Like many people around the world I have been watching the rise and fall of share prices in the US retail chain GameStop with interest. There is plenty of narrative and interpretation of what happened around and time will provide many details that are currently opaque, but here is my brief summary of what happened. This is based purely on what I (as a non-specialist in the field but a sceptical numbers professional) have gleaned from what has appeared in various media:

- GameStop has had a few non-profitable years, hit hard by the secular decline in bricks-and-mortar retail (in the face of online competitors such as Steam) and by Covid-19. This has led to a fairly low share price and a lot of short-selling by investors expecting it to go even lower, possibly to bankruptcy.
- In 2020 there was a small groundswell of investors who thought the price had gone too far down and the firm still has growth potential, most prominently successful e-commerce entrepreneur Ryan Cohen; and insurance marketer Keith Gill (who had, unknown to his employer, a YouTube channel under the name of Roaring Kitty and a Reddit identity as DeepFuckingValue, through both of which he expounded his faith in the long term value of GameStop).
- A number of market manipulators organising on Reddit identified the opportunity of heavily short-sold stock that might actually be slightly undervalued. Through co-ordinating the bulk purchase of options to purchase stock in future they engineered in early 2021 a gamma squeeze (many purchases of options at high future prices leads to market makers covering their liabilities by purchasing stock, which leads to more increases in price) and a short squeeze of at least one prominent hedge fund that had short-sold the stock.
- Facing this squeeze or squeezes, the price of the stock quickly jumped to $100 and beyond – well beyond its fundamentals-based value, attracted lots of attention including in non-specialist media, and mischievous (my interpretation) amplification by Elon Musk. This lead to an old-fashioned speculative bubble featuring many new and naive investors who expected the squeeze combination to take the price "to the moon" – $10,000 or more (for a stock probably worth $5-$20, based on profitability and revenue).
- That bubble is in the process of collapsing as we speak, and eventually (although no-one knows when, or via what other resurgences and deflations on the way) will get back down to the stock's actual fundamentals-based value, which is probably more than it was priced at in 2020 but less than a twentieth of what people were paying for it its peak.
- In general (exceptions will apply) hedge funds and the like who thrive on volatility doubtless made large amounts of money both on the way up and on the way down, at the expense of more naive investors.

There are a bunch of side-issues that attracted attention too, including dubious claims that many of the predatory traders were involved for LOLs or political reasons rather than to make money; the ability of the Robinhood app to support its customers during high volatility; and the different types of traders on different sides of the short and long ends of the market. These are in my opinion red herrings. I quite like the explanation of those red herrings and how they led to some odd mistakes by various commentators (including from the left who should have known better) in this New York Intelligencer piece by Eric Levitz.

None of this is financial advice to short GameStop further (you'd have to be brave!); I am not a financial professional; I don't hold any assets relevant to this discussion including but not limited to a short or long position in GameStop; but as they (don't) say on Reddit, "I just don't like the stock".

Anyway, following this saga as it unfolded prompted me to think about how best to present at-a-glance stock prices and volumes, the two most widely available relevant metrics in cases like

this (a third critical metric, the short interest in the stock, is less available and I will leave for consideration at another time).

Here is the most common way that daily stock prices and volumes are presented. It is a two panel chart, with the top panel showing daily prices and the bottom daily volume. The prices are shown by "OHLC" bars, which acronym stands for open-high-low-close. In this particular version, the prices at the opening and closing of trading on the day are shown by a solid bar, and the high and low prices (when these go beyond the open and close prices, as they usually do) by extending vertical lines. I have opted to use a logarithmic scale for the vertical axis, which emphasises relative rather than absolute growth in price day to day. The volume panel at the bottom is a straightforward bar chart. The colour aesthetic for both the price and volume bars is mapped to the direction of price from closing to opening – red means prices went down over the day, blue means up.

This is actually a pretty efficient and effective visualisation of a five dimensional time series. The fact that it is the standard means that people who follow these stock prices and volumes get very accustomed to reading them. This means they can add all sorts of cluttering annotations that would usually be thought of as too complex if it weren't for the familiarity of the underlying chart.

Here is the code to download that GameStop (GME) stock price and volume data from Yahoo Finance and present the chart. Both the download and chart visualisation functionality come from Ryan and Ulrich's quantmod R package.

Post continues below R code

```
#----------------Functionality------------------
library(quantmod)
library(tidyverse)
library(scales)
library(Cairo)
library(glue)
library(lubridate)

#------------Download stock data--------
symb <- "GME"
symb_name = "The GameStop stock bubble of 2021 ($GME)"

this_stock <- getSymbols(symb, src = 'yahoo', auto.assign =
FALSE)

# will use just the last 100 values for illustration
d <- tail(this_stock, 100)
names(d)  <- gsub(glue("{symb}\\."), "", names(d))

annotation_lab <- "Highest plausible actual value based on
firm's revenue, etc.: $20"
annotation_col <- "steelblue"


#---------OHLC Chart plus volume bars-----------
chartSeries(d, theme = chartTheme("white",
                                  bg.col = 'grey96',
```

```
                                        fg.col = "grey70",
                                        grid.col = "white",
                                        border = "transparent"),
                up.col = 'darkblue', dn.col = 'tomato',
                log.scale = TRUE,
                name = symb_name)

    # 32 is the number below chosen to appear as 20 on the y axis
    of top panel, and
    # it totally depends on device size, scale, etc. Caution -
    set by hand.
    # Basically like drawing a line on by hand.
    abline(32, 0, col = annotation_col)
    text(100, 40, annotation_lab, col = annotation_col, cex =
    0.8)
```

What are some of the alternatives? In other contexts (ie non-financial data) for general users, it wouldn't have occurred to me to use an OHLC chart like the traders use. I would have chosen just one of the four prices (probably the closing price) and made a line chart out of it; and if I wanted to show the volume as well would have mapped volume to size of points as a layer over the line chart. This gets me this result:

It's quite nice. You can see the volume of sales in a qualitative sense, and the up and down of the price is clear - probably clearer than in the more cluttered OHLC chart. Both charts show (for example) the burst of trades and price jump in October 2020 and then in January 2021; both give a good sense of the massive rise and fall in price in 2021 with big volumes on the way up and down; both make the peak in price pretty obvious. The OHLC chart makes the peak in volume clearer (because it's easier for the eye to measure distance vertically than size of points); and the colour is a nice touch too.

If I wanted to show what was happening to an audience that was numerate (and hence was familiar with notions such as the logarithmic vertical scale, and mapping size of points to a variable) but not familiar with finance visualisations in particular, this would probably be my starting point.

Here's the code to make that chart. I'm now using Wickham's `ggplot2` for the visualisations.

```
    #----------------Line series with bubble points-------------
    d %>%
      as_tibble() %>%
      mutate(date = time(d)) %>%
      ggplot(aes(x = date, y = Close)) +
      geom_hline(yintercept = 20, colour = "steelblue") +
      geom_line() +
      geom_point(aes(size = Volume / 1e6), alpha = 0.5) +
      annotate("text", x = Sys.Date() - 50, y = 32,  size = 3,
    hjust = 1,
             colour = annotation_col,
             label = annotation_lab) +
      scale_y_log10(label = dollar_format(accuracy = 1)) +
      scale_x_date(date_labels = "%b %Y") +
      scale_size_area(label = comma) +
      labs(x = "",
```

```
        y = "Closing price (log scale)",
        size = "Volume (m)",
        caption = "Source: Yahoo Finance, via quantmod R
package",
        title = symb_name,
        subtitle = "Line plot for price with volume shown as
point size. Intuitive, but not great for showing
relationship.")
```

The weakness of both the line chart and the OHLC chart is that the relationship between volume and price can be difficult to parse. The best way to show two continuous variables' relationship is nearly always some kind of scatter plot, and when the two variables are moving over time it usually means one should use a connected scatter plot. Here's my version of one of them, again drawn with `ggplot2`:

I've nicked the idea of colouring each day's trading as "bearish" or "bullish" from the standard OHLC chart; in my case mapped to the colour aesthetic for the point and text layers while leaving the path connecting points grey for readability.

I think this is a great visualisation of the relationship and one I'd definitely use if I were going to model the relationship between price and volume, but I'd never use it in a stand-alone way. Instead, it's a complement for the much more familiar and intuitive representation in one of the two preceding charts. Some variant of a time series chart with date/time on the horizontal axis should be used to establish the overall pattern over time before we look at the more complex bivariate relationship.

Here's the code for that final chart. The advantage of the `ggplot2` paradigm is that constructing this chart used many of the same notions and code snippets as the time series one, a massive saver in development time.

```
#---------------Connected Scatter Plot-------
d %>%
  as_tibble() %>%
  mutate(date = time(d)) %>%
  group_by(month(date), year(date)) %>%
  mutate(direction = ifelse(Close > Open, "Bullish",
"Bearish"),
         lab = ifelse(day(date) == min(day(date)) | Volume >
100e6 | Close > 300,
                      format(date, "%e %b %y"), "")) %>%
  ungroup() %>%
  ggplot(aes(x = Volume / 1e6, y = Close, label = lab)) +
  geom_path(colour = "grey") +
  geom_point(aes(colour = direction), alpha = 0.5) +
  geom_text(aes(colour = direction), hjust = 0, size = 3,
nudge_x = 0.015) +
  scale_x_log10(label = comma_format(suffix = "m", accuracy =
1),
                limits = c(1, 300)) +
  scale_y_log10(label = dollar_format(accuracy = 1)) +
  scale_colour_manual(values = c(Bullish = "darkblue",
Bearish = "tomato")) +
  labs(x = "Volume (log scale)",
```

```
        y = "Closing price (log scale)",
          title = symb_name,
          colour = "Direction of sentiment over the day:",
          subtitle = "Connected scatter plot. Not intuitive, but
   shows relationship well.")
```

Well, that's all for now. I think all three of these visualisations are good ones for different purposes and audiences. I'm impressed in particular with the efficient way the OHLC/volume chart shows a five dimensional time series in quite an intuitive way. But the other two have their uses too.