than 20x in a few days, Melvin Capital lost billions of dollars. Citadel, another hedge fund, pumped \$2B into Melvin Capital to keep it afloat as it was running out of liquidity. Soon after, Robinhood began halting the purchases of GameStop stock. This led to a massive panic that not only killed, but also reversed, the momentum of the short squeeze. As the price of Gamestop fell, Melvin Capital and Citadel, which were still able to trade and cover shorts at lower prices, narrowly avoided bankruptcy. Below illustrates the relationships among the brokerages, market makers, and hedge funds. Current, an SEC investigation is underway to determine if Citadel Securities and Robinhood had conflicts of interest in halting trades, which significantly benefited their partering hedge funds.

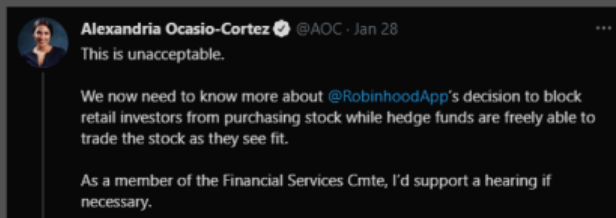


The previous visualizations were concerned with quantitative data underlying the short squeeze. While these help the reader understand what happened financially, there is more at play. This visualization adds to the narrative by looking past the quantitative data and introducing the halted trades. The relationships among all the financial institutions are visualized to drive home just how deep a network a retail investor must face. The goal of this visualization is to describe the conflict of interest that may have directly led to halting of trades and visually present the inherent disadvantage that retail investors face. We see hedge funds bleeding money and being bailed out by others, co-ownerships between hedge funds and market makers, and financial relationships between brokerages and hedge funds. None of this drew much attention before 2021, so light is

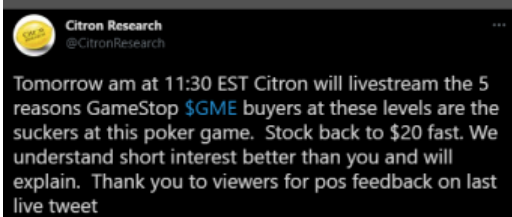
being shed on this financial web for the first time.

Aftermath

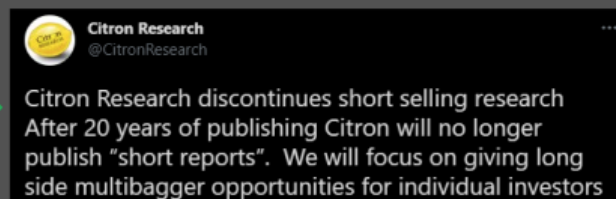
Congresswoman AOC and Senator Cruz condemn the system that favors financial institutions and support investigations. A rare bipartisan agreement.



Citron Research, a cornerstone of stock research for years, talks down to retail investors. 10 days later, they move away from supporting overshooting and aim to help retail investors with long-term investments.



10 days
later



The movement inspired people to donate games/consoles to children's hospitals and adopt gorillas (among many other animals).



The previous visualizations lead directly into the first pair of tweets presented in the

“Aftermath.” After seeing the conflict of interest, the reader is presented with lawmakers expressing bipartisan interest in investigating the financial institutions responsible for the potentially corrupt practices. The next pair of tweets reflects the feelings of many investors, and members of the financial community, that feel the predatory practice of short squeezing is outdated. Over the course of a month, Citron Research, a well-respected institution for stock research, decided to move away from pushing short trades to curb the number of companies being shorted into bankruptcy. Instead, they are beginning the transition towards helping retail investors make safe, long-term investments. After presenting the audience with a series of dark remarks regarding the stock market, it’s nice to end on a happy note. The final two tweets show how many retail investors have used their profits to donate their communities; while the cult-like atmosphere of r/WallStreetBets isn’t necessarily good as it coerces uneducated investors to make risky trades, it did inspire many people to do good as well.

Data

The sources of data can be split up into two categories: public financial data and Reddit web scraper data. The public financial data was retrieved via download while the Reddit web scraper data was retrieved using an API (described in the next section).

Financial data includes stock prices, company ticker names listed on the NASDAQ and NYSE, and short interest data. Stock price data for GameStop, AMC, and Volkswagen were acquired from Yahoo! Finance (finance.yahoo.com) and includes fields such as Date, Open Price, Close Price, High Price, and Low Price. Our ticker data were acquired from Datahub (datahub.io) and was only used to compile a list of all the ticker symbols to be used to analyze the web scraper data. The short interest data was acquired from Ortex (ortex.com) using our available free page views and includes the relevant fields Settlement Date and Short % of Free Float.

The web scraper data that was acquired using the API was limited in scope to the first 100 days of 2021 in order to collect only temporally relevant data. The important fields that were collected included ID, Title, Num_Comments, Score, SelfText (the text in the body of the post), Subreddit_Subscribers, and TimeStamp. The web scraper was filtered to look at only the submissions posted to r/WallStreetBets. It was then further filtered to consider only posts that

exceeded a score of 10 (calculated on Reddit by using [number of upvotes - number of downvotes]) and had more than 5 comments. This web scraper also considered only days that had more than 100 posts to ensure that we were looking at active days on the subreddit. This data reduction also made the code run much faster and reduced runtime for transformation tasks farther down the data pipeline. After the raw data was collected, we used various Python libraries to extract the tickers mentioned in the post and then aggregated that data to see how many posts per day mentioned tickers like GME and AMC.

Tools

_____The visualizations in this project were created entirely in Tableau and Figma. Tableau provided a clean and simple way to hold all our data sources in one place while allowing us to manage color schemes that were used across the central design. The copy as image function also facilitated cohesiveness between Tableau and Figma, easily allowing us to transfer graphs and resize them to fit in Figma while preserving image quality. All visualizations that used actual data were made in Tableau, while the visualizations that were used to demonstrate a concept were created in Figma. We used Noun Project (thenounproject.com) for our iconography and the various tools in Figma to create diagrams explaining financial concepts.

All code used for this project was done using Python via Jupyter Notebook. The main libraries used were Pandas, NumPy, datetime, and re (regex) for data transformation and storage. The API used for web scraping was the Pushshift Reddit Web Scraping API (github.com/pushshift/api). This API works seamlessly in a Jupyter Notebook, using a URL to access posts on Reddit and storing the data in a JSON format, which was then converted into a Pandas DataFrame and written to a CSV.

Results

Our usability study was instrumental in providing helpful feedback in order to assess interpretability, design aesthetic, and information quality. We selected 3 participants that had a diverse range of experience with information visualization, finance, and the GameStop short

squeeze. Their demographics and summary statistics regarding their performance on tasks and their change in knowledge are shown in the table below.

User	A	B	C
Age	22	21	23
Gender	Male	Male	Male
Infoviz Experience	Fair	Excellent	Good
Financial Experience	Very Poor	Very Poor	Good
GME Short Squeeze Familiarity	Fair	Poor	Excellent
Avg Task Time	1:27	0:28	0:27
Avg Number of Errors	0.67	0.33	0
Avg Number of Questions Indicating Confusion	0.5	0.25	0
Avg Number of Hints Needed	0.25	0	0
Avg Number of Failed Tasks	0	0.33	0
Avg Likert Score of Pre-Assessment Questions	2.8	3	4.8
Avg Likert Score of Post-Assessment Questions	4	3.6	5

We saw that each participant's knowledge was augmented after interacting with our visualization, especially those with very limited experience in the subject matter. This was our goal, as our target audience was focused on inexperienced investors. The participants were able to learn about the entire narrative and understand the central concepts regarding the controversy

with Robinhood and Citadel. We learned that the narrative structure and speech bubbles were instrumental to the participants' understanding of the concepts. They especially excelled in understanding the graphs about r/WallStreetBets and how they related to the entire narrative. However, we found out that we failed to sufficiently explain concepts like short selling/short interest, and were confused about the y-axis/orange line on the short interest graph and the faceted ticker count graph. We needed to create a better visualization and text combination in order to fully explain the concept of short trading.

The first thing we changed was our explanation of long vs short trading. We changed the text in the speech bubble to emphasize that the investor borrows a share from someone, immediately sells it and then buys the share back at a lower price and returns it for a profit. Prior to this change, the text mentioned only that the investors borrows the share and then returns it back at a lower price. Although the change is subtle, we believe that this clears up most confusion without needing to take up unnecessary space to explain the concept in more depth. We made the short interest graph more clear by adding an annotation above the orange line that indicated over 100% of available GameStop shares were shorted -- complementing our description of naked short selling in the section above. We felt that this method was the most concise way to explain the y-axis and the orange line at once. In order to explain what ticker count meant on the faceted y-axis of the stock price graph, we added another annotation that pointed to a specific data point, emphasizing that February 4th, GME was mentioned in 295 posts on r/WallStreetBets. Again, we felt that this small annotation added the most information to the visualization without taking up too much space.

One of the central aspects of our usability study was to compare the effectiveness of two designs in our "Conflict of Interest" section. One design was a minimalist flowchart that explained how all the players involved were connected. The other design attempted to explain the same concept but with an iconographic design. The participants overwhelmingly preferred the simpler diagram over the more visually pleasing design. We cleaned up the flowchart and added some more text to fully explain the situation.

We also made some other minor changes to try to maximize interpretability. We changed the money flow diagram in the first section to show the flow from multiple retail investors in order to highlight that each investment involves a trade between two parties. We also cleared up the market maker's involvement by changing the text in the speech bubble to emphasize that they

assume monetary liability in any transaction among retail investors. Lastly, we switched our bar graph for top ticker mentions in January from vertical to horizontal to look more aesthetically pleasing.

Link to Demo

<https://www.figma.com/proto/fDVueyJ6cm1JNISNx7l9TX/GME-Project?node-id=1%3A2&scaling=min-zoom&page-id=0%3A1>

Partner Contribution

We communicated with each other quite thoroughly through the entire process. Every task was proof-read and discussed by both members every step of the way.

Task	Subtask	Alex	Kayvon
Project Proposal		50%	50%
Sketches		50%	50%
Affinity Designer	Layout	30%	70%
	Graph	70%	30%
Mid-Project Presentation		50%	50%
Data Sourcing	Finding Data	0%	100%
	Importing Data	100%	0%
Data Analysis	EDA	50%	50%
Data Visualizations	Short vs Long	100%	0%

	Flow of Money	0%	100%
	Short Interest	50%	50%
	r/WallStreetBets subscribers	100%	0%
	Top Ticker Mentions	100%	0%
	Volkswagen	0%	100%
	Stock Price	50%	50%
	Conflict of Interest	50%	50%
	Tweets	0%	100%
Usability Study Design	Likert Scale questions	50%	50%
	Pre/Post questions (comprehension)	50%	50%
	Post questions (design)	50%	50%
	Tasks	50%	50%
	Quantitative Data	50%	50%
	Competing Designs	50%	50%
Usability Study	Finding/Organizing Participants	50%	50%
	Conducting Interview	66.67%	33.33%
	Recording Answers/Comments	33.33%	66.67%
Usability Study Writeup	Introduction	0%	100%
	Methods	0%	100%
	Scenarios/Tasks	0%	100%

	Test Measures	0%	100%
	Results	100%	0%
	Qualitative Measures	100%	0%
	Discussion	100%	0%
	Appendix	100%	0%
Final Infographic	r/WallStreetBets mascot comments	50%	50%
	Hedge Funds/Short Selling Text Description	0%	100%
	Wall Street Bets Text Description	100%	0%
	What Happened Text Description	100%	0%
	Conflict of Interest Text Description	50%	50%
	Aftermath Descriptions	0%	100%
	Icons	50%	50%
	Color Theme	50%	50%
	Design	50%	50%
Final Paper	Project Goals	0%	100%
	Discussion of Related Work	50%	50%
	Visualization Description	50%	50%

	Data	100%	0%
	Tools	100%	0%
	Results	50%	50%
	Partner Contribution	0%	100%
Average		50%	50%