# Mean-Variance Portfolio Optimization: G-MAFIA BAT Q1 2020 - Q2 2021

Prepared by Justin Farris

# G-MAFIA BAT. What's in a name?

G-MAFIA BAT is an acronym for what are some of the world's most powerful companies: Google, Microsoft, Amazon, Facebook, IBM, Apple, Baidu, Alibaba, and Tencent.

Over the course of the COVID-19 pandemic, more than one received attention for their growing power and profits.

As an investment strategy, do G-MAFIA BAT portfolios yield promising returns? What does correlation between these assets look like, whether positive or negative? Just how much did Amazon make over the past year and a half?

I captured both exploratory data analysis and mean-variance optimization using Markowitz efficient frontier in a Jupyter notebook with LaTeX and Markdown.

Let's dive in!
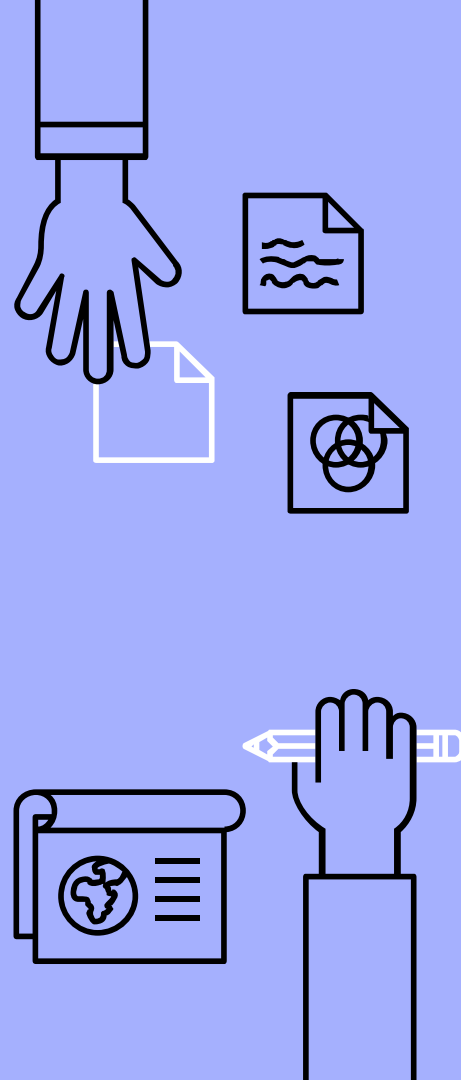
# Company profiles

G-MAFIA BAT companies can be considered technology companies. Though many are hypernational, Baidu, Alibaba, and Tencent primarily serve Asian markets, and G-MAFIA grew up serving American markets.

Several companies are in direct competition with each other, like Google and Apple. Many more are in less direct competition, like Amazon and Facebook.

All operate in similar or connected ecosystems, and all have enormous amounts of capital.
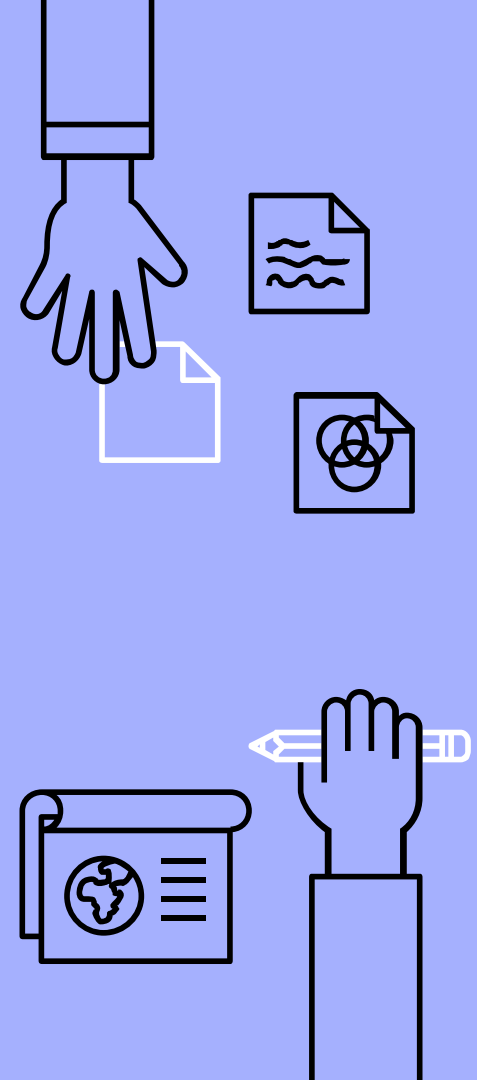
# Adjusted closing prices compared over Q1 2020 – Q2 2021

- **Closing prices adjusted for corporate actions such as stock splits and dividends**
- **GOOG and AMZN posted the largest gains over this period**
- **This graph provides context for the discussion around AMZN's controversial pandemic performance**



Adjusted Closing Prices over Time
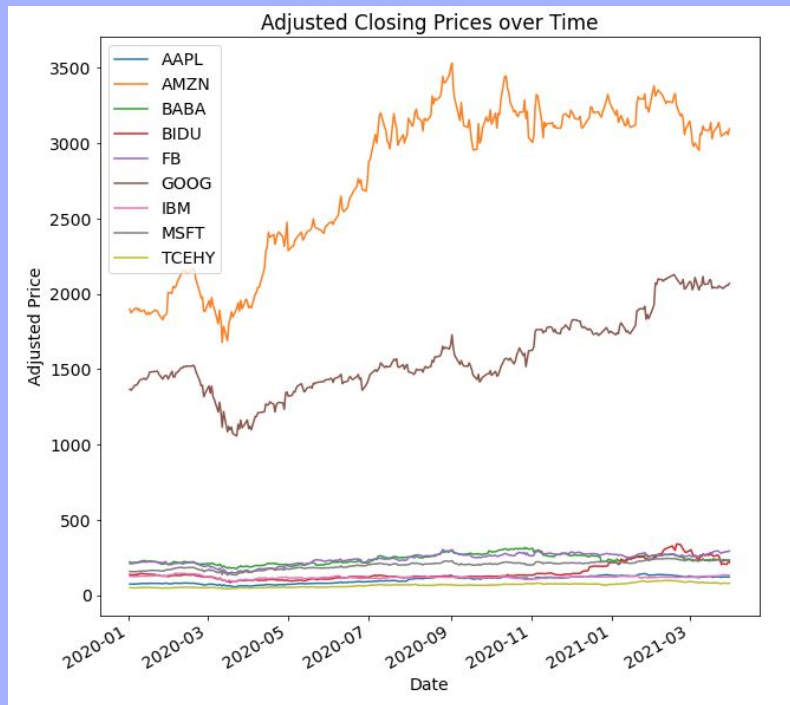
# Individualized simple returns over time

- **Note the recent dip in performance compared to 2020**
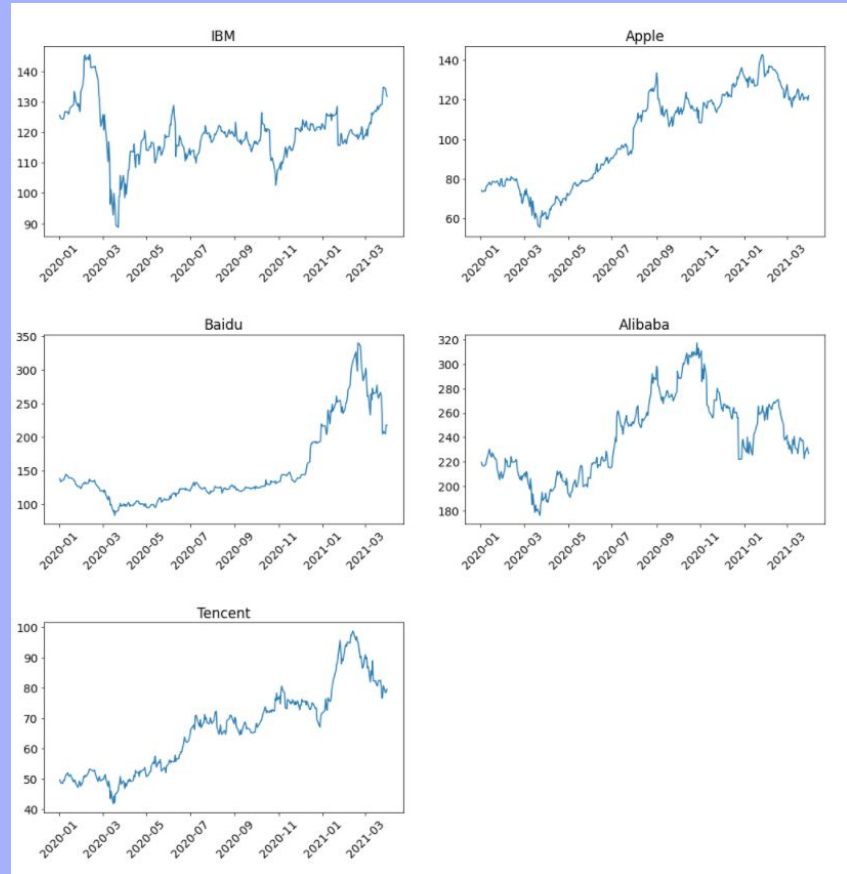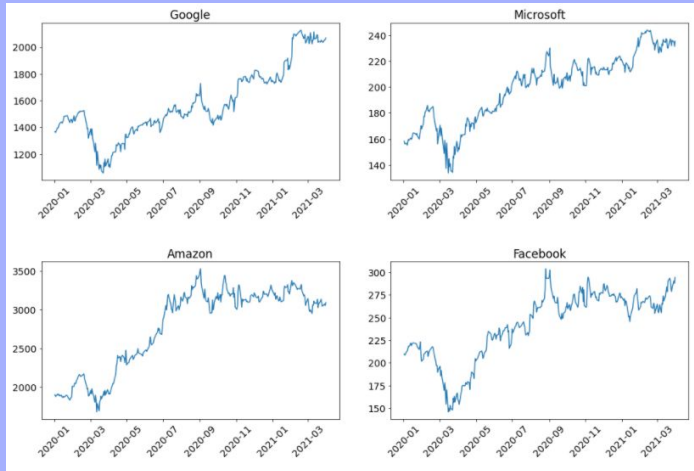
# Simple returns over time reframed as percentages

- **A cursory glance (taking y-axis scale into account) suggests IBM and Tencent have experienced the least day-to-day variance**
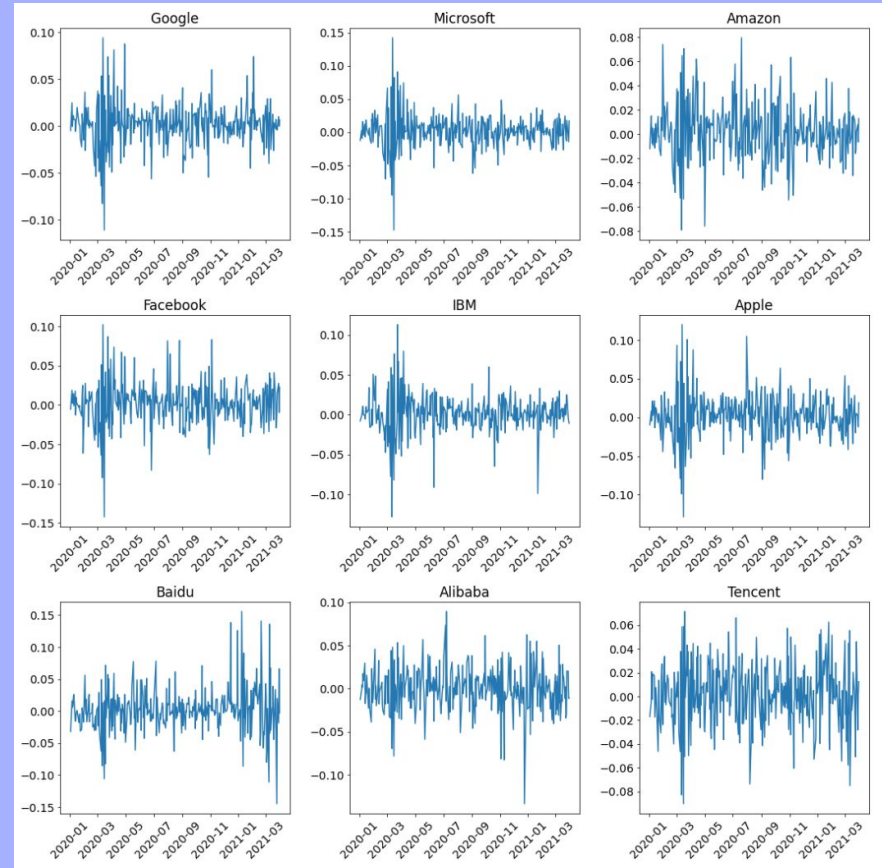
# Comparative summary of simple returns over time as percentages

- **This graph shows most of the extreme spikes in both directions coming from Baidu, as well as some from IBM and Google**



Daily Simple Rate of Return over Time

# Summary statistics broken out by company

- **Baidu has enjoyed the most gains over the selected time period, but also suffered the most variance**
- **Variance indicates fluctuation in investment returns**
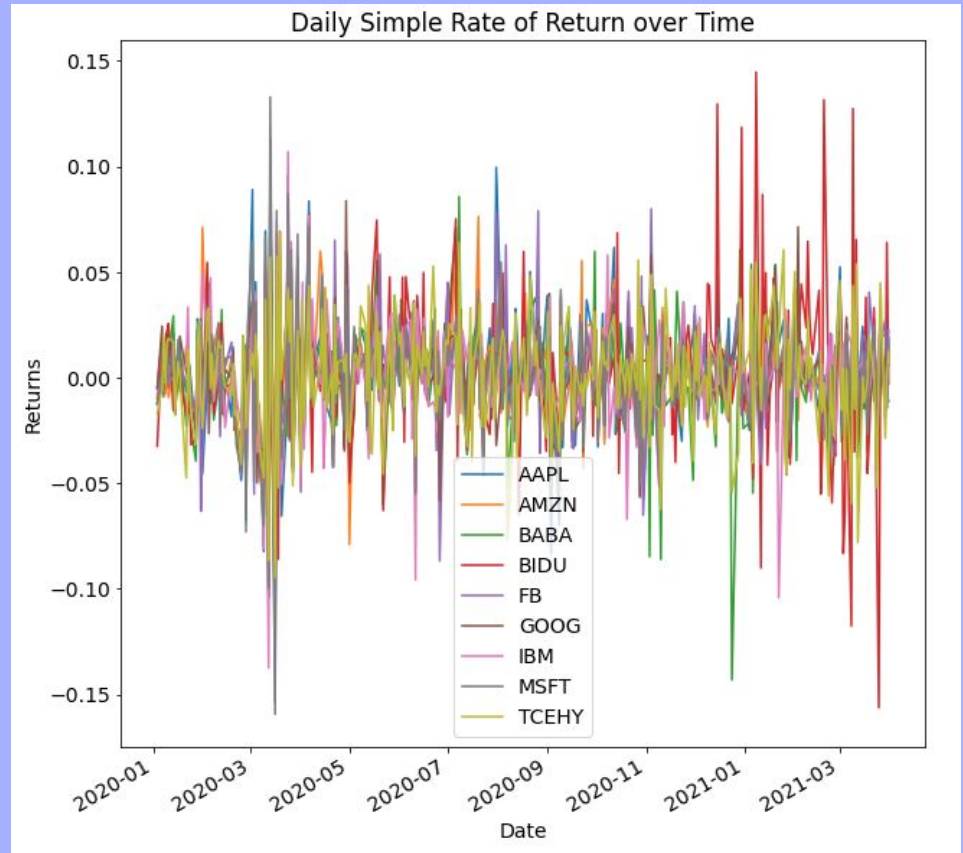- **Standard deviation transforms variance into human-understandable units**



Mean daily return per company Jan 1, 2020 - Apr 1, 2021

$$s^2 = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N - 1}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n - 1}}$$



Variance per company Jan 1, 2020 - Apr 1, 2021
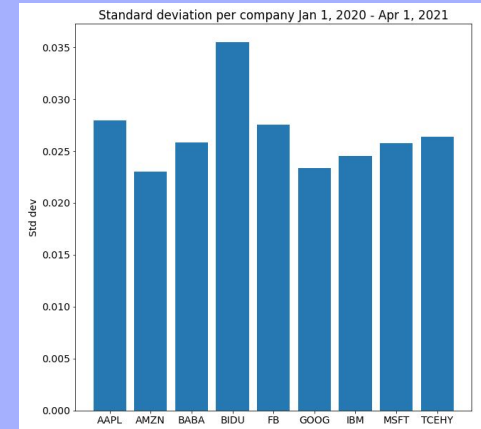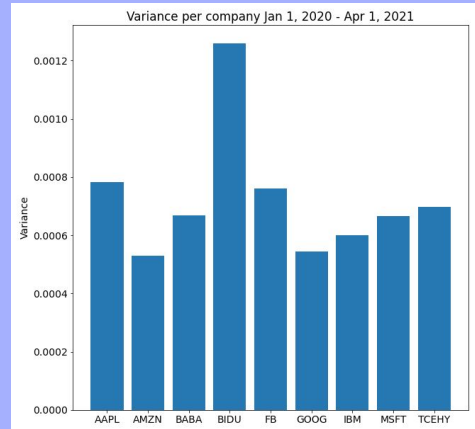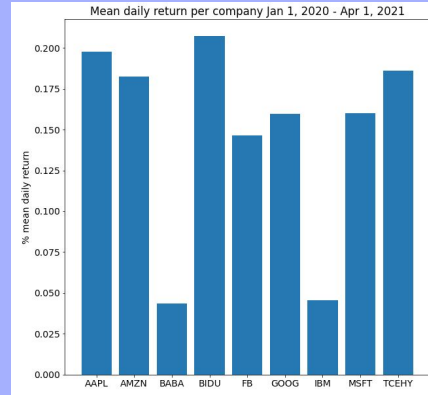


Standard deviation per company Jan 1, 2020 - Apr 1, 2021

# Covariance between assets

- Covariance communicates the directional relationship between the returns on two assets.
- Positive covariance indicates that assets move together (covariance value greater than 0).
- Negative covariance means they move inversely (covariance value less than zero). Uncorrelated assets have no quantifiable pattern, and their covariance value is 0.
- Generally, well-balanced portfolios reduce risk and volatility by pairing assets with negative covariance. If all stocks across a portfolio are positively covariant, poor performance in one will mean some measure of poor performance in the others, causing higher losses to the portfolio.
- Covariance is sometimes preferred to correlation for similarly-scaled variables, like those in our current portfolio.

# Covariance formula and heatmap

$$cov_{x,y} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{N-1}$$

- **From these results, we can see that the BAT companies are slightly less covariant with the G-MAFIA companies as a whole, which makes sense, since they operate in somewhat different markets. Another standout is that**
- **IBM appears to be the least covariant with any other stock**

Covariance Heatmap

| | AAPL | AMZN | BABA | BIDU | FB | GOOG | IBM | MSFT | TCEHY |
|---|---|---|---|---|---|---|---|---|---|
| AAPL | 0.00078 | 0.00045 | 0.00035 | 0.00041 | 0.00058 | 0.00047 | 0.00035 | 0.0006 | 0.00039 |
| AMZN | 0.00045 | 0.00053 | 0.0003 | 0.00029 | 0.00043 | 0.00036 | 0.00021 | 0.00044 | 0.0003 |
| BABA | 0.00035 | 0.0003 | 0.00067 | 0.00044 | 0.00034 | 0.00031 | 0.0002 | 0.00036 | 0.0005 |
| BIDU | 0.00041 | 0.00029 | 0.00044 | 0.0013 | 0.00036 | 0.00036 | 0.00025 | 0.00038 | 0.00048 |
| FB | 0.00058 | 0.00043 | 0.00034 | 0.00036 | 0.00076 | 0.0005 | 0.0003 | 0.00053 | 0.00035 |
| GOOG | 0.00047 | 0.00036 | 0.00031 | 0.00036 | 0.0005 | 0.00054 | 0.00033 | 0.0005 | 0.00031 |
| IBM | 0.00035 | 0.00021 | 0.0002 | 0.00025 | 0.0003 | 0.00033 | 0.0006 | 0.00036 | 0.00019 |
| MSFT | 0.0006 | 0.00044 | 0.00036 | 0.00038 | 0.00053 | 0.0005 | 0.00036 | 0.00067 | 0.0004 |
| TCEHY | 0.00039 | 0.0003 | 0.0005 | 0.00048 | 0.00035 | 0.00031 | 0.00019 | 0.0004 | 0.0007 |

# Correlation formula and heatmap

- **Correlation is a similar concept to covariance, but the correlation coefficient is standardized, so that the values always fall between –1 and +1, making it more intuitive.**
- **We can see that each company is correlated to itself, which is logical. Strong correlations appear to be segregated by geographic region. Microsoft appears to have the largest correlations overall (0.83 with MSFT and GOOG, respectively).**
-

$$\rho_{X,Y} = \frac{cov_{X,Y}}{\sigma_X \sigma_Y}$$

**Correlation Heatmap**

| | AAPL | AMZN | BABA | BIDU | FB | GOOG | IBM | MSFT | TCEHY |
|---|---|---|---|---|---|---|---|---|---|
| **AAPL** | 1 | 0.7 | 0.49 | 0.42 | 0.76 | 0.72 | 0.51 | 0.83 | 0.53 |
| **AMZN** | 0.7 | 1 | 0.51 | 0.35 | 0.68 | 0.67 | 0.37 | 0.74 | 0.5 |
| **BABA** | 0.49 | 0.51 | 1 | 0.48 | 0.47 | 0.51 | 0.32 | 0.54 | 0.73 |
| **BIDU** | 0.42 | 0.35 | 0.48 | 1 | 0.37 | 0.43 | 0.28 | 0.42 | 0.51 |
| **FB** | 0.76 | 0.68 | 0.47 | 0.37 | 1 | 0.78 | 0.44 | 0.75 | 0.48 |
| **GOOG** | 0.72 | 0.67 | 0.51 | 0.43 | 0.78 | 1 | 0.57 | 0.83 | 0.51 |
| **IBM** | 0.51 | 0.37 | 0.32 | 0.28 | 0.44 | 0.57 | 1 | 0.58 | 0.3 |
| **MSFT** | 0.83 | 0.74 | 0.54 | 0.42 | 0.75 | 0.83 | 0.58 | 1 | 0.58 |
| **TCEHY** | 0.53 | 0.5 | 0.73 | 0.51 | 0.48 | 0.51 | 0.3 | 0.58 | 1 |
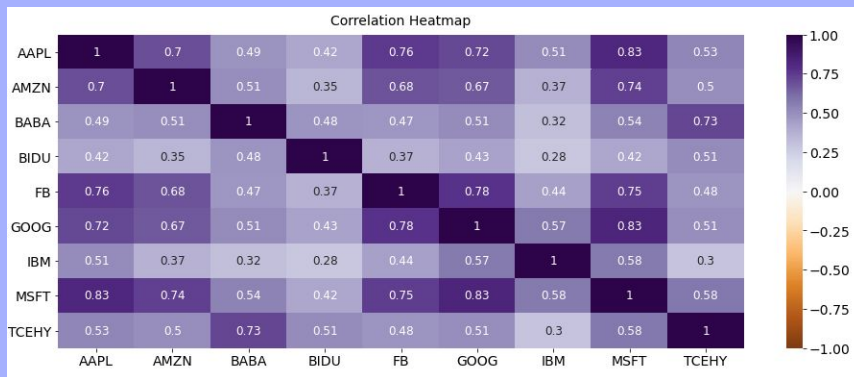
# Markowitz Efficient Frontier

- **Plotting the efficient frontier, we are able to see highest potential reward given an accepted level of volatility**
- **Plotted in red is each company's performance versus volatility. Alone, most do not reach the efficient frontier, but in composition, many of the plotted portfolios do**
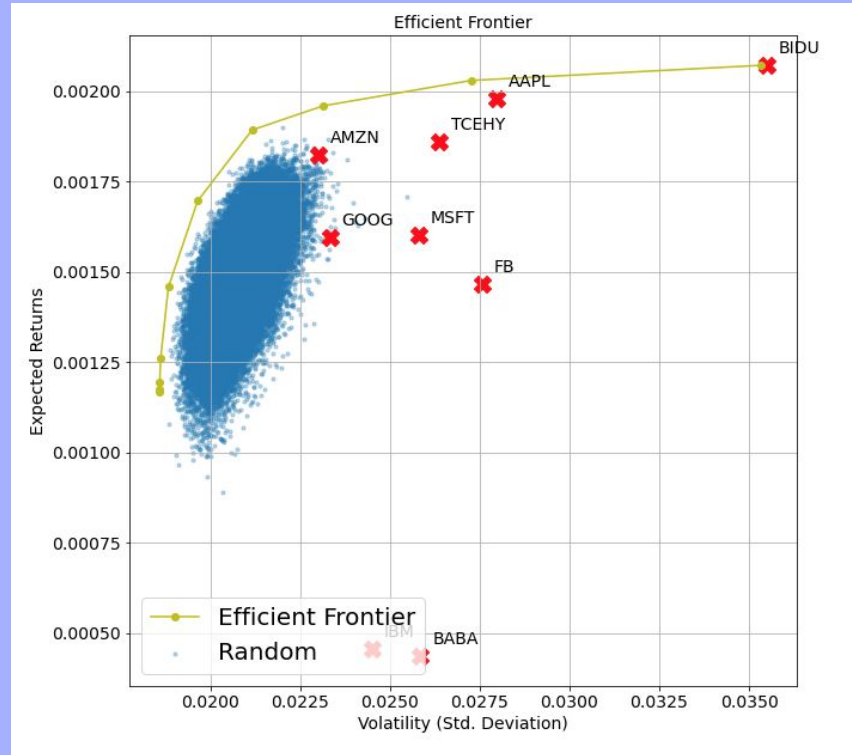
# Minimum volatility portfolio and optimal portfolio

$$S = \frac{R_p - R_f}{\sigma_p}$$

$R_p$ = return of portfolio

$R_f$ = risk-free rate
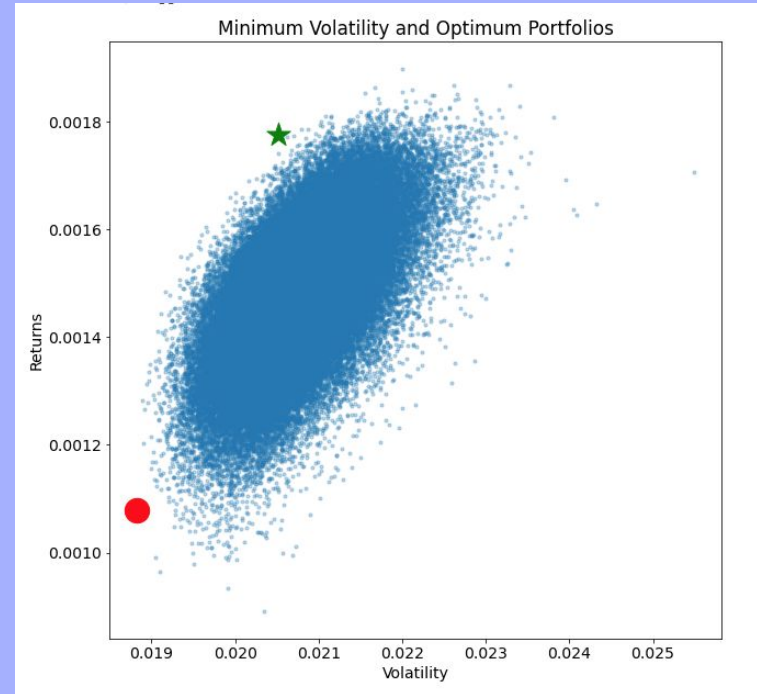
$\sigma_p$ = standard deviation of portfolio

- The minimum volatility portfolio is the same as the left-most portfolio on the scatterplot. Any portfolio to its right has increased risk.
- The optimal portfolio is found by determining which portfolio has the highest Sharpe ratio. The Sharpe ratio represents the additional amount of return that an investor receives per unit of increase in risk.



Minimum Volatility and Optimum Portfolios

# Conclusions

- **Investment decisions can be made according to multiple desires: minimum risk, highest Sharpe ratio, and highest return given specific risk criteria**
- **The minimum volatility portfolio is most heavily weighted toward IBM, while the optimal portfolio is most heavily weighted toward Google**
- **The return on the minimum volatility portfolio is quite low. Most investors want to trade some level of risk for higher potential reward to enjoy the rewards of stocks as an investment instrument.**
- **The volatility difference between the minimum volatility and optimal portfolios appears appears to be proportionally less than the associated increase in potential returns**
- **The final mean-variance optimization demonstrates that while G-MAFIA BAT are all individual powerhouses, a portfolio derived from only them doesn't provide a very extended efficient frontier. Options for investors seem to be less than a more balanced portfolio, where risk tolerance can be adjusted with more impact. Balancing could involve choosing less tech-centric companies and organizations not in direct competition with each other. Given the pandemic, the chosen time period may not be representative of performance in other circumstances, though all included stocks seem fairly reliable.**