

Assignment 2, exercise 1

```
In [2]: #Importing of packages
import nltk
from nltk import word_tokenize
import re
import pandas as pd
from collections import Counter
from nltk.corpus import stopwords
import numpy as np
import math
from wordcloud import WordCloud
english_words = set(nltk.corpus.words.words())
stopwords=stopwords.words('english')
```

```
In [3]: #Importing the dataset
df = pd.read_csv('/Users/espensivertsen/Downloads/wallstreet_subs.csv')
```

Creating the tags

Creation and implementation of the 'textsubmission' column

```
In [4]: #Creation and implementation of the 'textsubmission' column
x=df['title']
y=df['selftext']
x_1=[]
y_1=[]
for i in range(len(df)):
    x_1.append(x[i])
    y_1.append(y[i])
list1=[]
for i in range(len(df)):
    list1.append(x_1[i]+' '+y_1[i])

df['textsubmission']=list1
```

Finding every word starting with '\$'

```
In [5]: b=str(list(df['textsubmission']))
reg_ex=('\$\w+')

c=re.findall(reg_ex,b)
```

Removing the '\$'

```
In [6]: c = [e[1:] for e in c]
```

Removing every element containing a number

```
In [7]: test = [x for x in c if not any(x1.isdigit() for x1 in x)]
```

Making everything uppercase

```
In [8]: main_list=[x.upper() for x in test]
```

Finding the top 15 tags

```
In [9]: word_counter = {}
        for word in main_list:
            if word in word_counter:
                word_counter[word] += 1
            else:
                word_counter[word] = 1

        popular_words = sorted(word_counter, key = word_counter.get, reverse = True)
        top_15 = popular_words[:15]
        top_15
```

```
Out[9]: ['SPY',
        'TSLA',
        'SPCE',
        'PLTR',
        'MSFT',
        'ROPE',
        'AAPL',
        'AMZN',
        'NIO',
        'ZM',
        'AMD',
        'BABA',
        'GME',
        'DIS',
        'BA']
```

Exercise 1: Tokenizing and cleaning

Making the "selftext" column a list

```
In [10]: liste_selftext=[]
        for i in range(len(df['selftext'])):
            liste_selftext.append(df['selftext'][i])
```

Removing URLs

```
In [11]: liste_selftext_1=[]
        for i in range(len(liste_selftext)):
            liste_selftext_1.append(re.sub(r'http\S+', '', str(liste_selftext[i])))
```

Removing numbers

```
In [12]: def remove(list):
        pattern = '[0-9]'
        list = [re.sub(pattern, '', i) for i in list]
        return list

        liste_selftext_2 = remove(liste_selftext_1)
```

Making everything lowercase

```
In [13]: liste_selftext_3=[x.lower() for x in liste_selftext_2]
```

Removing stop words

```
In [14]: def remove_stopwords(data):
          output_array=[]
          for sentence in data:
              temp_list=[]
              for word in sentence.split():
                  if word not in stopwords:
                      temp_list.append(word)
              output_array.append(' '.join(temp_list))
          return output_array

          liste_selftext_4=remove_stopwords(liste_selftext_3)
```

Tokenizing the list and removing punctuation

```
In [15]: tokenizer = nltk.RegexpTokenizer(r"\w+")
          liste_token=[]
          for i in range(len(df['selftext'])):
              liste_token.append(tokenizer.tokenize(str(liste_selftext_4[i])))
```

remove stopwords that were "hidden" in punctuation

```
In [16]: liste_token2=[]
          for i in range(len(liste_token)):
              test=[]
              for j in liste_token[i]:
                  if j not in stopwords:
                      test.append(j)
              liste_token2.append(test)
```

removing duplicates within the same sublist

```
In [17]: list_tokens = []
          for i in range(len(liste_token2)):
              a=[]
              for elem in liste_token2[i]:
                  if elem not in a:
                      a.append(elem)
              list_tokens.append(a)
```

Adding the tokens column to the df

```
In [18]: df['tokens']=list_tokens
```

Exercise 2: Adding the stocks to the dataframe

Making the tags lower case so that they match the tokens (which we made all lower case)

```
In [19]: top_15_lower=[x.lower() for x in top_15]
```

Creating a list of the stocks within the tokens list (when there is no tag the list is empty)

```
In [20]: list_stocks=[]
for i in range(len(list_tokens)):
    append_list=[]
    for j in range(len(top_15_lower)):
        if top_15_lower[j] in list_tokens[i]:
            append_list.append(top_15_lower[j])
    list_stocks.append(append_list)
```

replacing all the empty items with 'Other'

```
In [21]: for i in range(len(list_stocks)):
        if list_stocks[i]==[]:
            list_stocks[i]= ['Other']
```

Adding the list of stocks to the dataframe.

```
In [22]: df['stocks']=list_stocks
```

If you look at the dataframe at the end of the document you can see that there is a column for "stocks". So by using this you can find any submission discussing the stock in question. (In regards to "Find submissions discussing at least one of the top 15 stocks you identified above.")

(Example of finding every submission title for submissions that contain 'gme')

```
In [64]: submissions_gme=[]
for i in range(len(df['stocks'])):
    if 'gme' in df['stocks'][i]:
        submissions_gme.append(df['title'][i])

submissions_gme
```

```
Out[64]: ['Here are all positions commented in the last 24 hours',
          'Thank You Wallstreetbets! **Warning to Gamestop Shorts**',
          'I just bought some GME (Gamestop), and it just might be autistic enough to work',
          'Boston authorities shut down GameStop and stores close statewide',
          'GME going down',
          'GME Options to Print',
          'DG You fuckheads',
          'GME option trades leading to net stock to buy',
          'ZM and PTON too the MOOON',
          '$GME to Zero?',
          'Puts on Home Depot or Lowe's?',
          'Here are all positions posted in the last 24 hours',
          'A play on GME earnings 5/28',
          'GME to the moon',
          'GME tanking the next 3 days',
          'Gamespot shitting the bed',
          'Actually meeting your heroes: Dr. Michael Burry',
          'Will GME finally die due to the Rona?',
          'STONK MARKET != ECONOMY',
          'GME short squeeze before EOW',
          '[DD] GME will moon this holiday season due to PS5 Launch',
          'GameStop (GME) - THE BIGGEST SHORT SQUEEZE OF YOUR ENTIRE LIFE',
          'PSA: Incoming short squeeze on GME',
          'Here are all positions posted in the last 24 hours',
```

'You tards need to buy reddit awards and not GME shares',
 'GME Goes 🚀🚀🚀🚀🚀🚀🚀',
 'In honor of the previous retards request, I don't want any free money to YOLO into PLTR. Please only pay me in Reddit awards.',
 'Shorting GME on Think Or Swim',
 'Why Cohen is Vital to a \$GME Success Story',
 'EVs wil fly again',
 'Ryan Cohen Tweet Analysis \$GME',
 'If you have money for useless reddit rewards and GME shares then you have money for Gamestop Power Up Pro',
 'After 15 consecutive red days, 2 green ones',
 'After 15 consecutive green days, two red ones',
 'Should the 1B market cap rule apply to SPACs?',
 'Should the 1B market cap rule apply to SPACs?',
 'Shilling, shadow bans and censorship!',
 'GME fomo',
 'GME 🚀🚀🚀🚀🚀🚀🚀\u200d🚀🚀\u200d🚀🚀\u200d🚀🚀\u200d🚀🚀',
 'Ryan Cohen new tweet',
 'You guys need to stop wasting money on useless Reddit awards and buy GME shares',
 'The unlikely things that just might happen in 2021. GME, BABA, LMT, GOLD, GILD',
 'Have faith in RC, GME retards!',
 'Have faith in RC, GME retards!',
 'Ryan Cohen is Probably Increasing His Stake in GME',
 'A shallow dive into r/GameStop',
 '🚀GME DD PT2 🚀🚀🚀🚀 + Coordinated Charity for Autism and Kids + making it rain at GME + Coordinated Social Media',
 'GME 2021 timeline predictions, institutional bullshit, and how to spot when the short squeeze will happen.',
 'I'm so glad I met you retards',
 'I'm so thankful I've you retards',
 'Last song for you this year',
 'WSB ---2021 New Years Resolutions',
 'Ryan Cohen is About to Light This Shit Up 🚀🚀🚀',
 'Yolo my whole account into GME ATH',
 'Why PLTR and GME are going to Moon',
 'The Mother of All Short Squeezes: GME, is it a myth?',
 'ARKG and LRNZ - And The Future of Retards',
 'GME redemption story: battle against sleazy hedge funds',
 'Happy new years degenerates. I wish y'all a new year full of 10 baggers as sexy as NIO and GME 🚀🚀🚀',
 'Super Serial Reasoning Why GME And PLTR Will Moon in 2021',
 'Help I have too much money.',
 'Gamestop. com is a f*cking mess.',
 'You tards need to buy reddit awards and not GME shares',
 'GME Goes 🚀🚀🚀🚀🚀🚀🚀',
 'In honor of the previous retards request, I don't want any free money to YOLO into PLTR. Please only pay me in Reddit awards.',
 'Shorting GME on Think Or Swim',
 'Why Cohen is Vital to a \$GME Success Story',
 'EVs wil fly again',
 'Ryan Cohen Tweet Analysis \$GME',
 'If you have money for useless reddit rewards and GME shares then you have money for Gamestop Power Up Pro',
 'After 15 consecutive red days, 2 green ones',
 'After 15 consecutive green days, two red ones',
 'Help!!! How can I get intentionally assigned stock, I need the low-risk leverage.',
 'GameStop Short Squeeze [Discussion]',
 'Wall Street Week Ahead for the trading week beginning June 8th, 2020',
 'GME GANG Still Ridin - The DD',
 'Stop pretending Hedge Funds and Big Banks aren't made of retards.',
 'A wee bit DD On The new PS5 and how it affects SNE and Gamestop',

'Possible big GME moves?',
 'Who bought GME puts?',
 'Why did my options lose value?',
 'Just a retard looking for guidance',
 'Bankrupting Institutional Investors for Dummies, ft GameStop',
 'FOLLOW ME TO TENDY TOWN',
 'More GME good news: Ryan Cohen wants to turn GameStop into Amazon rival',
 'Failure to deliver threshold list',
 'Once You Pop You Know You Can't (Game)Stop!',
 '"Two financial detriments: a Fat Tail and a fat ass."- Nassim Nicholas Taleb (probably).',
 'GameStop Pump?',
 'GME - Where have all the Options Gone?',
 'Can someone explain NKLA option pricing to a retard?',
 'Get out of VALE (or don't): Clarity on the Dark Pool and Insane Volume",
 'GameStop to become dog food company DD',
 'GameStop and Dog Food DD',
 'Gamestop (GME) 9/11 Puts',
 'CORRECTION: It's not GM buying an 11% stake in NKLA, it's GME.',
 'DisputablyGreen's First Big Green LEAP",
 'Next generation consoles and what that means for Gamestop...',
 'I'm on a roll - SCS is next",
 'Cheap GME calls for a potential lottery win',
 'Where the man who YOLOed 20k onto GME calls?',
 'GME - Michael Burry Suggestion',
 'What's your worst bets?",
 'GE mooning',
 'Fuck all the GME 🌈💩 's',
 'Gamestop: What Robinhood Doesn't Tell You',
 'GameStop will not be saved by consoles',
 'Cry For Help',
 'Listen up Autists',
 'Too late to ride the GME train?',
 'CHWY GME and Ryan Cohen Hype',
 'GME is the next Hertz',
 'Investing in Gamestop short term?',
 'GME short squeeze incoming, last chance to jump ship at the dip',
 'Gamestop (GME) Short Squeeze is still alive and well',
 'Reg SHO list update',
 'GameStop (GME) - The Last Stand (Short Squeeze Earnings Play 09/09)',
 'Why Gamestop Might Squeeze But Virgin Galactic Won't",
 'I want my \$1k, /u/ankitpatel2945',
 'FOR RETAIL EYES ONLY. NO MM'S ALLOWED. DD on GME",
 'Fuck your zoom and docu gains fellow autists!! These are real gains on actual YOLOs!!!',
 'GME has now been memed and is destined to fail 💩',
 'Everyone here jerking off tsla',
 'GME calls',
 'Why GE will be the next r/wsb meme stock',
 'Will this work?',
 'GME 9/25 \$14 volume is insane',
 'GME short interest',
 'Hey Mods .. are you going to ban all these idiots pumping the GME "short squeeze" like the VALE idiots?',
 'The Rise Of GME And Why You Should Hop On',
 'Long GME (a balance sheet DD)',
 'GME Information For People interested in forcing this squeeze',
 'GME - Further down the rabbit hole',
 'aw the chewy CEO bought GME today(he already had 10%, bought more)',
 'Terminally Ill Optionable Companies',
 'Retail Weakness over Christmas',
 'I'm another Senior in HS!!! I'm totally living the dream too!",
 'WallStreetBets YOLOs of the month (top YOLOs) for the month of January 2020',

```

'WallStreetBets YOLOs of the month (top YOLOs) for the month of January 2020',
'WallStreetBets YOLOs of the month (top YOLOs) for the month of January 2020',
'WallStreetBets YOLOs of the month (top YOLOs) for the month of January 2020',
'WallStreetBets YOLOs of the month (top YOLOs) for the month of January 2020',
'WallStreetBets YOLOs of the month (top YOLOs) for the month of January 2020',
'WallStreetBets YOLOs of the month (top YOLOs) for the month of January 2020',
'WallStreetBets YOLOs of the month (top YOLOs) for the month of January 2020',
'WallStreetBets YOLOs of the month (top YOLOs) for the month of January 2020',
'WallStreetBets YOLOs of the month (top YOLOs) for the month of January 2020',
'WallStreetBets YOLOs of the month (top YOLOs) for the month of January 2020',
'Gamestop and Ford. This is the way',
'$GME - Short Bubble',
'UBER options play',
'GME about to moon 🌕',
'GME Puts?',
'I wrote some code to scrape US financial disclosures and track stock trading by politicians. Here are 5 congressmen who definitely browse this sub.',
'I've been scraping US financial disclosures to track stock trading by politicians. Here are 5 congressmen who definitely browse this sub.',
'Flairs and making money, I have a dream speech',
'GME console week',
'GameStop leads google trends in the US',
'GME Reality check part 2',
'GME - Let the games begin',
"GME short call: it won't go tits up... ?",
'The Most Popular Tickers on Reddit for Today, 2020-11-14',
'The Market Wizards: Beat the Market with These 9 Weird Tricks!',
'Calls of GME. Employees DANCING just to get more hours due to increased sales.',
'So, what's some stocks that we can turn into hype stocks in this sub?',
'Estimated ~70,000 orders per day on their E-commerce platform for 4Q20, HUGE upside potential with limited downside risk',
'GME Bull Case Mid-November 2020',
'Banbet: GME',
'Game consoles may haul in $45 billion this year, and big changes are coming',
'What if GME God u/DeepFuckingValue is actually Ryan Cohen?',
'GME Technical Analysis',
'Dr. Michael Burry is gonna iLUMNate your portfolio with tendies.',
'Let Dr. Michael Burry illUMNate your tendies.',
'GME gang is back. TO THE MOON 🚀🚀🚀',
'GAMESTOP WONTSTOP POPPING',
'GME - Share Lending FAQ',
'Buy GME Monday or after earnings?',
'Gme options for a beginner',
'Phase all the money from PLTR into GME',
'GME Flood',
'GME short squeeze trigger strategy',
'WSB CANCEL CULTURE',
'Gme',
'GME SHORT SQUEEZE INBOUND. 103% Short Positions. More shares shorter than there are actual shares. Prepare your anus',
'Big retard seeks assistance to make moon gains or eventually sleep under a bridge',
'Retarded screeching GME',
'A sHoRt sQuEeZe has been on its way for months',

```


'When PLTR recovers, I am hopping on the GME train',
 'Best way to donate to charity: shares not cash!',
 'GME get in while you can 🚀🚀🚀',
 'Should we add stop losses to our position? Protect our tendies (PLTR, GME, T
 SLA)',
 'GME Earnings Specifics?',
 'GME Stock/Options',
 'In for 230k of GME shares',
 'WSB Autist Game Plan Week of November 30 (All-in-One Thread)',
 'Someone should calculate collective GME ownership of WSB',
 'Game Plan Week of Nov 30 (All in One Thread)',
 'Case Against GME Short Squeeze',
 'The Phases of WSB Meme Stocks',
 'GME has several times more revenue than NIO. If GME traded at the same rati
 o, shares would cost 3.2k',
 'WSB Calls the Top',
 'Put in a premarket order for GME tomorrow or wait a few hours to see what ha
 ppens?',
 'Is it possible to sell calls during a short squeeze?',
 'People pumping GME: Melvin Capital Management hates you a lot',
 'GME VS PLTR',
 'Your GME options?',
 'Limit Order for GME',
 'First time buyer',
 'Welcome to Short Squeeze Monday',
 'Is GME a meme or are you guys for real?',
 'Gme short squeeze??',
 'Is it one minute before midnight?',
 'SNE and GME console release pump',
 'This company is the #2 trending Google Search in the US (1m+ searches)',
 '[CRSR] Corsair \$40c May 21, 2021 Exp',
 'Short Interest % Volume for GME?',
 'GME Options',
 'If you liked Burry on GME, you're gonna love him on DISCA",
 'When will PSTH print? Got my three day notice and need cash asap.',
 'GameStop Announces Voluntary Early Redemption of Senior Notes.',
 "GME short squeeze won't happen anytime soon and it probably won't be as big
 as you think, stop buying weeklies and get shares instead",
 'Nothing Can Stop What Is Coming - GME',
 'GME Squeeze Finally??',
 'GME?',
 'GME Gang dead?',
 'Breaking Robinhood News and DD Loss Porn',
 'Breaking Robinhood News and DD Loss Porn',
 'FLY ON THE GME ROCKET blow up hedge funds Let Cohen rain his tendies down on
 us all',
 'Meme! Stocks! Update!',
 'What's with all the hype around GME?',
 '\$RIDE',
 'GME - WSJ Exclusive Article about Cohen near 10% stake just dropped',
 'Obscure stocks to make tendies on',
 'PSA: Pulling Scrabble Tiles from your Asses isn't DD',
 'GME',
 "I hope GameStop gets what's coming to them.",
 'GME - Warning to all Shorts and Gay Bears, you will not be warned again.',
 'Is GME gonna moon or what? 🚀🚀🚀',
 'Reminiscences of a Stock Operator: The Evolution of the Average WSB Autist',
 'The Most Popular Tickers on Reddit for Today, 2020-11-20.',
 "GME DD Because there wasn't enough of it already",
 'Our time has come, GME shall rise again.',
 'New here',
 'How do short squeezes work?',
 'Why GME will survive and prosper',
 'I just got my paycheck - what should I yolo it into?',

"If you keep talking about it, it won't come true",
 'WSB is in desperate need of a paper trading competition',
 'GME- what I think will happen next week',
 'There are literal market manipulation attempts in this sub right now and mods are doing nothing about it.',
 'Former Chewy CEO tries to push GameStop to become the Amazon of the video-game industry',
 'Retard vs Autist',
 'Will GME squeeze be called off?',
 'Yall sleeping on Tim Apple - the next Tesla',
 'Yall sleeping on Tim Apple - the next Tesla',
 'LAST CALL for GME',
 'Don't be dumb, buy GME',
 'Special thanks (GME)',
 'My autism',
 'PETITION FOR U/DEEPFUCKINGVALUE TO DO A GME PORTFOLIO UPDATE AT MARKET OPEN',
 'GME...',
 'WSB mods gay keep taking down GME posts',
 "This is your last chance (probably), don't miss the rocket 🚀",
 'Thoughts on FLUX',
 'Anyone got some decent ideas for calls I can make on Monday with about 2k?',
 'Last day for GME tendies?',
 'Should I should I sell my 50k portfolio and buy GME?',
 'REPOST: PAY ATTENTION. It isnt a meme, its a solid play. Gamestop is a once in a lifetime >3000% return opportunity. DD READ THIS RETARDS. SHORT SQUEEZE MEANS MORE RETURNS THAN PLTR, NIO, XPEV, EVERY MEME STOCK COMBINED. GME 🚀🚀🚀🚀🚀🚀🚀🚀🚀',
 'GME FULL STORY',
 'GME support last console cycle was around \$17.74. Now premarket is at \$17.90. You can still hop in on this rocket',
 'GME Up Premarket Reaching \$20',
 'Why's everyone creaming over GME?',
 "GME From South Africa Let's go BALLS DEEP GME WOOO",
 'GME time',
 'GME YOLO - Roth IRA',
 'Rockets and shit',
 'What the fuck is going on with GME?',
 'Turn \$500 into \$1000 on a weekly. No meme stocks. Whatcha got?',
 'It isnt a meme, its a solid play. Gamestop is a once in a lifetime >3000% return opportunity.',
 'Would the retard who was formally known as k2ebgo pm me?',
 'Much respect',
 'I guess GME short squeeze is over??',
 'I guess the GME short squeeze is over??',
 'Papa Cohen about to send GME to the moon 🌕',
 'Time to milk GME shorts - A.k.a short squeeze',
 'Yet Another GME DD - Just look at this graph -',
 'The real reason PLTR is down.',
 'GME Shorts?',
 'GME IV and Vega',
 'Please help GME get to the 🌕',
 'BAN BET: GME will not hit the high of around \$19.40 again this week. If it does, ban me instantly.',
 'Letter to retards below',
 'Peoples opinions on GME',
 'Holiday spreads for the black friday and christmas craze.',
 'Ok boys GME is rocket ship and it is nearly fueled enough (What is a short squeeze?)',
 '🚀🚀🚀GME TO PLUTO 🚀🚀🚀',
 'BLNK takeover rumors',
 'We are the captains! Be a part of Short Squeeze Monday',
 'Let's Squeeze GME',
 'Is the WallStreetBets sub-reddit fools gold or are we a diamond in the rough'

h? I think the latter',
 'TLS the the moooooon!',
 'The FED will pay us for the GME short squeeze',
 'The only bear case for GameStop is COVID-19. The vaccine is a dick at bear s',
 'What reason could someone have to purchase 1 million shares in GME in a single open market order? (1.5% of the company)',
 'Gamestop WSB Ultimate Owners list',
 'GME>PLTR*',
 'Where the fuck is /u/DeepFuckingValue?',
 'GME target price?',
 'Should I buy GME?',
 'GME 🚀',
 'What are your stock markets plays and predictions for the upcoming January 5th Senate runoff elections in Georgia?',
 "BUY CNK! You haven't missed the boat. (Squeeze inbound!)",
 'Do YOU want a GME short squeeze price target? Then this is for you.',
 'Why are people so interested in GME?',
 'What time should I buy shares tomorrow morning?',
 'GME ShortSqueeze is not happening, but GME is unvalued.',
 'GME.',
 "For those who aren't already all in on GME, here's why you should start thinking about picking up a few shares",
 'GME bagholders, I just found a treasure trove of GME porn',
 'Ryan Cohen before he sold Chewy for \$3 Billion',
 'Ryan Cohen roasts GME's board. Holding GME is getting riskier by the day',
 'Reminder !! Market only open 3.5 hours today, double ur positiond to make sure u print as much !!',
 'Missed GME? Don't Sleep on bby',
 'WSB GME ROTATION',
 'GME earnings on 12/8',
 "GME 15 EOD or I'll eat one of my socks and film it",
 'BREAKING NEWS: GME is voted next meme stock!!',
 'Should we buy GME now or wait till the market opens Monday?? Going in 1K 🚀🚀🚀🚀',
 'Should we buy GME now or waiting till the market opens Monday?? Going in 1K 🚀🚀🚀🚀',
 'GME and citron',
 'I'm convinced by you guys with buying GME but worried about one thing...',
 'I will be retired by 2025',
 'What about CRSR?',
 'How to ride momentum/meme stocks like PLTR/NIO/CRSR (and soon GME) without getting burned.',
 'Michael Burry of the Big Short owns GME (4.3%) and so should you',
 'Quick, mods and GME short holders are sleeping',
 'Thoughts on any I should add?',
 'GME gang',
 'Make bigger 🚀🚀🚀🚀',
 'Why you believe GME is bound to have a squeeze',
 'Atleryx (AYX) this weird name company is going to the 🤔🤔',
 'Question about apex tech and GME.',
 'Yolo everything in GME?',
 "It's all in, balls deep time. What is the choice: PLTR or GME?",
 'My first month of trading',
 'GME short squeeze question',
 'Will GME Take The Title of Mobile Gaming King? DD Inside',
 'How does premarket work',
 'SO WE ARE ALL THINKING ABOUT BUYING GME',
 'PLTR, MRNA & GME 🚀🚀🚀🚀',
 'How to play the GME squeeze',
 'Holy Shit (GME Shitpost)',
 'The GME Short Squeeze and You; an autists guide to infinite tendies',
 'Get it in GMHI/Luminar NOW!!',
 'New Spin on an Old Game: Load up on ELY (Callaway) for Top Golf Merger',

'Case and Point - Mods banned my last post calling into question new Policy C
hanges',
 'FOMO purchases today',
 "GME Short Squeeze LET'S GO 🚀🚀🚀",
 'bet on 12/4 GME calls like a retard',
 'Already Tendie Makers -- Let's take a break from the bleeding and talk WSB",
 'GME Short % of Float',
 'Question about otm options',
 'fucking hold gme',
 'GME Short Squeeze',
 "Cirton is a garbage company, a \$50 website blog and he has been wrong 100% o
f the time. Andrew Left couldn't predict his piss going in the toilet bowl.",
 'Short squeeze explained for room-temperature IQ',
 'GME train - All abooooaaard',
 'You had one job',
 'Take your winnings and switch tables - NIO, BLNK, LI, XPEV, etc to BABA, GM
E, FSLY, OSTK',
 'The right way to play the Cannabis hype!',
 'GME TO THE MOON???',
 'WE NEED TO RALLY TOGETHER PUMP GME ONE LAST SQUEEZE',
 'What are we doing with GME?',
 'What positions are you all taking for this hopeful GME squeeze?',
 'If PLTR drops to \$20 by tomorrow close I'll eat my shit',
 'Can someone explain GME to me?',
 'Request to pump different meme stocks',
 'What's the lowest PLTR could dip to?',
 'I turn 40 in January, where do you want to drop 40k on a stock YOLO buy unti
l at least January?',
 'We printing or we printing?',
 'Get this guys',
 'If you've ever made some variation of the comment "I TOLD you all to sell. Y
ou should have listened." Then you should probably just go ahead and kill your
self.',
 'Play for tomorrow',
 'GME Calls - what to put?',
 'GME short squeeze setup',
 'First Post',
 'Full Yolo',
 'How long approx to turn say £1k into £50k ; Shares only?',
 'First time buying stocks',
 '3K --> 2.5M This Yolo Is Free',
 "Fuck y'all and also I love you (first post)",
 'GME gang question',
 'I have a vision of this crazy new Flair for WSB',
 'Y'all need to chill out with GME and APHA",
 "Don't confuse today's selloff with your long term goal.",
 'Why GME will keep going up',
 '🚩GME DEC 8TH Q3 EARNINGS🚩',
 'Hello brethren',
 'GME yolo at the top',
 "What's the fucking deal with GME???",
 'GME',
 'Keep The Loss Porn Coming',
 'Everyone stop freaking out FFS',
 'Bought 500 shares of GME at \$19. Fuck me! Ha!',
 'The Lil Yachty Prophecy has once again come to fruition',
 'PLTR, GME mooooooning',
 'GME TO FUCKING PLUTO 🚀🚀🚀',
 'Want free money? Buy the dip in CRSR.',
 'For those new to GME gang - this is a war, not a cake walk',
 'GME shorted 144,87%, how?',
 'GME shorted 144,87%, is it time to short my life and go all in?',
 'Can someone educate this noob retard why many stock prices dropped today at c
ca 16:44 CET (10 :44 in freedom time)',

```

'Where is the guy that said he would eat his own sock?',
'Next meme GME',
'GME',
'GME and PLTR Chasing',
'Why are we going in on PLTR and not GME',
'DD for retards: buying GME at $15 is like buying TUP at $8',
'I called PLTR, here are my next picks',
'GME shorters is in trouble.',
'Remember to sell every now and then',
'FLOAT and GME after market close.',
'GME reminder to put limit sell orders way out of the money',
'🚀🚀',
'Welp I'm bag holding GME',
'the PLTR rocket is having some problems',
'You're all pussies',
'GME will import my filipina girlfriend',
'ATTENTION PLTR RETARDS',
'Where the F is GME gang at??',
'Alright Boys time for some DD on MOMO',
'GME is ripping to the 🌪️!',
'GME or PLTR',
'Don't Miss Out .... Agree or Disagree?",
'Game Stop is going to fucking 0',
"Can't wait to see GME YOLO guy's account after Chewy founder buys 10% stake",
'GME short squeeze',
'GameStop +31%, Buy Puts',
"Gamestop's short float over 100%",
'I USED TO BE A GAY BEAR',
'ZOOM suspended because investors were confusing it with ZM',
'All the DD',
'Corporate Bond Defaults are coming',
'Corporate Bond Defaults are coming',
"Corporate Bond defaults are coming. Here's how to short them.",
'I just realized a hilarious blunder I made while stoned yesterday',
'GME is about to get hit hard',
'$SPY Fibonacci Retracements and Future Outlook',
'Biggest UVA-C radiation lamps and devices manufacturer?',
'Gamestop GME earnings beat expectations - 10% up after hours',
'GameStop to close 300 stores permanently.',
'Wall Street Week Ahead for the trading week beginning March 23rd, 2020',
'Essential Retail and Petitions',
"Night shift? Where you at? I'll show you mine if you show me yours.",
"Today's Stock Market News [Tuesday, March 24th, 2020]",
'GME - GameStop is a dying retailer in a dying market (Kung-Flu)',
"iT's PRicEd In",
'We are doomed',
"Today's Stock Market News [Monday, March 23rd, 2020]",
'$GME stands for everything I hate',
"Today's Pre-Market Movers & News [Thursday, March 26th, 2020]",
'Speculation',
"Today's Stock Market News [Wednesday, March 25th, 2020]",
"Today's Pre-Market News [Friday, March 20th, 2020]",
'Coronavirus, Becky, Reggie, and the return of Gamestop',
'$SPY Technical Analysis -Bearish $215p 4/17',
'[DD] $GME the Blockbuster of video games (and murderous douchebags too)',
"We are so early, we haven't hit the real issues the economy will face...",
'People talking about the shutdown',
"GME Update : No I'm not Burry, and so can you",
'GME Long Thesis - Short Squeeze',
'Gamestop calls',
'GameStop is totally screwed',
"Battlefield triage is not supposed to fix the patient, it's simply to make them stable enough to move.",

```

'Just a matter of time until GameStop shifts the bed, right?',
 '\$F',
 'Pupil of the WSB',
 'From \$AAPL to \$AAWW',
 'GME to the moon!',
 'This is a test. Long \$GME',
 'Why not YOLO stocks from real companies with like profits maybe?',
 'GME - Puts Jan 8 vs Jan 15',
 'Listen up retards',
 'Beat the Shorters? Puts on GME',
 'GME 12/1 - 12/18 forecast',
 'PROOF /all gang has completely infiltrated WSB',
 'I Have A Dream',
 '\$GME is a SCREAMING buy. But AFTER earnings....',
 'GME Limit Price?',
 'Stop slacking',
 'Special Education for the War',
 'DD for the \$GME Tribe (you \$PLTR weirdos are welcome to listen in too)',
 '"diamond hands" and "paper hands // GME',
 'Have these 4 stocks to buy YOLOing/leveraging 5000\$ - CRM; BABA; GME; TDOC;
 QRVO',
 'Made a Robinhood for my autistic 13 year old brother',
 'The thing most people miss with GME',
 'GME Next Steps',
 'Wall Street Week Ahead for the trading week beginning December 7th, 2020',
 'PSA to anyone playing GME: Buy shares!',
 'GME Gang checking in',
 'Something I don't understand about the potential GME squeeze',
 'Please explain to this gullible retard why I should sell my PLTR shares and
 buy GME this week.',
 'Best option plays for week of 12/7?',
 'Need advice for GME earnings play for next week.',
 'GME Elliott wave 3 Short Squeeze is coming!!!',
 'Publicly discussing and indicating willingness to participate in GME short s
 queeze - what are the effects?',
 'u/r8cobra checking in with my 43 shares of GME',
 'PSA: GME short interest and share price',
 'GME squeeze happening the day of Q4 Earnings',
 'The real WSB GME Survey - Limit Sell 420.69',
 '\$GME',
 'GME TO THE MOON',
 'Do you want to bet against reddit? I have the answers',
 'GME 🚀 Short squeeze explanation from my last post',
 'GME 🚀🚀🚀 what's our play with earnings tomorrow??? How do we get this squeeze
 started?!?!',
 'Could a genius explain to me how borrow fees work',
 'The Big Boys',
 '#WeWantCohen GME 2021',
 'Short Interest Update on GME for 11/30/20 with commentary of the latest conf
 erence call',
 'Short Interest Update and some light DD',
 'THE GME SHELF OFFERING IS NOTHING. GME IS STILL GOING TO THE MOON 🚀🚀🚀🚀',
 '12/11 SPY and TSLA POSSIBILITIES or STOP LOSING YOU'RE MONEY or I'M BACK FRI
 ENDS and I MISSED YOU!',
 'GME: Bull case vs bear case, from both perspective',
 'What would it look like if Ryan Cohen tried to take over GME (GME porn with
 STORY inside)',
 'Find Other Market Opportunities - SIMMPs Unite!',
 'GME Earnings Thread',
 'Banbet part II',
 'Banbet part II',
 'GME - a bit for everyone',
 'Ford',
 'GME bull thesis',

'I AM HEREBY DEMANDING JIM CRAMER TO DO AN AMA ON WSB',
 'Will GME continue to fall or rise at open tomorrow?',
 '[DD] GME shareholder DD and support thread',
 'GME - Short Squeeze Still On 🚀',
 'It should be illegal for a company to be both authorized to buy back shares and have an ATM shelf offering outstanding (GME)',
 'What to do?',
 'Airbnb & DoorDash IPOs | CHEWY & GameStop EARNINGS | APPLE news | STOCK MARKET NEWS [12-09]',
 'TSLA to the moon! Part 1 (the non DD)',
 'GME: personal experience today',
 'post your GME loss porn pussies',
 'GME Bull run thesis explained and why it should interest you. rokit rokit mu n',
 'GME SHIT SHOW, BETTER GET COHEN NOW !',
 'RALLY THE TENDIES BOYS (GME)',
 'VLDR; (Velodyne), our shiny chance to visit Pluto 🚀🚀🚀',
 'Thanks guys 😊',
 'GME DD crossing the Shermancon edition.',
 'Listen my GME brethren: they can keep the price down all they want, but we win in the end.',
 'GME gang is in a battle with the shorts. Here's how the shorts are doing',
 'GME Interesting Squeeze Facts & Market Cap & Future Valuations',
 'Ryan Cohen is coming to tendy town with an early christmas present 🚀🚀🚀',
 'Big Rock Partners Acquisition will give you big cock tendies!!! BRPAW DD inside 🚀🚀🚀',
 'A Christmas Carol for my fellow bulls: WSB Wonderland',
 'A Christmas Carol for my fellow bulls: WSB Wonderland',
 'GME 4Q20 Financial Model 🚀🚀🚀',
 'GME Today',
 'GME Longterm. Is Digital Distribution of Video Games Bound to Take Over?',
 'GME breaking the \$20 wall premarket',
 '\$GME - Regulation SHO Threshold List - shorts fked?',
 'Snow incoming!',
 'GME: Melvin Capital executed 48,000 July puts early? \$25 million realized losses?',
 'Where is my GME short squeeze?',
 'Little gains...',
 'Wut stock',
 'Why GME is going to the Moon (Compressed information + Links)',
 'GME Infinity squeeze definitive proof.',
 'Grandpa is a boomer Autist',
 'GME Donation Drive Interest Check',
 'Options question',
 'GameStop GME Background DD and Price Target 69.69 no squeeze and at least 420.69 on a squeeze.',
 'When to buy GME shares?',
 'YOLO on GME 🚀🚀🚀 - will decide if I get married or if I die broke and alone',
 '4th trade: GME bag holding',
 'Ryan Cohen fucks a stripper (gone sexual) (🚀🚀🚀🚀)',
 'Just Turned 18 and can do options!!!',
 'Listen up frog brains. GME 🚀',
 'Path to 7-Figures, The Unsexy Way',
 'GME is done rising no more FOMO, but you can still ride it back down for massive \$\$\$\$',
 'Retard Mode (activated) - what now?',
 'Why doesn't Donald Trump pump a stock and make billions and laugh?',
 'Flywheel of Basic Bitches DD on \$STIC aka BarkBox, my position is \$370k of shares',
 'GME ICR Conference - Words Matter',
 'RC bought more GME 🚀🚀🚀',
 'Stocks whose TA suggests a breakout: BBY, BYND, NVDA',
 'Gamestop GME documentary says vaccine will be delivered through stores, bull


```

ish if true',
'Any advice on finding DD sources to get me out of my shitty boomer stocks?',
'My Story, and a Christmas Praryer; Please feel free to share your story / pr
ayers!',
'gamestop.com website access denied',
'GME squeeze and iBorrowDesk',
'Too late to buy gme?',
'Now that 2020s end is upon us, how well did you do in the market fellow auti
sts?',
'Testing r/wallstreetbets vs r/stocks',
'Thank you Wallstreetbets. Merry Christmas and god bless you all.',
'Merry Christmas WSB - A quick look at this year's top 50 most mentioned tick
ers!",
'Thoughts on GME's situation & a question for GME Elders & Autists",
'🚀Christmas Miracle / GamerGate 2020: GME shorts negotiating with PE/IBs to
unwind their shorts 🚀',
'Let's get serious about GME',
'Stock prices not changing after refresh',
'Wash Sale 30 Days calculation',
'My gift to you - a list of things you can do to stop losing all your fucking
money.',
'Gabe Plotkin - confirmed enemy and known fuckup for shorting GME (Melvin Cap
ital)',
'Why GME could squeeze as early as tomorrow',
'GME down today up tomorrow',
'GME Short Raid on 12/29 - Proof is in the NYSE SRR List + Chart',
'GME',
'GameStop GME Short Interest Update for 12/15/20',
'BYND',
'A comprehensive list of all the stuff pumped on this sub in 2020. Congrats r
etards!❤️',
'GME call options?',
'PSTH Bang Bros Rumor',
'Non existent (for now) nuclear energy play',
'What's with all the GME and PLTR autism",
'Why you shouldn't research stocks before buying",
'LETS GO TO THE MOOOOOOOOOONNNNNNNN',
'GME, hypnotism, and general ideals of manifestation. The squeeze HAS to happ
en, and here is why...',
'Devils Advocate to GME mooning 🚀🚀🚀 What are good reasons to why it will
not take off?',
'What are you top 5?',
'Robinhood deleted article on GME?',
'Lesson learned: You have to be willing to lose $350k to make money',
'$ShitFix ($SFI) is going down the crapper',
'If the GME short squeeze hits, how long will the squeeze last? Several weeks
like Shkreli's squeeze or how long are we talking about?',
'Gains were made. Buying GME on monday 🚀🚀🚀',
'Fucking genius idea : we create a WSB trust fund for when the next tesla com
es around 🚀🚀🚀',
'Help Me Understand GME',
'Wtf is the DD behind GME?',
'Does GME still have some run up, or is has it reached the zenith?',
'So who you got, PLTR OR GME?',
'$GME, a word to the paperhands',
'PLTR or GME?',
'GME Explanation',
'WE ARE STRONGER THAN WE THINK!!!',
'I like feet',
'Why is no one talking about GMHI?',
'[DD] GME earnings and option positioning',
'What is your #1 play for Tomorrow?',
'Is GME a good choice for a stock right now? and when should i sell it if i g
ot it?',

```


'A warning to my friends and fellow retards.',
 'GME challenge: No Sell December',
 'CARR Global Carrier',
 'CRSR, BRQS or GME?',
 'What happen to GME ?',
 'I have extra \$200',
 'VTRS TO THE 🍌',
 "If GME hits 100\$, I'll tattoo the most upvoted comment",
 'HOW TO WIN THE \$GME WAR - Strategy Discussion',
 'GME',
 'Is NOKIA still la buy?',
 'You idiots',
 'Today is a green day for most every single stock except for wallstreetbets p
 icks lol',
 'FUCK YOU',
 'Campaign the SEC to look into short sales of GME',
 'Panic over Chinese stocks due to vote in the house',
 'GME breakout imminent = short squeeze imminent = tendies imminent',
 'WSB owns 5.8% of GME (GME survey update)',
 "GME - 'The Big Squeeze'",
 'Liquidity question',
 'GME',
 'Exactly how the GME squeeze will go for you.',
 'GME limit sell',
 'Gay bias for GME short squeeze',
 'CALLING ALL BAG HOLDER AND PROFIT TAKER, Donate to the less fortune this hol
 iday!!!',
 'GME 4Q Bottom Up E-Commerce & Financial Model',
 'I Yolo and achieved tendie crumbs. How to report on taxes?',
 'GME Tendies Preheating',
 'I just want to say thanks to the guy who posted the DD about ZI the other da
 y',
 'GOLDEN NUGGET ONLINE GAMING MERGER VOTE DATE MAKE MONEY ROCKET GO FSSSSSSSSH
 SHHHHHSHSHSHHHH🚀🚀🚀🚀🚀',
 'Covid exemption to withdraw from 401(k)',
 'What's the best play next week?',
 'Starting a Reverse ETF to Hedge all you Fuckin Mamas Boys',
 '12/3 Possibilities and Plays: PLTR, AMD, TSLA, NIO, SPY, and GME or WSB will
 be in Heaven',
 'WSB owns GME (but how much?)',
 '50% daily volume first hour for Chinese EV's',
 'Anyone traded a short squeeze before?',
 "When's the GME investor meeting for its 5th largest holder?",
 'GME- DO NOT SELL AFTER EARNINGS',
 "It's time for you to short BCRX",
 'A Thoughtful GME Squeeze Question',
 'XOM Calls and Print Money',
 'GME - the short Squeeze.',
 '\$GME Earnings 12/8/20 - Will Gamestops's Shops Shock the Stock?',
 'NightMan No More!!!!!!!!!!!! THE SPY PROPHECY: WEEK of 12/07',
 'If you want to sell options, these are what you should be selling...',
 'GME is undervalued cuz price to sales ratio',
 'I can prove WSB and fellow retards actually have power in the market',
 'Question about GME Short Interest',
 'Why you should Short DOORTRASH',
 'I'm proud of you guys but don't let your guard down',
 'Should I buy GME 12/11 \$20?',
 'Everything into TSLA forever',
 'Introducing WSB Roulette! Randomly generated options plays because this is a
 casino!',
 'Earnings pre GME thread',
 'GME',
 'Coincidence? I think not',
 'New Predictions for \$GME stock price. Open Price for Tomorrow, EOW/EOY Price

```

s?',
'GME Management gone WSB Autistic!',
'INVERSING WSB TMROW 20C VLDR',
'BUY THE DIP - $GME',
"Simple click-clack on GME so you don't have to",
'Your Tendie Stimulus IV: Don't Buy the Dip (yet), Jerome has forsaken us',
'GME MOON MISSION CONFIRMED',
'GME play to save taxes & be charitable per the CARES Act',
'What should I put 3k in?',
'Panic selling GME is worst than FOMO',
'Bought GME with leverage',
'New to stocks',
'Thought experiment on GME',
'Thought experiment on GME',
'The squeeze is coming anytime now',
'Uranium stonks will be the next big thing',
'Brutal week after 2 months of drinking WSB DD Koolaid - conviction shattere
d',
'I bought $BB calls and I'm worried",
'GME Q3 Call Thoughts on the clash between Cohen and Sherman [Insert Rocket E
moji Here]',
'$VSPR HydraFacial - Major Becky Stock + Undervalued SPAC set to 🚀',
'Someone please explain to me why I would not use poor mans covered calls for
theta gang?',
'GME Rises from the Dead - 12/14 - 12/31 Possibilities',
'GME Gang... Can ya feel it? (Fresh confirmation bias within)',
'WSB + Cohen = GME Hostile Takeover?',
'High Leverage Makes Money (Up Another $120K)',
'Down about $220K from mid-day peak.',
'Found the app for the Autists',
'DOUBLED ON BOOMER OPTIONS; What's Next?",
'How many GME shares do we have?',
'Big steppa',
'GME',
'What does GME mean?',
'GME Short Squeeze Target Prices',
'Brand new SPAC headed by Netflix VP/DraftKings CEO and rumored to merge with
OnlyFans: $HZON',
'What is the average time you hold positions?',
'What the fuck is happening with GME. Time for a support group ?',
'Gang Stop',
'I'm declaring myself a moderator',
'We need the tri-weekly GME live stream back.',
'We need the tri-weekly GME live stream back.',
'Hey everyone, it's me - JP',
'GME🚀 employee improvements w/ Cohen? 🚀🚀🚀🚀',
'What to do with 500k? Ark ETFS or PLTR/GME?',
'Formal GME exit price & strategy',
'TRUMP is gonna moon GME🚀🚀🚀',
'BBY - Your next 10 bagger',
'GME Technical Analysis: 2 Strong and Perfectly Positioned Bull Flags 🚀🚀🚀🚀
🚀🚀🚀',
'Ackman, Cohen, and Takeover Strategy',
'How bullish 0% interest rate world really is - Answer: VERY BULLISH',
'Wife's boyfriend's girlfriend's only fans',
'GME Jan 15 Calls: J take the wheel. Thoughts?',
'@cramer and wsb retards',
'It's too bad we can't talk about one of the hottest tickers of the week beca
use of the ridiculously arbitrary market cap rule",
'Lantern - Follow the smart money',
'Pulling out of my girl TSLA (was hitting it raw for a while)',
'Something will happen by early January to GME.',
"The Perfect Crime (hypothetically, ironically hahaha I don't mean this at al
l)",

```

'Borrowed against my 800 shares of GME to buy Steel stocks on Margin in true WSB fashion. Why? Because Steel is hard. Almost as hard as my hands. 💎👊🏻🌈🤪',
 'How good of an idea is it to buy super far OTM GME calls given Ryan Cohen might buy out GME? (30+ strike)',
 'Understanding the GME short squeeze',
 'Christmas miracles do happen',
 'Why the time to buy is now...',
 'Market Makers sent us a message in the VIX!',
 'Can you feel the tingle....',
 'Selling Put Leaps on AMC',
 "GME Discussion: If Cohen is buying GME at 16 then I'm willing to do my part",
 "While everyone's talking about steel...",
 'Future possibilities',
 '\$GME: ShyGlizzy Concedes Defeat (StockTwits)',
 'Is WSB committing market manipulation?',
 'GME DD Fresh Perspectives Only Pt.1',
 'GME short squeeze is a myth that assumes massive incompetence from people that actually know what they are doing.',
 'I have a dream',
 'New to the market, what to do with \$1000?',
 "GME GANG DON'T BE SCARED 🍌💎🐔",
 'I'm stupid and need advice',
 'Reading news paper is NOT research. You CANNOT do research on stocks and anyone that says otherwise is lying.',
 'Was yesterday's GME short raid engineered around uptick rule? (SHADOW WARZ)',
 'What will it take to change GME shorts investment thesis?',
 'GME autists. Please DO NOT let the SHORTS Manipulate you like this again.',
 'Take your money and get out!',
 'Why a \$150 price target for GME is realistic',
 'GME Debt Payment',
 '[Advice Request] How do I get started as an international university student?',
 'The GME short squeeze OTM options play fit for WSB',
 'Holding \$GME to send a message. Join us or fear us. 🚀🚀🚀',
 'Can u tell me how to make lots of monies',
 'Brand new close to NAV SPAC merging with a fintech unicorn: AJAX',
 'GME 🐻 Short position',
 'Why GME will continue crashing on Monday using Technical fundamentals. (Not enough retail to break the balance board.)',
 'Is there a f. way as a European wanna be retard to buy options on companies located in the USA?',
 'Gme 🚀🚀🚀🚀=gold?',
 'What's our best ALL-IN play to close out the year?',
 'WHY THE FUCK WONT ROBINHOOD LET ME OPTIONS TRADE.',
 'GME 🚀🚀🚀🚀🚀 = gold?',
 'VALE - hold or dump for GME?',
 'Best investment right now?',
 'GME mooning on Monday and Tuesday 🚀🚀🚀🚀🚀🚀🚀🚀🚀🚀🚀',
 'Autism in the movies - a retard's watchlist for the holidays',
 'Put on GME (advice from a 15 year old)',
 'GME Position Tracker',
 'GLSI made a silent but deadly blast to space leaving only scorched earth behind',
 'Goodbye GME - HELLO GPRO - RIDE THIS B to \$20!!!',
 'How do u feel about GoDaddy(\$GDDY) as an investment?',
 'So..... fuck Uncle Sam, right?',
 'First yearly christmas gains/loss Post!',
 'Quick! You've caught a case of the "Paper Hands!"',
 'Why \$GME can still pop next week',
 'I'm really retarded.',
 'Merry Christmas to everybody (except....)',
 "Fringe GME DD: Melvin Capital's 2nd largest *long* holding dumping in Hong K

```

ong trading",
'GME (Help)',
'letdowns suck',
'Is SCR the next DKNG?',
'Multiple expansion is how GME gets to $100+',
'WSB MIGHT GO DOWN',
'Game changer for GameStop',
'New Member. Too late to get in on the GME action?',
'GME about to do an offering.',
'Mod team compromised - what r we gonna do',
'GME DD Fresh Perspectives Pt 2. Gamestop vs Amazon',
'Gather Around Little Speds, it's GME Time.",
'When should I sell GME?',
'The put options that messed up our GME run today',
'MM's holding RKT to 20.75 = unsustainable 🚀🚀🚀🚀🚀🚀",
'First data about PlayStation 5 sold = GME 🚀😄',
"GME Gang, don't let Robinhood lend your shares to the shorts and Melvin! DD
inside for more tendies",
'Real talk about GME, no rocket emojis.',
'5th trade: GME Gang rise up',
'To be honest, I don't know shit.',
'Is there or can there be an ETF for this sub?',
'Why do you all blindly follow $GME?',
'You guys make me broke lol - More GME and PLTR ?',
'122% GME Shares Held By Institutions. WTF & Does It Matter? 🚀🚀🚀',
'GME battle of the autistics vs hedge fund shorts. Threshold securities editi
on.',
'Talking about pump&dumps: starting to become tired of GME posts.',
'Stimmy checks to take GME to the moon',
'GME 🚀🚀🚀 QUESTION',
'help me pick 4 plays to sell covered calls on',
'How fucked are my GME calls????',
'Brand new SPAC at NAV with Uber SVP and Google CEO: $XPOA',
'I mean sure -- anything's possible",
'Hello $BABA holders, please don't $ROPE",
'PSTH Bumble rumor',
'Confession from a r/WSB doubter',
'Sold 7 18$ 12/24 GME covered call contracts and don't want to lose my share
s!',
'For all those brains that think they are too smart for GME',
'GME: This SHaREz not Options thing is retarded',
'Desperate to make back losses on a YOLO. What short term positions will prin
t?',
'I don't understand the GME hype',
'House votes to give 2000$ stimulus checks bill now goes to congress for reje
ction',
'GameStop DD and potential profits.',
'Question on Incentive Structures for HF',
'Gamestop evolves into technology, what would you like to see them do?',
'GME DD fresh perspectives Pt 3 Future of Gamestop + Squeeze teaser. Christma
s Edition.',
'Changes to Calls/Puts Open Interest Holdings Wednesday, December 23 End of D
ay',
'Need Help. Lost my entire portfolio due to bad call options. Wanting to add
my savings to my portfolio so I can get back the money lost as well as more. A
ny suggestions or advice on what to do next?',
'***Daytrading GME, Options trading GME, Buying GME stock',
'This market has taught me that a much simpler strategy is needed',
'FIRST GAINS',
"SIR_JACK_A_LOT's next move: ALL in STIC -> BARK merger (BarkBox -- bitche
s love spoiling their dogs)",
'GME out @ $21.85',
'WSB Lurker Seeks Advice On How To Become More Autistic',
'Holding CRSR is fine even with the dip',

```

'My analysis on why GameStop will beat Amazon. The answer - Microsoft GPUs.'
 'CSRS holders who are panicking at the drop assemble',
 "What's your excuse?",
 'Meme Stock ETF PROPOSAL',
 "Whats the game plan for GME when Citron shows up soon? He's attacked every single meme stock that wsb has gone hard on i'm sure he already wrote up a hit piece on GameStop and is waiting for just the right moment to buy puts and release his report.",
 'VAR1 DE000A0TGJ55 - Short Squeeze inc?',
 'Fuckers. I have always hated GME and now I hate y'all. Christmas cash from in-laws I also hate going into GME calls at market open so wife can hate me too.',
 'Dump all NIO and full retard into GME?',
 'Is it worth selling my position in PLTR to get more GME shares?',
 'Does the day trade limit prevent any bad behavior or just make it harder for us poor slobs to more easily make tendies?',
 'Flywheel of Basic Btches DD on \$STIC aka BarkBox, \$370k position in shares',
 'Need help, TSLA, IWM',
 'Why is Amazon performing so poorly?',
 'Aphria Tilray Merger',
 'The day of GME short squeeze',
 'The boomer pessimist long term potential of GME',
 'GME DD',
 'Lost 70% of Portfolio',
 'Good price chart website/software?',
 'Charge of the Bill Brigade; GME to infinity and beyond',
 'Lesson of the day - How options work - What are the benefits of call options, and why buy them instead of the stock?',
 'HOLD THE LINE',
 'What did expect?',
 'GME is dead. Only Autists and Boomers buy games in store. PEG is garbage. You can watch it bounce down the stairs. Has anyone told you about our lord and savior 🚀PLTR🚀?',
 'Unlimited money making guide',
 'Is GME the Most Shorted Stock in Wall Street History?',
 'Meme stock explosion',
 'For all of the paper handed pussy ass one day holding GME sellers',
 'Why are we sleeping on Moderna?',
 'Tesla joining the S&P will crash the market. I promise I'm not a 🌈🐻.',
 'Calculating what the GME price should be for shorters to get in the red',
 'GME is the Most WSB Move in the History of WSB',
 '"GameStop is the Real Game, not just the place you go to buy Games"',
 'if I am dying, can I yolo with borrowed funds?',
 'CARR (Global Carrier) must Buy Now DD',
 'A message to the paper-handed sissies and 🌈🐻',
 'I sold GME and it immediately climbed up 1% AH',
 'Predictions on Year End price of GME and PLTR',
 'All is quiet on the GME front...',
 'GME question',
 'We are going to be ok',
 'Why are you guys holding GME?',
 'I went ALL IN on GME puts today',
 'Potential GME gain this week vs current positions',
 '\$GME Price Action, Short Raids, and You',
 'I'm an absolute retard and fucked up on GME calls',
 'Buy GME stocks before Market close',
 'GME stocks before earnings call',
 'Next Major Position',
 'GME reverse tax-selling',
 'GME short squeeze and earnings report',
 'Set A High Sell Order on GME in Margin Accounts so Brokers CANT Lend to Shorters',
 'CRM for Retards',

'GME Call Option Pricing',
 'How many people in here are actually AUTISTIC?',
 'GME...Ready For Takeoff',
 'It is time for GME longs to analyze one or two things',
 'Trending tickers on WSBS daily chat, Dec 9',
 'Did every GME gang member become retarded after earnings?',
 'Im DOUBLING down on GME puts',
 'GME is probably Michael Burry's largest position',
 'BBBY: Another fat girl ready for a short squeeze',
 'Lost 35k last 2 days and hitting breaking point, tell me what to put 75k into for recovery play...',
 'GME MOON MISSION',
 'GME : 2 Bulls 1 Cup',
 'GME entry price...',
 'GME Paper Hands Question',
 'Why GME earnings has me worried',
 'Real world gme ? Why? Short?',
 'If NIO hits \$60 within the next 2 weeks and XPEV hits \$80, I'm going to bed my hot friend",
 'PLTR may never be this cheap again',
 'Potentially Solid Run Up to ER Play with FedEx - Shipping it on this play',
 "🚀 Buy high, sell low 🚀 GME Dec 18 '20 17.00 Put (20K open interest)",
 'GME mooning wednesday?',
 'Taking the cue from Cohen',
 'It is today boys. GME make or break!!🚀🚀🚀🚀',
 'Gamestop Shelf Offering',
 "🚀🚀🚀 DD - I had a vision on Christmas Eve GME 🚀🚀🚀",
 'I'M SORRY GME...P.S. I love you long term",
 'Let's destroy Melvin capital",
 'Petition Ryan Cohen to take over GameStop (GME)',
 'Evaluation of my first week trading, hype effect & international options trading platform',
 'GME Post E/R',
 'GME',
 'Message to those who suck Cohen Wenis',
 'It's starting to feel like 1999",
 'I did the click-clack for GME so you don't have to",
 'GME Survey 2: Diamond Hands, or Paper Hands?',
 'Merry Christmas fuckers \$GME bought me \$ROPE',
 'The difference between TSLA and GME',
 'Tomorrow's the day you make up your lost money on GME',
 'GME Earnings Call',
 'To those looking to make extra profit on GME and those who want to lower the average price of bought the stock high',
 'GME 🚀🚀🚀🚀',
 'Next Big Bet',
 'Predictions for GME',
 'Made myself some apparel to let the world know about my recent autism diagnosis...',
 'Making of Diamonds: What you need to be aware of going into this week for GME',
 'Trying to convince my brother to join the gme autist gang',
 'Here is how im gonna play for the next couple months',
 'GME, why?',
 'What are your plays for this week?',
 'Nio thread because no one else is talking about it🚀🚀🚀',
 'What for Yolo',
 'A sign that GME is gonna moon',
 'GME INCOMING SQUEEZE 12/28🚀',
 'STIMULUS APPROVED LAUNCH THREAD 🚀🚀🚀',
 'A summary of the GME meme machine',
 'A Happy Ending',
 'Told my 49.5k Instagram followers to comment "we want Cohen" on @gamestop Instagram 🚀',

Exercise 3: Creation of large documents for each stock

In [117...

localhost:8888/nbconvert/html/Køben/CSS/Assignment 2%2C Part 1.ipynb?download=false


```
for i in range(len(list_stocks)):
    if 'spce' in list_stocks[i]:
        words_spce.append(list_tokens[i])

words_pltr=[]
for i in range(len(list_stocks)):
    if 'pltr' in list_stocks[i]:
        words_pltr.append(list_tokens[i])

words_msft=[]
for i in range(len(list_stocks)):
    if 'msft' in list_stocks[i]:
        words_msft.append(list_tokens[i])

words_rope=[]
for i in range(len(list_stocks)):
    if 'rope' in list_stocks[i]:
        words_rope.append(list_tokens[i])

words_aapl=[]
for i in range(len(list_stocks)):
    if 'aapl' in list_stocks[i]:
        words_aapl.append(list_tokens[i])

words_amzn=[]
for i in range(len(list_stocks)):
    if 'amzn' in list_stocks[i]:
        words_amzn.append(list_tokens[i])

words_nio=[]
for i in range(len(list_stocks)):
    if 'nio' in list_stocks[i]:
        words_nio.append(list_tokens[i])

words_zm=[]
for i in range(len(list_stocks)):
    if 'zm' in list_stocks[i]:
        words_zm.append(list_tokens[i])

words_amd=[]
for i in range(len(list_stocks)):
    if 'amd' in list_stocks[i]:
        words_amd.append(list_tokens[i])

words_baba=[]
for i in range(len(list_stocks)):
    if 'baba' in list_stocks[i]:
        words_baba.append(list_tokens[i])

words_gme=[]
for i in range(len(list_stocks)):
    if 'gme' in list_stocks[i]:
        words_gme.append(list_tokens[i])

words_dis=[]
for i in range(len(list_stocks)):
    if 'dis' in list_stocks[i]:
        words_dis.append(list_tokens[i])

words_Other=[]
for i in range(len(list_stocks)):
    if 'Other' in list_stocks[i]:
        words_Other.append(list_tokens[i])
```

Removing non english words from the tokens and flattening the list for the 5 chosen tags (couldnt remove non-english words when cleaning of the tokens because tags such as 'tsla', 'gme' would get removed then)

In [25]:

```
words_gme1=[]
for i in range(len(words_gme)):
    test2=[]
    for j in words_gme[i]:
        if j in english_words:
            test2.append(j)
    words_gme1.append(test2)
words_gme_flat = [x for l in words_gme1 for x in l]

words_spy1=[]
for i in range(len(words_spy)):
    test3=[]
    for j in words_spy[i]:
        if j in english_words:
            test3.append(j)
    words_spy1.append(test3)
words_spy_flat = [x for l in words_spy1 for x in l]

words_tsla1=[]
for i in range(len(words_tsla)):
    test4=[]
    for j in words_tsla[i]:
        if j in english_words:
            test4.append(j)
    words_tsla1.append(test4)
words_tsla_flat = [x for l in words_tsla1 for x in l]

words_spce1=[]
for i in range(len(words_spce)):
    test5=[]
    for j in words_spce[i]:
        if j in english_words:
            test5.append(j)
    words_spce1.append(test5)
words_spce_flat = [x for l in words_spce1 for x in l]

words_amzn1=[]
for i in range(len(words_amzn)):
    test6=[]
    for j in words_amzn[i]:
        if j in english_words:
            test6.append(j)
    words_amzn1.append(test6)
words_amzn_flat = [x for l in words_amzn1 for x in l]
```

Exercise 4: TF and IDF

Choosing the stocks: 'gme', 'spy', 'tsla', 'spce' and 'amzn' and finding the 5 most common terms from the token list and how many times they occur for each chosen stock

In [214...]

```
#gme
words_to_count_gme = (word for word in words_gme_flat)
top_gme=Counter(words_to_count_gme).most_common(len(words_gme_flat))

#spy
words_to_count_spy = (word for word in words_spy_flat)
```

```

top_spy=Counter(words_to_count_spy).most_common(len(words_spy_flat))

#tsla
words_to_count_tsla = (word for word in words_tsla_flat)
top_tsla=Counter(words_to_count_tsla).most_common(len(words_tsla_flat))

#spce
words_to_count_spce = (word for word in words_spce_flat)
top_spce=Counter(words_to_count_spce).most_common(len(words_spce_flat))

#amzn
words_to_count_amzn = (word for word in words_amzn_flat)
top_amzn=Counter(words_to_count_amzn).most_common(len(words_amzn_flat))

```

In [215...

```

print(top_gme[0:5])
print(top_spy[0:5])
print(top_tsla[0:5])
print(top_spce[0:5])
print(top_amzn[0:5])

[('buy', 409), ('short', 393), ('going', 388), ('like', 387), ('get', 378)]
[('spy', 7038), ('market', 2447), ('like', 2242), ('p', 2129), ('going', 1939)]
[('like', 1249), ('buy', 1094), ('get', 1032), ('market', 1020), ('stock', 1017)]
[('like', 391), ('go', 364), ('stock', 335), ('buy', 333), ('get', 331)]
[('like', 343), ('market', 299), ('going', 265), ('time', 262), ('one', 260)]

```

Describe similarities and differences between the stocks: We can see from the results that a lot of the same words go again and again such as 'buy', 'going', 'like', 'stock', 'market' which makes sense since this subreddit is about the stock market. However, the 'p' appeared and I'm not sure of, the 'p' is not in the stopwords list, maybe it's an abbreviation for something like prize or something? Other than that all the results are quite alike, and expectable.

Why aren't the TFs not necessarily a good description of the stocks? TF is how many times a word appears in a text. When discussing on reddit a lot of words not necessarily used to describe the stock may be used. Even though we've done some cleaning steps there will be words left. It's a lot of talk about 'buying', 'going', the 'market' etc which will occur just because the stocks are talked about which isn't really describing the stock.

Calculating the TF (number the given word occurs / total words in the document) for the top 5 tokens for the 5 chosen stocks.

In [237...

```

#gme
tf_gme=[]
for i in range(len(top_gme)):
    tf_gme.append((top_gme[i][0], top_gme[i][1]/len(words_gme_flat)))

#spy
tf_spy=[]
for i in range(len(top_spy)):
    tf_spy.append((top_spy[i][0], top_spy[i][1]/len(words_spy_flat)))

#tsla
tf_tsla=[]
for i in range(len(top_tsla)):
    tf_tsla.append((top_tsla[i][0], top_tsla[i][1]/len(words_tsla_flat)))

#spce

```

```
tf_spce=[]
for i in range(len(top_spce)):
    tf_spce.append((top_spce[i][0], top_spce[i][1]/len(words_spce_flat)))

#amzn
tf_amzn=[]
for i in range(len(top_amzn)):
    tf_amzn.append((top_amzn[i][0], top_amzn[i][1]/len(words_amzn_flat)))
```

In [253...

```
print(tf_gme[0:5])
print(tf_spy[0:5])
print(tf_tsla[0:5])
print(tf_spce[0:5])
print(tf_amzn[0:5])
```

```
[('buy', 0.005571144468357534), ('short', 0.0053532023864657965), ('going', 0.005285095485874629), ('like', 0.005271474105756395), ('get', 0.005148881684692293)]
[('spy', 0.018888635418204264), ('market', 0.006567276338213389), ('like', 0.006017095852175897), ('p', 0.005713825633043035), ('going', 0.005203902255739992)]
[('like', 0.005998031070665354), ('buy', 0.005253679736835787), ('get', 0.004955939203303959), ('market', 0.004898312003265542), ('stock', 0.004883905203255936)]
[('like', 0.0067212156633547635), ('go', 0.00625709079657579), ('stock', 0.005758586310035411), ('buy', 0.005724206690274005), ('get', 0.0056898270705126)]
[('like', 0.004922573515693394), ('market', 0.0042911063591612965), ('going', 0.0038031544654774036), ('time', 0.0037600998866229424), ('one', 0.0037313968340533014)]
```

IDF calculation

Calculation by using the formula: $\log(\text{Number of documents}/\text{number of documents the word appear in})$. where the number of documents will be the length of the list_tokens.

Need to find number of documents the words appear in. (Since IDF for one word is the same, only needs to calculate one time each word)

In [193...

```
#buy
docs_with_buy=0
for i in range(len(list_tokens)):
    if 'buy' in list_tokens[i]:
        docs_with_buy+=1

#short
docs_with_short=0
for i in range(len(list_tokens)):
    if 'short' in list_tokens[i]:
        docs_with_short+=1

#going
docs_with_going=0
for i in range(len(list_tokens)):
    if 'going' in list_tokens[i]:
        docs_with_going+=1

#like
docs_with_like=0
for i in range(len(list_tokens)):
    if 'like' in list_tokens[i]:
        docs_with_like+=1

#get
docs_with_get=0
for i in range(len(list_tokens)):
```

```

        if 'get' in list_tokens[i]:
            docs_with_get+=1
#spy
docs_with_spy=0
for i in range(len(list_tokens)):
    if 'spy' in list_tokens[i]:
        docs_with_spy+=1
#market
docs_with_market=0
for i in range(len(list_tokens)):
    if 'market' in list_tokens[i]:
        docs_with_market+=1
#p
docs_with_p=0
for i in range(len(list_tokens)):
    if 'p' in list_tokens[i]:
        docs_with_p+=1
#stock
docs_with_stock=0
for i in range(len(list_tokens)):
    if 'stock' in list_tokens[i]:
        docs_with_stock+=1
#go
docs_with_go=0
for i in range(len(list_tokens)):
    if 'go' in list_tokens[i]:
        docs_with_go+=1
#time
docs_with_time=0
for i in range(len(list_tokens)):
    if 'time' in list_tokens[i]:
        docs_with_time+=1
#one
docs_with_one=0
for i in range(len(list_tokens)):
    if 'one' in list_tokens[i]:
        docs_with_one+=1

#know
docs_with_know=0
for i in range(len(list_tokens)):
    if 'know' in list_tokens[i]:
        docs_with_know+=1
#money
docs_with_money=0
for i in range(len(list_tokens)):
    if 'money' in list_tokens[i]:
        docs_with_money+=1
#would
docs_with_would=0
for i in range(len(list_tokens)):
    if 'would' in list_tokens[i]:
        docs_with_would+=1
#also
docs_with_also=0
for i in range(len(list_tokens)):
    if 'also' in list_tokens[i]:
        docs_with_also+=1

```

In [194...

```

#Making a list of all the numbers from above to have easier code to iterate t
list_with_numbers=[('buy', docs_with_buy), ('short', docs_with_short), ('going

```

```
In [195... IDF_values=[]
for i in range(len(list_with_numbers)):
    IDF_values.append((list_with_numbers[i][0],math.log(len(list_tokens)/list

IDF_values
```

```
Out[195... [('buy', 1.7092449465395803),
('short', 2.516102451794195),
('going', 1.6616262711589307),
('like', 1.8765250164535898),
('get', 1.6292332442144688),
('spy', 2.458342081577906),
('market', 1.632774170417683),
('p', 2.5026453109141555),
('stock', 1.7644847241540806),
('go', 1.738141600724539),
('time', 1.7769143434253172),
('one', 1.8765250164535898),
('know', 1.7780650915384952),
('money', 1.7356554115986131),
('would', 1.9368320519235682),
('also', 2.129326636430573)]
```

Exercise 5: TF-IDF

Top 10 tf for the 5 stocks

```
In [254... print(tf_gme[0:10])
print(tf_spy[0:10])
print(tf_tsla[0:10])
print(tf_spce[0:10])
print(tf_amzn[0:10])

[('buy', 0.005571144468357534), ('short', 0.0053532023864657965), ('going', 0.005285095485874629), ('like', 0.005271474105756395), ('get', 0.005148881684692293), ('know', 0.00479472580161822), ('stock', 0.00464489062031765), ('time', 0.004440569918544147), ('go', 0.004345220257716512), ('one', 0.004249870596888877)]
[('spy', 0.018888635418204264), ('market', 0.006567276338213389), ('like', 0.006017095852175897), ('p', 0.005713825633043035), ('going', 0.005203902255739992), ('get', 0.005123388038271092), ('go', 0.005104601387528347), ('money', 0.004887213000362314), ('time', 0.00481206639739134), ('buy', 0.004656405576951463)]
[('like', 0.005998031070665354), ('buy', 0.005253679736835787), ('get', 0.004955939203303959), ('market', 0.004898312003265542), ('stock', 0.004883905203255936), ('going', 0.004715825869810551), ('go', 0.0045621533363747685), ('money', 0.00453333973635556), ('one', 0.004514130669676087), ('time', 0.004485317069656878)]
[('like', 0.0067212156633547635), ('go', 0.00625709079657579), ('stock', 0.005758586310035411), ('buy', 0.005724206690274005), ('get', 0.0056898270705126), ('going', 0.0056210678309897896), ('money', 0.005242892013614329), ('market', 0.004761577336954653), ('time', 0.0047443875270739505), ('one', 0.004349021899817788)]
[('like', 0.004922573515693394), ('market', 0.0042911063591612965), ('going', 0.0038031544654774036), ('time', 0.0037600998866229424), ('one', 0.0037313968340533014), ('stock', 0.0036165846237747384), ('would', 0.003530475466065816), ('get', 0.0035017724134961755), ('also', 0.003487420887211355), ('buy', 0.0034300147820720732)]
```

```
In [271... #The function makes it able to sort by the 2nd element, the tf and not
#alphabetically by words
```

```
def takeSecond(elem):
    return elem[1]
```

Top 10 TF-IDF for \$GME. Calculated by TF*IDF for the given word.

In [292...

```
TF_IDF_gme=[]
for i in range(len(tf_gme[0:10])):
    for j in range(len(IDF_values)):
        if tf_gme[i][0]==IDF_values[j][0]:
            TF_IDF_gme.append((tf_gme[i][0], tf_gme[i][1]*IDF_values[j][1]))

TF_IDF_gme.sort(key=takeSecond, reverse=True)
TF_IDF_gme
```

Out[292...

```
[('short', 0.013469205649537127),
 ('like', 0.009892053033039193),
 ('buy', 0.009522450528982053),
 ('going', 0.008781853504912757),
 ('know', 0.008525334571356286),
 ('get', 0.008388729211227686),
 ('stock', 0.008195838544917065),
 ('one', 0.007974988491752528),
 ('time', 0.007890512381244087),
 ('go', 0.007552608094248071)]
```

Top 10 TF-IDF for \$SPY

In [293...

```
TF_IDF_spy=[]
for i in range(len(tf_spy[0:10])):
    for j in range(len(IDF_values)):
        if tf_spy[i][0]==IDF_values[j][0]:
            TF_IDF_spy.append((tf_spy[i][0], tf_spy[i][1]*IDF_values[j][1]))

TF_IDF_spy.sort(key=takeSecond, reverse=True)
TF_IDF_spy
```

Out[293...

```
[('spy', 0.04643472731215443),
 ('p', 0.01429967892791626),
 ('like', 0.011291230893007202),
 ('market', 0.010722879175030045),
 ('go', 0.008872520026779225),
 ('going', 0.008646940700680792),
 ('time', 0.008550629803039663),
 ('money', 0.008482517691713945),
 ('get', 0.008347194114962014),
 ('buy', 0.007958937701443007)]
```

Top 10 TF-IDF for \$TSLA

In [294...

```
TF_IDF_tsla=[]
for i in range(len(tf_tsla[0:10])):
    for j in range(len(IDF_values)):
        if tf_tsla[i][0]==IDF_values[j][0]:
            TF_IDF_tsla.append((tf_tsla[i][0], tf_tsla[i][1]*IDF_values[j][1]))

TF_IDF_tsla.sort(key=takeSecond, reverse=True)
TF_IDF_tsla
```

Out[294...

```
[('like', 0.011255455353569446),
 ('buy', 0.00897982554092396),
 ('stock', 0.00861757612536173),
 ('one', 0.008470879129187574),
```



```
( 'get', 0.008074380906328579),
( 'market', 0.007997837317578874),
( 'time', 0.00797002423588372),
( 'go', 0.007929668502837236),
( 'money', 0.007868315646020557),
( 'going', 0.007835940155488127)]
```

Top 10 TF-IDF for \$SPCE

In [302...

```
TF_IDF_spce=[]
for i in range(len(tf_spce[0:10])):
    for j in range(len(IDF_values)):
        if tf_spce[i][0]==IDF_values[j][0]:
            TF_IDF_spce.append((tf_spce[i][0], tf_spce[i][1]*IDF_values[j][1])

TF_IDF_spce.sort(key=takeSecond, reverse=True)
TF_IDF_spce
```

Out[302...

```
( 'like', 0.012612529333264923),
( 'go', 0.010875709813039024),
( 'stock', 0.010160937576780297),
( 'buy', 0.0097840713582989),
( 'going', 0.009340113979938983),
( 'get', 0.00927005541711055),
( 'money', 0.00909985389585686),
( 'time', 0.008430370247625872),
( 'one', 0.008161048392112596),
( 'market', 0.007774580486225774)]
```

Top 10 TF-IDF for \$AMZN

In [295...

```
TF_IDF_amzn=[]
for i in range(len(tf_amzn[0:10])):
    for j in range(len(IDF_values)):
        if tf_amzn[i][0]==IDF_values[j][0]:
            TF_IDF_amzn.append((tf_amzn[i][0], tf_amzn[i][1]*IDF_values[j][1])

TF_IDF_amzn.sort(key=takeSecond, reverse=True)
TF_IDF_amzn
```

Out[295...

```
( 'like', 0.009237332347530552),
( 'also', 0.007425858187583479),
( 'market', 0.0070064076257536296),
( 'one', 0.007002059505416745),
( 'would', 0.006837938041206071),
( 'time', 0.006681375421252215),
( 'stock', 0.006381408322261059),
( 'going', 0.006319421373112654),
( 'buy', 0.005862735432812751),
( 'get', 0.005705204029741104)]
```

Since I only have the IDF for the top words the words doesn't change, but the order of them change. This makes sense as now they are dependant on other documents as well.

(was told to only find the IDF for the top words of each stock)

Exercise 6: Word cloud

Creation of the wordcloud. The wordclouds can be found in a zip separate folder in the assignment. I did one wordcloud based on TF-IDF and one based on TF.

```
In [281... wc= WordCloud(  
    background_color='white',  
    height=600,  
    width=400  
)
```

GME:

```
In [298... wc.generate(str(TF_IDF_gme))  
wc.to_file('GME.png')
```

Out[298... <wordcloud.wordcloud.WordCloud at 0x7f90b72b8340>

```
In [313... wc.generate(str(top_gme))  
wc.to_file('GME_tf.png')
```

Out[313... <wordcloud.wordcloud.WordCloud at 0x7f90b72b8340>

SPY:

```
In [299... wc.generate(str(TF_IDF_spy))  
wc.to_file('SPY.png')
```

Out[299... <wordcloud.wordcloud.WordCloud at 0x7f90b72b8340>

```
In [314... wc.generate(str(top_spy))  
wc.to_file('SPY_tf.png')
```

Out[314... <wordcloud.wordcloud.WordCloud at 0x7f90b72b8340>

TSLA:

```
In [300... wc.generate(str(TF_IDF_tsla))  
wc.to_file('TSLA.png')
```

Out[300... <wordcloud.wordcloud.WordCloud at 0x7f90b72b8340>

```
In [315... wc.generate(str(top_tsla))  
wc.to_file('TSLA_tf.png')
```

Out[315... <wordcloud.wordcloud.WordCloud at 0x7f90b72b8340>

SPCE:

```
In [303... wc.generate(str(TF_IDF_spce))  
wc.to_file('spce.png')
```

Out[303... <wordcloud.wordcloud.WordCloud at 0x7f90b72b8340>

```
In [316... wc.generate(str(top_spce))  
wc.to_file('SPCE_tf.png')
```

Out[316... <wordcloud.wordcloud.WordCloud at 0x7f90b72b8340>

AMZN:

In [304...

```
wc.generate(str(TF_IDF_amzn))  
wc.to_file('AMZN.png')
```

Out[304...] <wordcloud.wordcloud.WordCloud at 0x7f90b72b8340>

In [317...

```
wc.generate(str(top_amzn))  
wc.to_file('AMZN_tf.png')
```

Out[317...] <wordcloud.wordcloud.WordCloud at 0x7f90b72b8340>

Lastly I've just added the dataframe to show that the textsubmission, tokens and stocks columns has been added

In [319...

```
df
```

Out[319...

	created_utc	title	selftext	score	textsubmission	tokens
0	1586173811	What is the Fed actually buying?	Okay, I may actually just be retarded. On my d...	1	What is the Fed actually buying? Okay, I may a...	[okay, may, actually, retarded, defense, every...
1	1586173320	I didn't learn about puts because I was lazy	Beginning of the this virus shit, everyone was...	1	I didn't learn about puts because I was lazy B...	[beginning, virus, shit, everyone, talking, pu...
2	1586173268	HOT TAKE	Literally everyone has free time on their hand...	1	HOT TAKE Literally everyone has free time on t...	[literally, everyone, free, time, hands, mean,...
3	1586172639	Fuck you Gordon	Gordon I believed in you, I can't even begin t...	1	Fuck you Gordon Gordon I believed in you, I ca...	[gordon, believed, even, begin, describe, disa...
4	1586171822	Can't find a picture	Someone uploaded a ohoto of the stock market h...	1	Can't find a picture Someone uploaded a ohoto ...	[someone, uploaded, ohoto, stock, market, hist...
...
82237	1602007302	Hurricane Delta (BECN) 🌪️	\nHurricane Delta is looking like it is going ...	1	Hurricane Delta (BECN) 🌪️ \nHurricane Delta is ...	[hurricane, delta, looking, like, going, hit, ...
82238	1602006818	Made 40k on Nike. Next play? CROCS motherfucker	# 1. Introduction\n\n[Proof that I'm lucky](h...	1	Made 40k on Nike. Next play? CROCS motherfucke...	[introduction, proof, lucky, way, going, struc...

	created_utc	title	selftext	score	textsubmission	tokens
82239	1602006029	Please screenshot the whole timeline, not just...	I could nut over your retarded failures just f...	1	Please screenshot the whole timeline, not just...	[could, nut, retarded, failures, fine, really, ...
82240	1602005968	What is your price target for Tesla in 40 years?	I am 26 and currently max out my roth each yea...	1	What is your price target for Tesla in 40 year...	[currently, max, roth, year, robinhood, specul...
82241	1601822856	White House infected folks will infect Senators	They all gonna get the rona, \n\nThis means no...	1	White House infected folks will infect Senator...	[gonna, get, rona, means, checkies, stimulus, ...

82242 rows × 7 columns