I've been spending some time researching the practice of
[sub-pennying](https://www.investopedia.com/terms/s/sub-pennying.asp) and how it might be used to
negate buying pressure on GME, but now I feel that I've gone down a rabbit hole and honestly don't
understand what I've uncovered after digging through the tape for these subpenny priced trades. I'm not
sure how to proceed at this point so while I think of something new to try, I thought I'd share some findings
with the hopes someone else might find this useful (or better yet, have an explanation for it).

For context, sub penny trades are those which transact at a price with less than a penny precision. In other
words, securities that are worth more than $1 typically trade in whole penny increments ($185.25 and not
$185.251234); the SEC in fact has a rule on the books preventing the listing of orders with subpenny
precision so you as an investor could never place a limit order at $185.251234 for example. However,
orders placed in dark pools effectively have an exemption to this rule and in fact nearly all retail trades get
executed in dark pools at subpenny prices. This is especially unfair to retail investors because their trades
don't go through at just any subpenny fraction, in fact when a retail investor submits a 'buy' order it will get
transacted at a fraction > 0.005000, usually around 0.009000. Conversely, a 'sell' order will hit the tape at
a fraction < 0.005000, typically closer to 0.001000. What's going on here is an HFT stepping in front of the
NBBO to foist some 'price improvement' on you in the smallest possible amount to legally justify their line
cutting since they are acting with the knowledge of how the order book will look in the immediate future
and how to profit from it (with order flow information they paid for *cough cough*) and the cherry on top is
preventing the retail investor from moving the NBBO since the order was chopped up by the broker into
odd lots to make them trade through exempt (I've seen this happen with my own round and mixed lot
orders lol).

There's a lot more to read on subpenny trades by this great research firm, Nanex:

- [What Every Retail Investor Needs to
Know](http://www.nanex.net/aqck2/What-Every-Retail-Investor-Needs-to-Know.pdf)
- [Sub-penny for your thought](http://www.nanex.net/aqck2/3540.html)
- [Where is the Sub-penny Competition?](http://www.nanex.net/aqck2/3533.html)
- [Sub-penny Price Anomaly](http://www.nanex.net/aqck2/3517.html)
- [Salami Slicing Sub-Penny Style](http://www.nanex.net/aqck2/3519.html)
- [Can you jump in front of the Best Bid or Offer by 1/100th of a
penny?](http://www.nanex.net/aqck2/4618.html)
- [Subpenny Flash Crash](http://www.nanex.net/aqck2/3544.html)

So what I've done to measure subpenny volume over time is:

1. Download all trades from the tape from 2020-01-01 to 2021-09-30 for the given symbols: AMBA, SFIX,
CLOV, BBBY, KOSS, AMC, GME and filter them according to:
- within normal market hours (no pre/post market volume)
- with an error code of 0 (no error) or 1 (corrected in place)
2. Bucket all volume based on the subpenny fraction in the trade price
- prices without a subpenny fraction are bucketed into '0000'
- prices at midpoint (ex. $12.345000) are bucketed into '5000'
- all other prices are bucketed by their subpenny fraction into ranges of 1000, that is, a trade with volume
45 at price $185.251234 and one with volume 105 at $209.461445 would be grouped together to
contribute 150 volume to the bucket covering $X.XX1000-$X.XX1999.
3. Filter all volume by trade condition of interest
- Trade conditions: (not all are explicitly marked on tape, some are derived from context of trade like '100
Lot')
- 100 Lot: Trade had exactly 100 volume, just enough to affect NBBO

- ISO: Intermarket Sweep Order
- Odd Lot: Trade had less than 100 volume
- Regular Sale: Trade _without_ a trade condition on the tape
- Round Lot: Trade had volume in multiple of 100
- TTE: Trade was marked Trade Through Exempt for any reason
4. Divide the volume of all trades having the condition of interest occurring at a price within a given subpenny bucket by the daily volume of all trades having the condition of interest. This gives the distribution of volume of all trades across subpenny buckets relative to all volume of the condition of interest.

Below are albums of candlesticks and heatmaps showing the distribution of daily volume at the subpenny level by trade condition. The candlesticks are fairly straightforward to read but the heatmaps are information dense; recall they display the percentage of volume transacting at prices having a subpenny fraction residing in a bucket covering a range of 1000. **Volume with no subpenny fraction (bucket 0000) and the midpoint fraction (bucket 5000) are omitted from the heatmaps because the vast majority of volume transacts at those fractions and would make the heatmap more difficult to read. That volume is """clean""" because it is more transparently and fairly priced for retail (IEX transacts at midpoint for example). We're looking at patterns in the """dirty""" volume.** The only exception is 2 heatmaps for KOSS which would be nearly empty _without_ those buckets.

- [AMBA](https://imgur.com/a/qZp0fpN) (random midcap)
- [SFIX](https://imgur.com/a/8xIvtgU) (random midcap)
- [CLOV](https://imgur.com/a/1blEJ1m) (random shorted? midcap)
- [BBBY](https://imgur.com/a/eo0puk2)
- [KOSS](https://imgur.com/a/J9r24Mv)
- [AMC](https://imgur.com/a/icdsRaI)
- [GME](https://imgur.com/a/LFY56kQ)

It is very interesting that the distributions in the meme stocks change very dramatically at the end of January but they appeared to have different patterns in the lead up.