Quantitative Finance Mini-Projects

Himanshu Raj

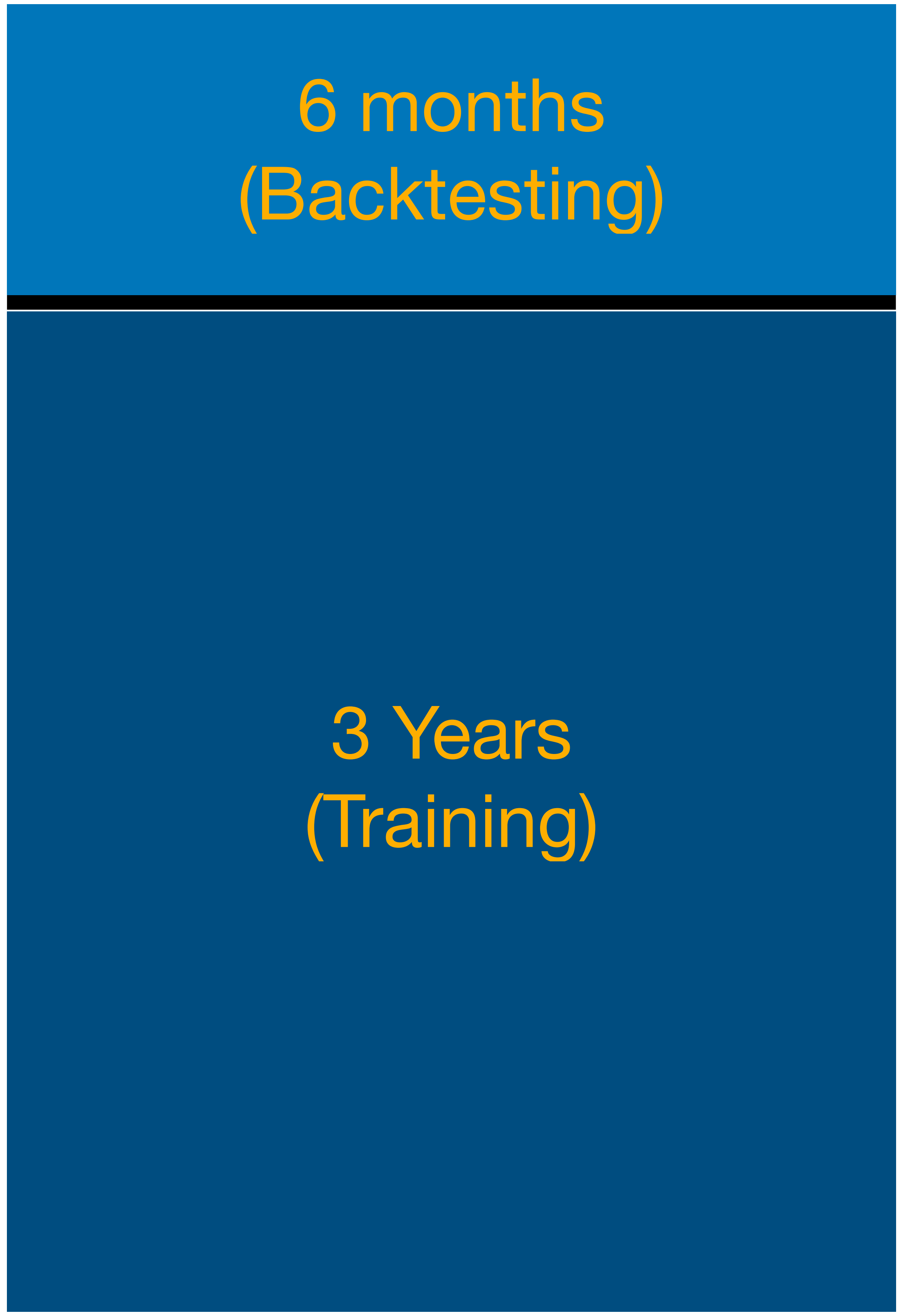
THE ERDŐS INSTITUTE

Helping PhDs get and create jobs they love at every stage of their career.

Project-1
Portfolio Risk

Using current stock data create two potentially profitable investment portfolios.
One that is higher risk and one that is lower risk.

Stock Dataset

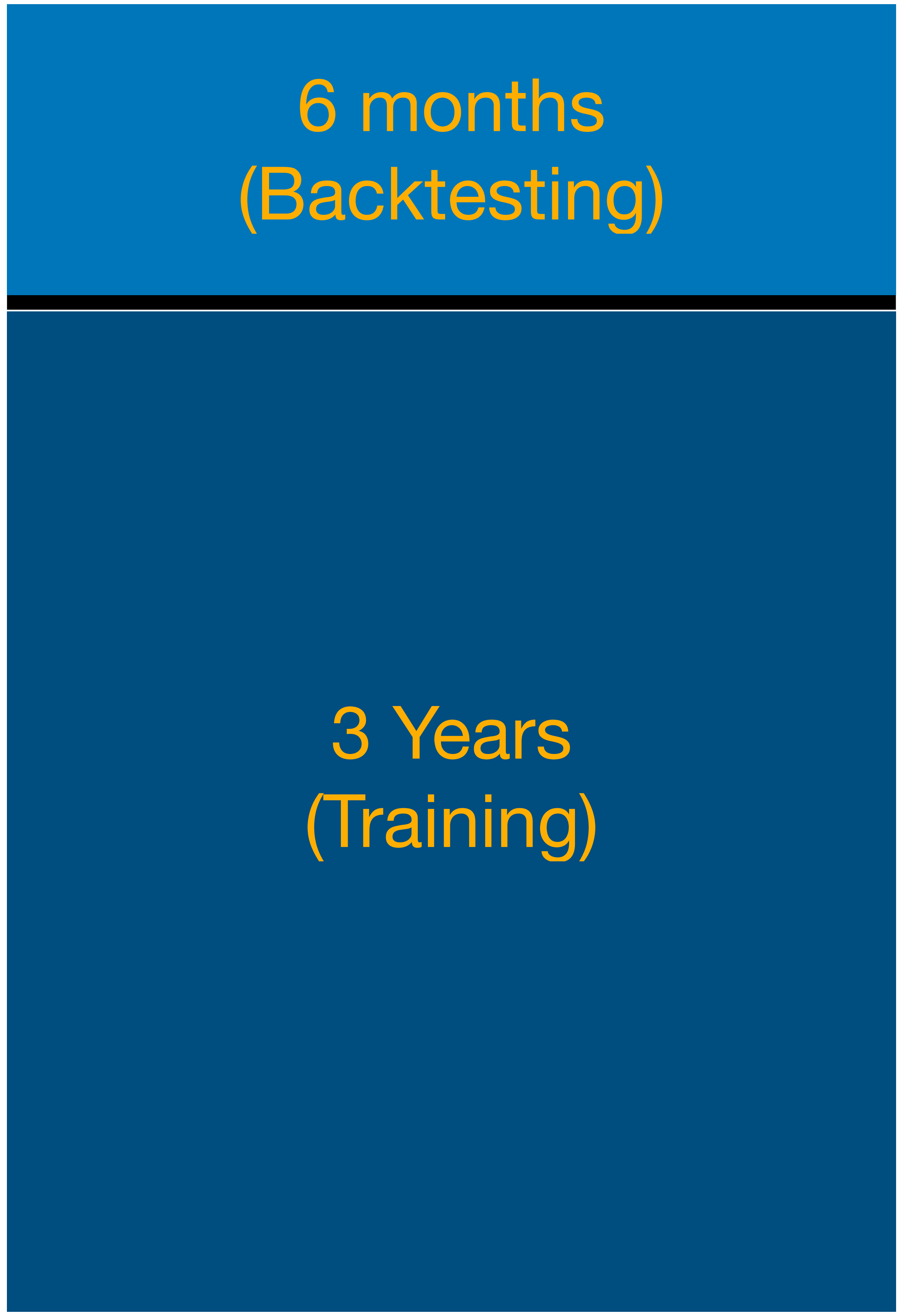


Two baskets of stocks

Basket 1 (Diversified by Sector)

```
diversified_tickers = [
    'AAPL', 'MSFT', 'GOOGL', 'AMZN', 'META',  # Tech
    'JPM', 'BAC', 'GS', 'MS', 'WFC',          # Finance
    'JNJ', 'PFE', 'MRK', 'UNH', 'ABBV',       # Healthcare
    'XOM', 'CVX', 'COP', 'SLB', 'EOG',        # Energy
    'PG', 'KO', 'PEP', 'WMT', 'COST'          # Consumer Staples
]
```

Stock Dataset



Two baskets of stocks

Basket 2 (Tech heavy)

```
tech_tickers = [
    'AAPL', 'MSFT', 'GOOGL', 'AMZN', 'META', 'NVDA', 'TSLA',  # Big Tech
    'ADBE', 'CRM', 'ORCL', 'INTC', 'AMD', 'AVGO', 'QCOM',
    'CSCO', 'IBM', 'TXN', 'AMAT', 'NOW', 'SHOP', 'UBER',
    'SNOW', 'ZM', 'DOCU', 'PYPL', 'NFLX'
]
```

Portfolio Construction: Optimisation

$r_t^i \equiv \frac{P_t^i}{P_{t-1}^i} \in \mathbb{R}^{T \times N}$: return series of the i' th stock

w_i : weight of the i' th stock in the portfolio $\sum_i w_i = 1$

$w_i \in [-1,1]$: $w_i > 0$ $(w_i < 0)$ long (short) position in the i'th stock

$R_t = r_t^i w_i$: return series of the portfolio

$\bar{R} = \mathbb{E}[r_t^i] w_i$: Expected portfolio return

$\mathbb{V}[R] = w_i C^{ij} w_j$: Portfolio variance

$C^{ij} = \mathbb{E}[(r^i - \bar{r}^i)(r^j - \bar{r}^j)]$: covariance matrix of the returns

Optimisation problem

Minimize: $-\mathbb{E}[R] + \lambda |w|^2$

- Risk tolerance: $\mathbb{V}[R] < \sigma^2$
- L2 regularization (to promote diversification)
- Long-only portfolios: $w_i \geq 0.01$
- Long-short portfolios: $w_i \in [-0.2, 0.2]$

Portfolio Construction: Types

We construct a total of 8 portfolios
Following four types in each basket:

1. Low-risk, long-only
2. Low-risk, long-short
3. High-risk, long-only
4. High-risk, long-short

Specifications:

Low risk: $\sqrt{\mathbb{V}[R]} < 0.08$

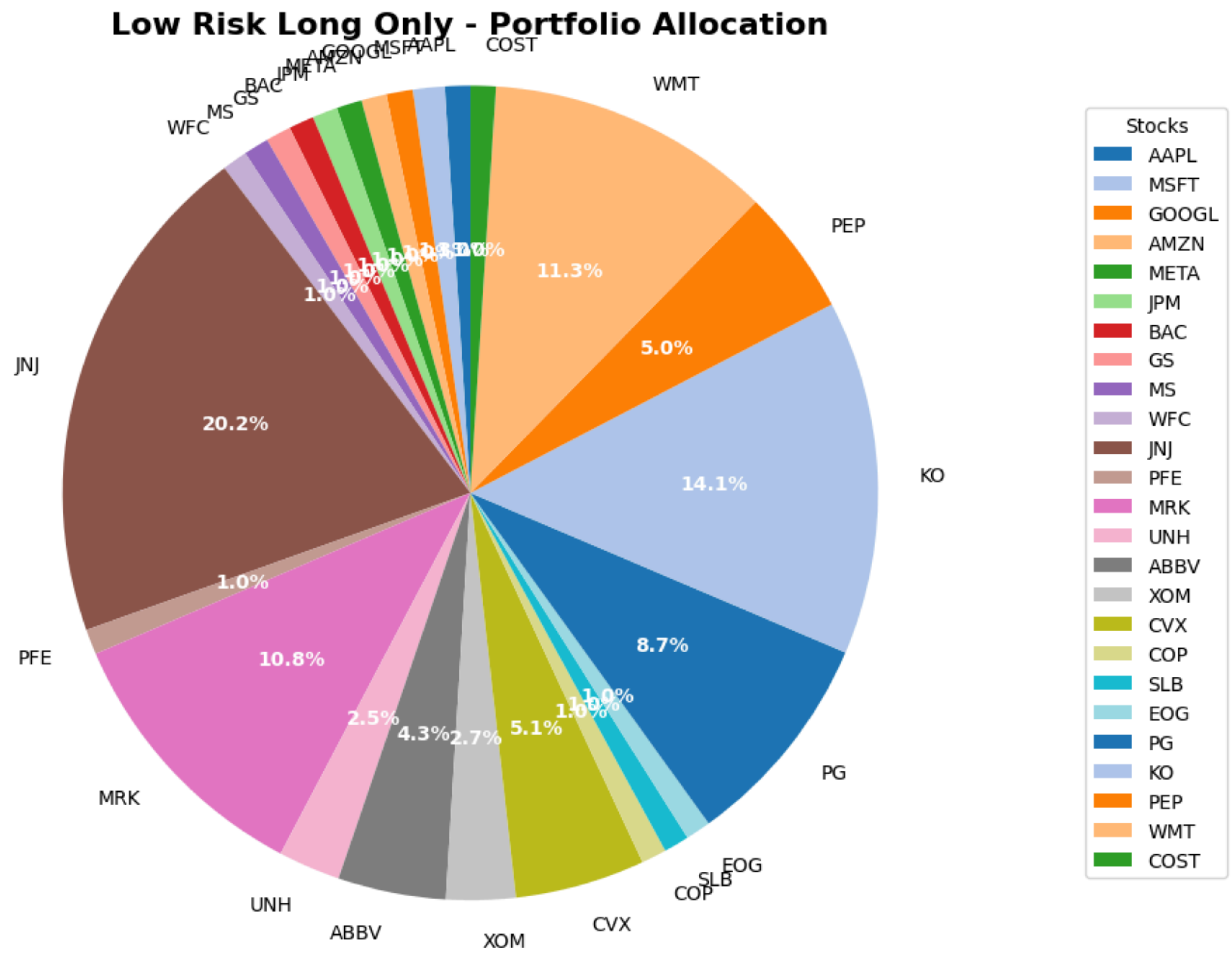
High risk: $\sqrt{\mathbb{V}[R]} < 0.25$

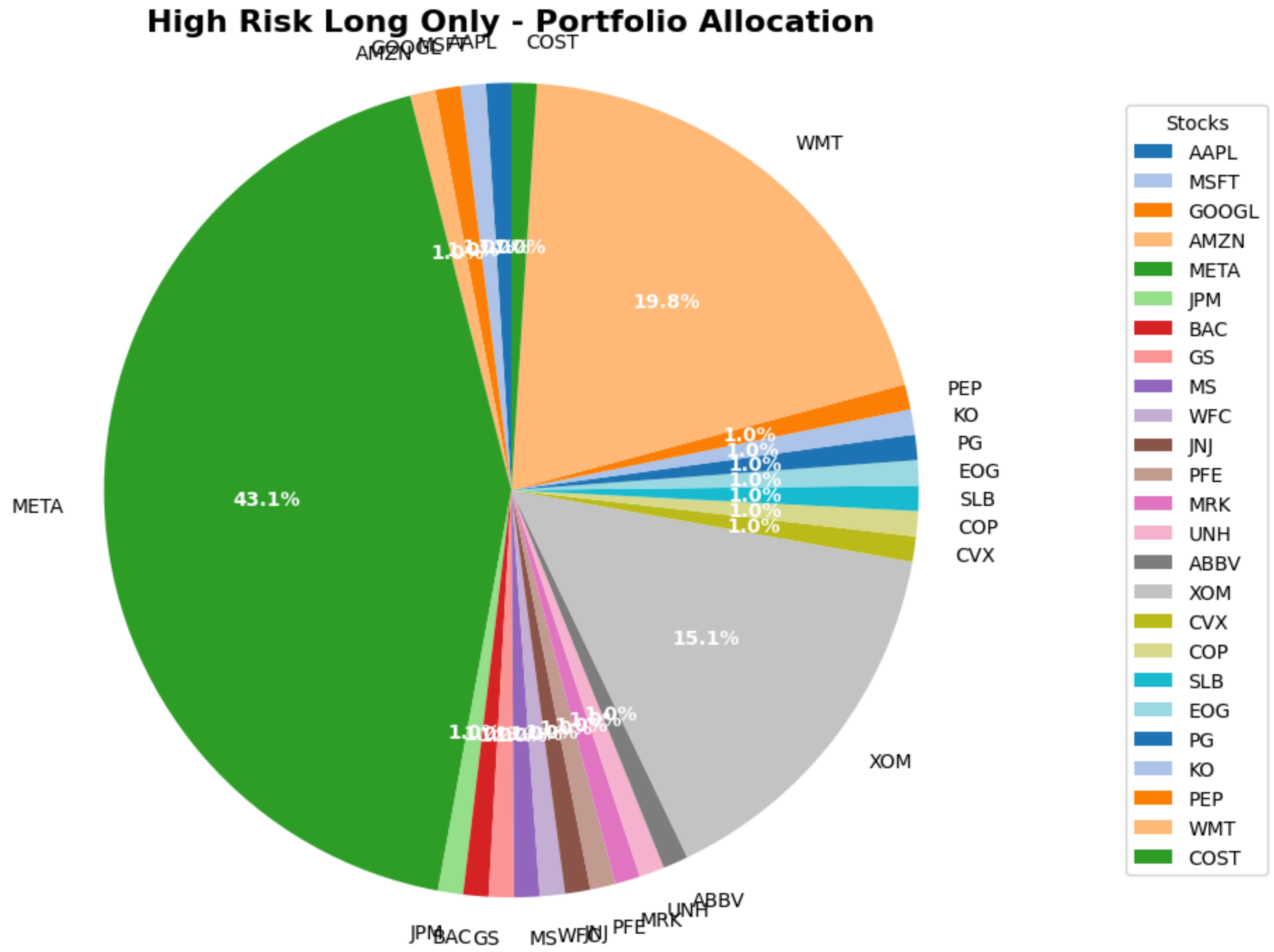
Long-only: $w_i > 0$

Long-short: w_i allowed to be negative but bounded below

Portfolio Construction: Allocation

Basket 1 (diverse sector): Long-only



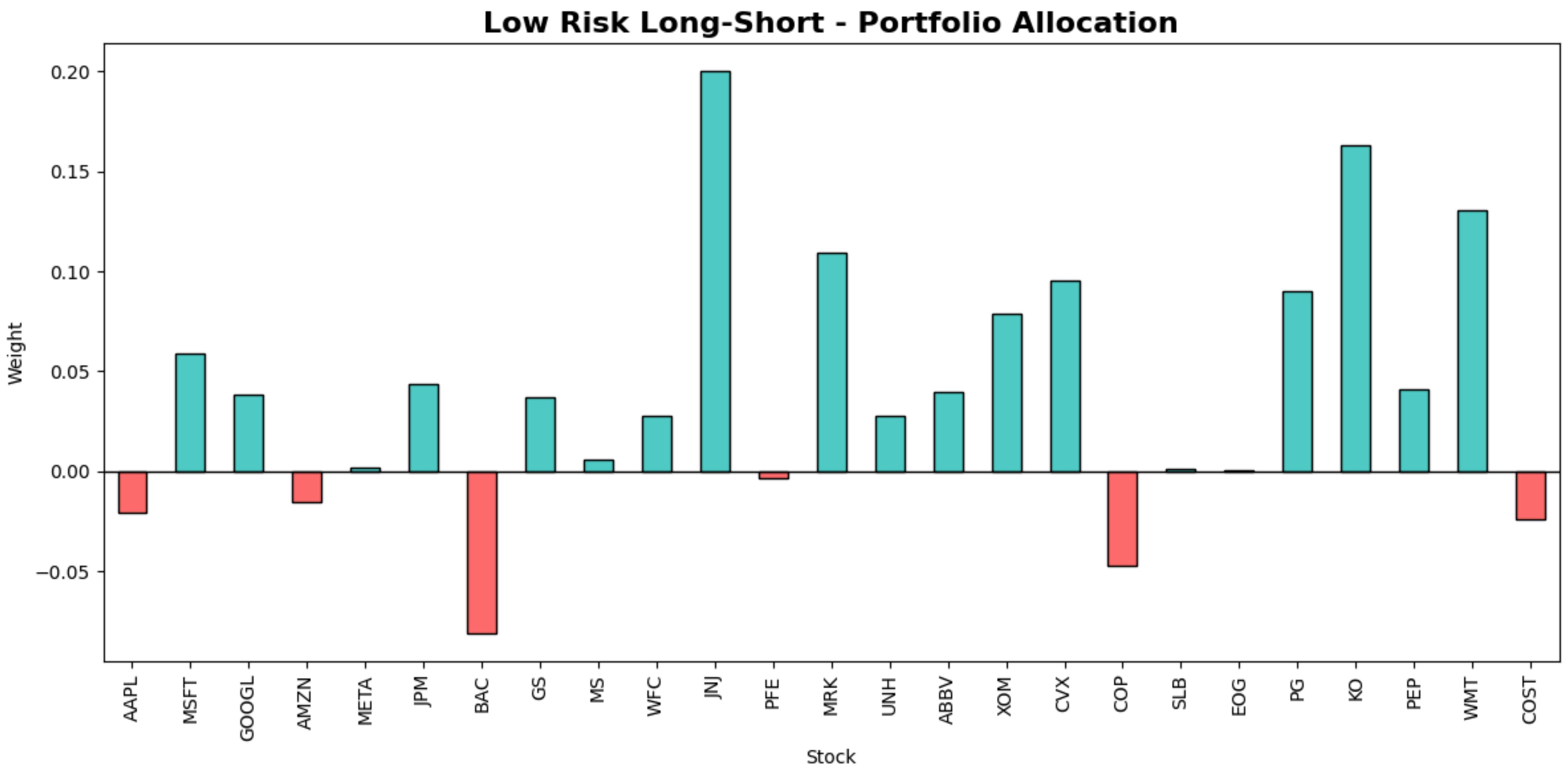


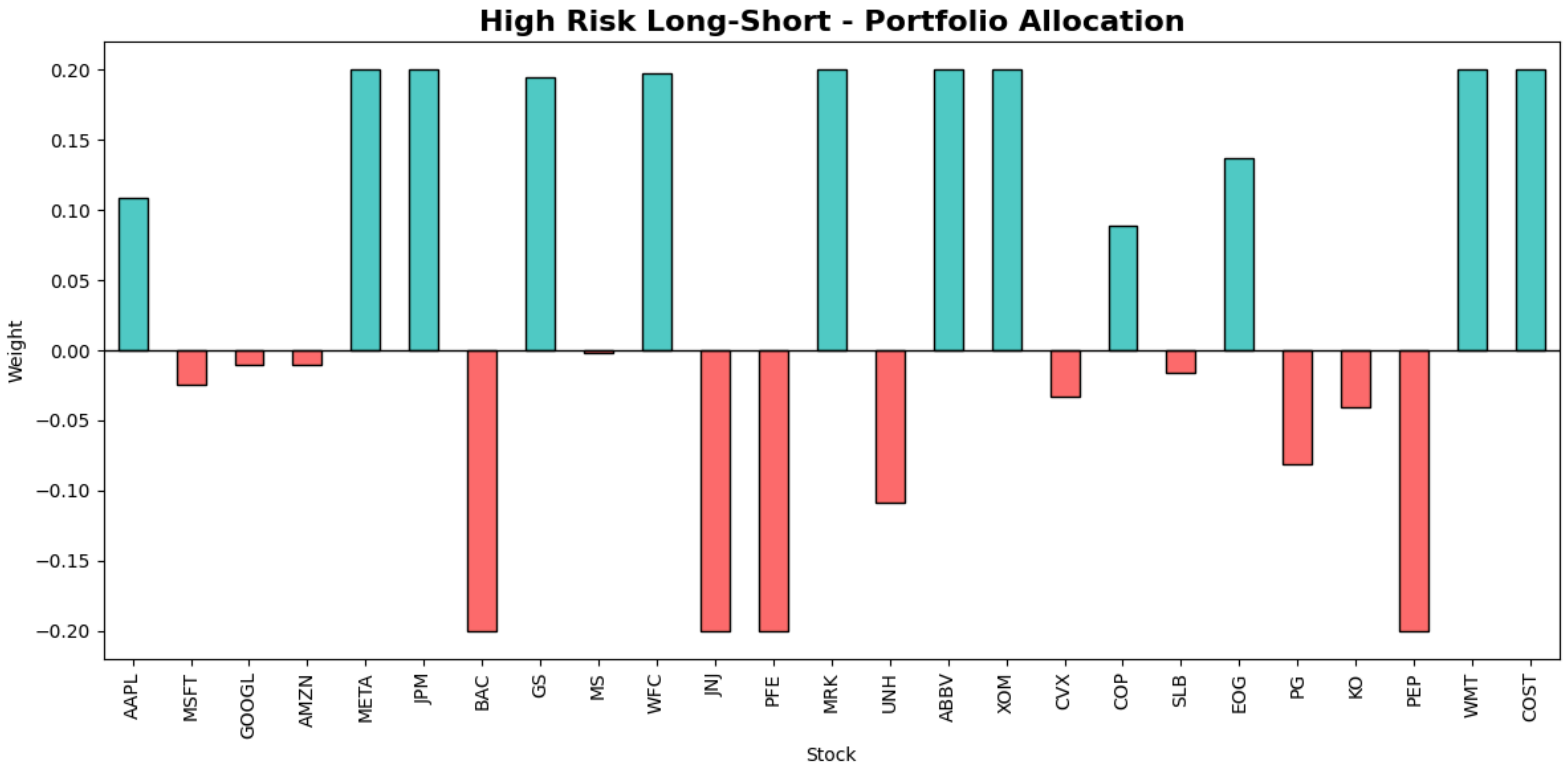

```
Annual Return Projections for Low Risk Long Only (Based on $100 investment):
Expected Annual Return: 10.30%
Annual Volatility: 11.90%
Sharpe Ratio: 0.66
Expected Value after 1 year: $110.30
95% Confidence Interval: $87.79 - $139.96
```

```
Annual Return Projections for High Risk Long Only (Based on $100 investment):
Expected Annual Return: 25.05%
Annual Volatility: 24.66%
Sharpe Ratio: 0.91
Expected Value after 1 year: $125.05
95% Confidence Interval: $79.23 - $208.30
```

Portfolio Construction: Allocation

Basket 1 (diverse sector): Long-short



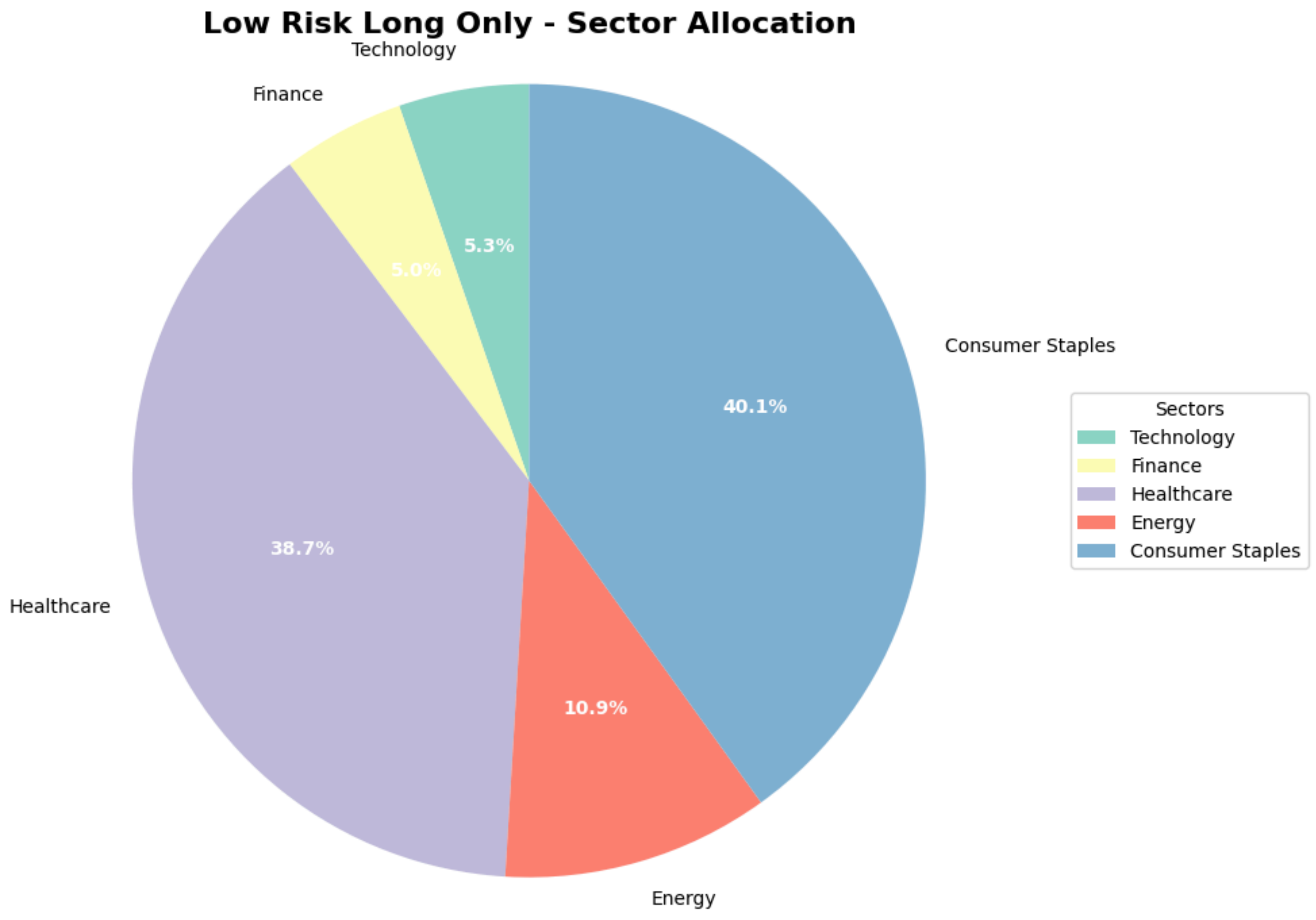


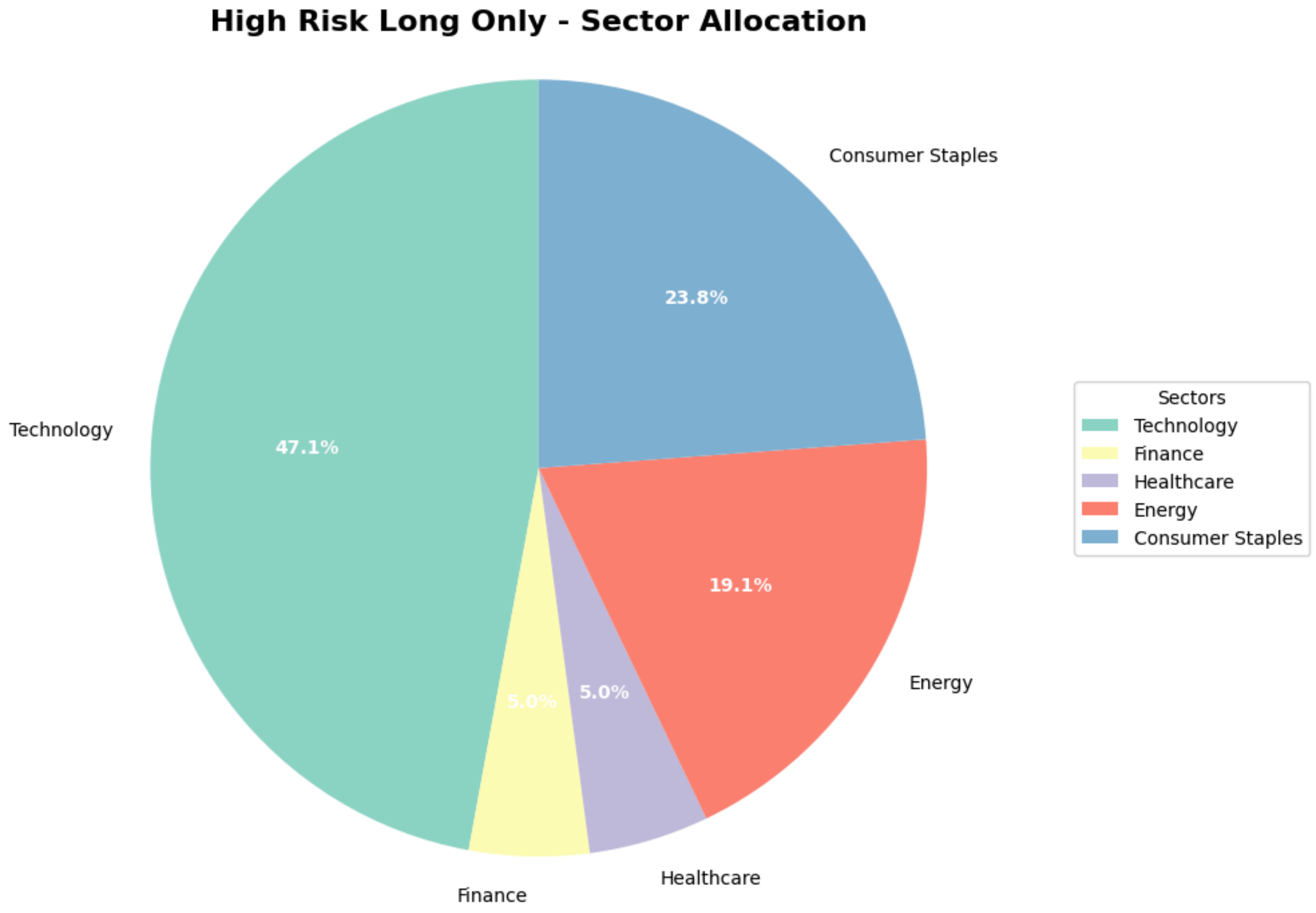

```
Annual Return Projections for Low Risk Long-Short (Based on $100 investment):
Expected Annual Return: 11.67%
Annual Volatility: 11.70%
Sharpe Ratio: 0.78
Expected Value after 1 year: $111.67
95% Confidence Interval: $89.34 - $141.35
```

```
Annual Return Projections for High Risk Long-Short (Based on $100 investment):
Expected Annual Return: 44.96%
Annual Volatility: 25.00%
Sharpe Ratio: 1.70
Expected Value after 1 year: $144.96
95% Confidence Interval: $96.04 - $255.89
```

Portfolio Construction: Allocation

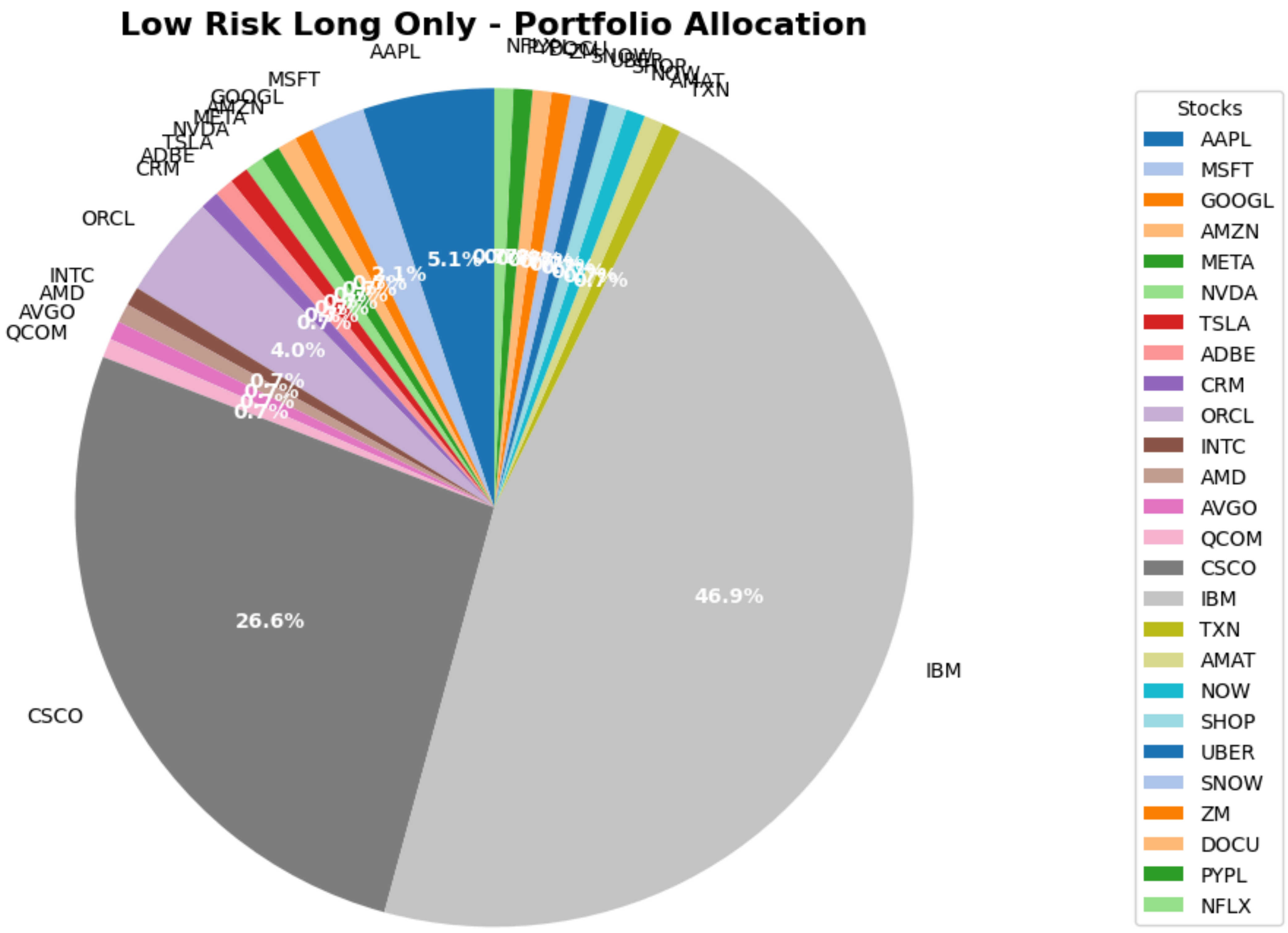
Basket 1: Allocation by sector



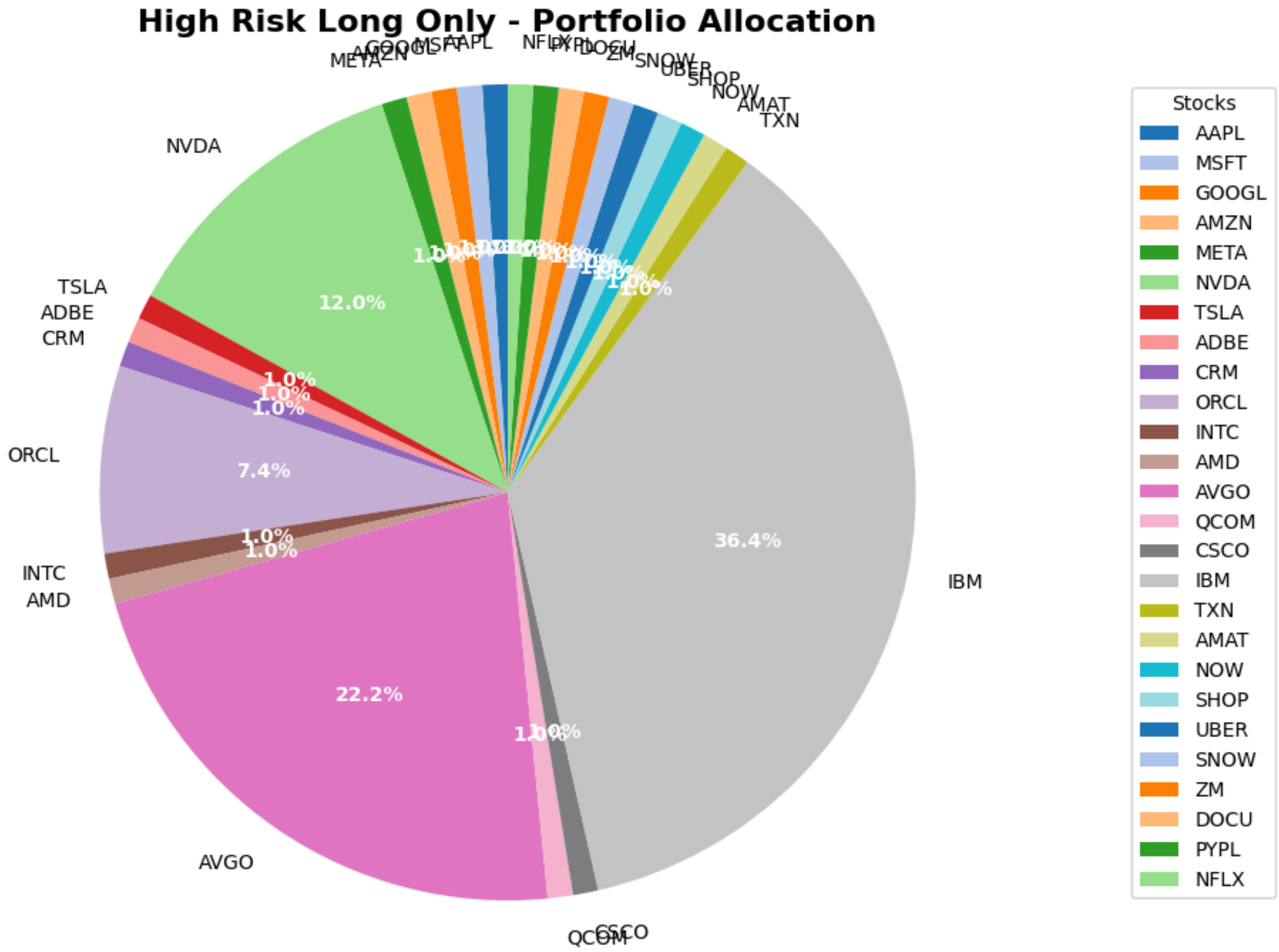


Portfolio Construction: Allocation

Basket 2 (tech-heavy): Long-only



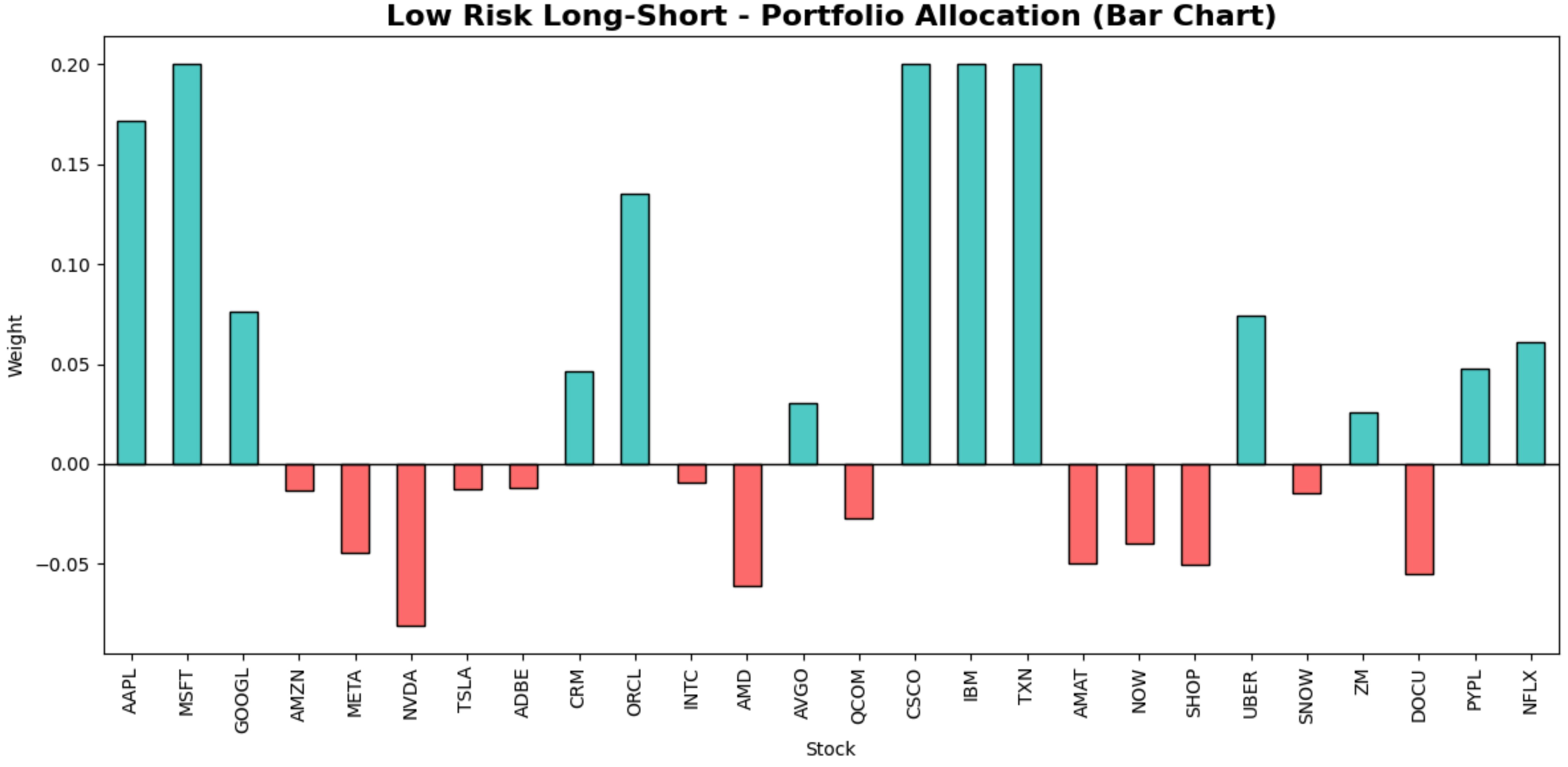
Annual Return Projections for Low Risk Long Only (Based on $100 investment):
Expected Annual Return: 23.13%
Annual Volatility: 25.22%
Sharpe Ratio: 0.82
Expected Value after 1 year: $123.13
95% Confidence Interval: $76.87 - $206.61

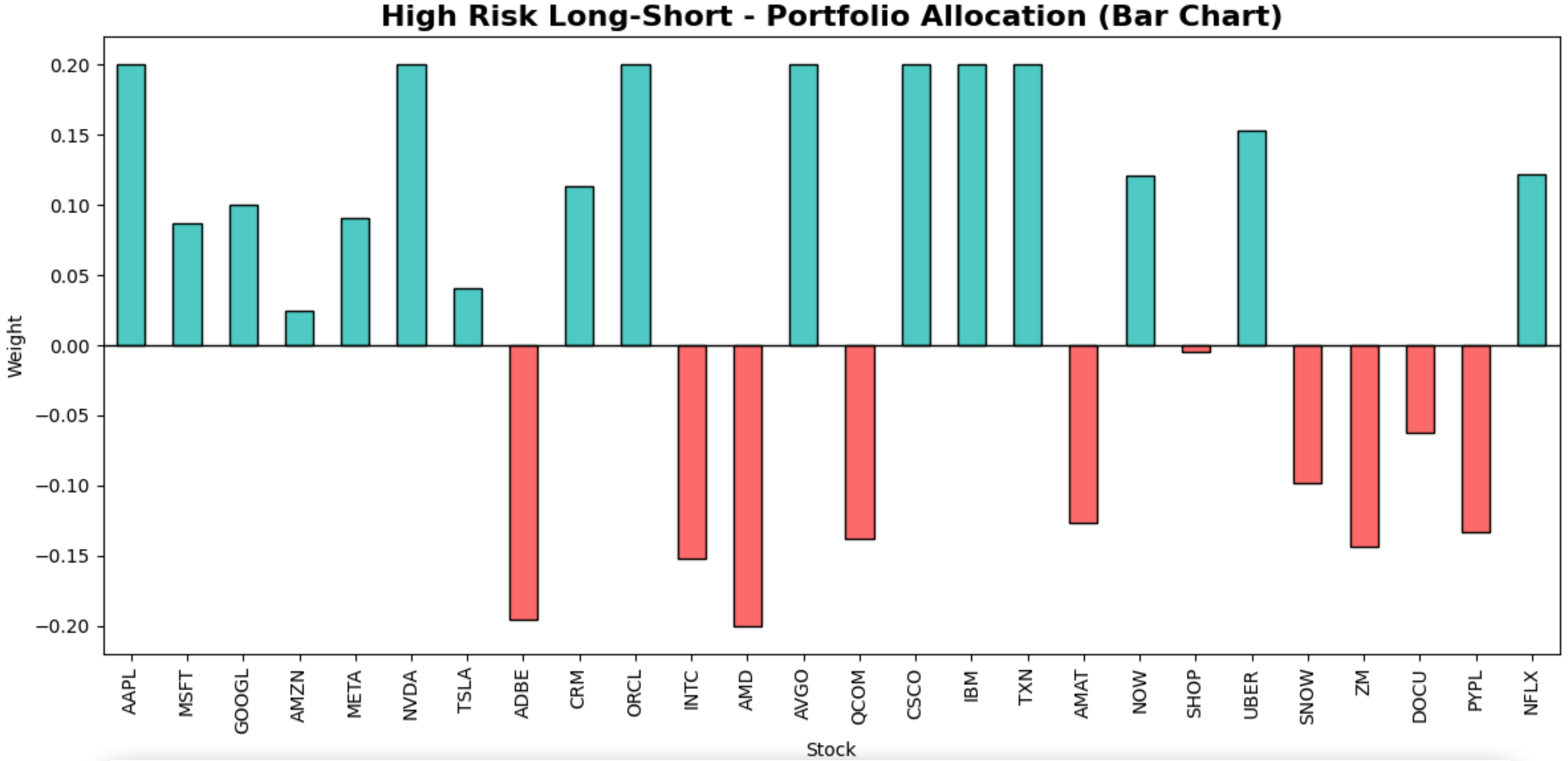


Annual Return Projections for High Risk Long Only (Based on $100 investment):
Expected Annual Return: 32.82%
Annual Volatility: 25.00%
Sharpe Ratio: 1.21
Expected Value after 1 year: $132.82
95% Confidence Interval: $85.06 - $226.63

Portfolio Construction: Allocation

Basket 2 (tech-heavy): Long-short



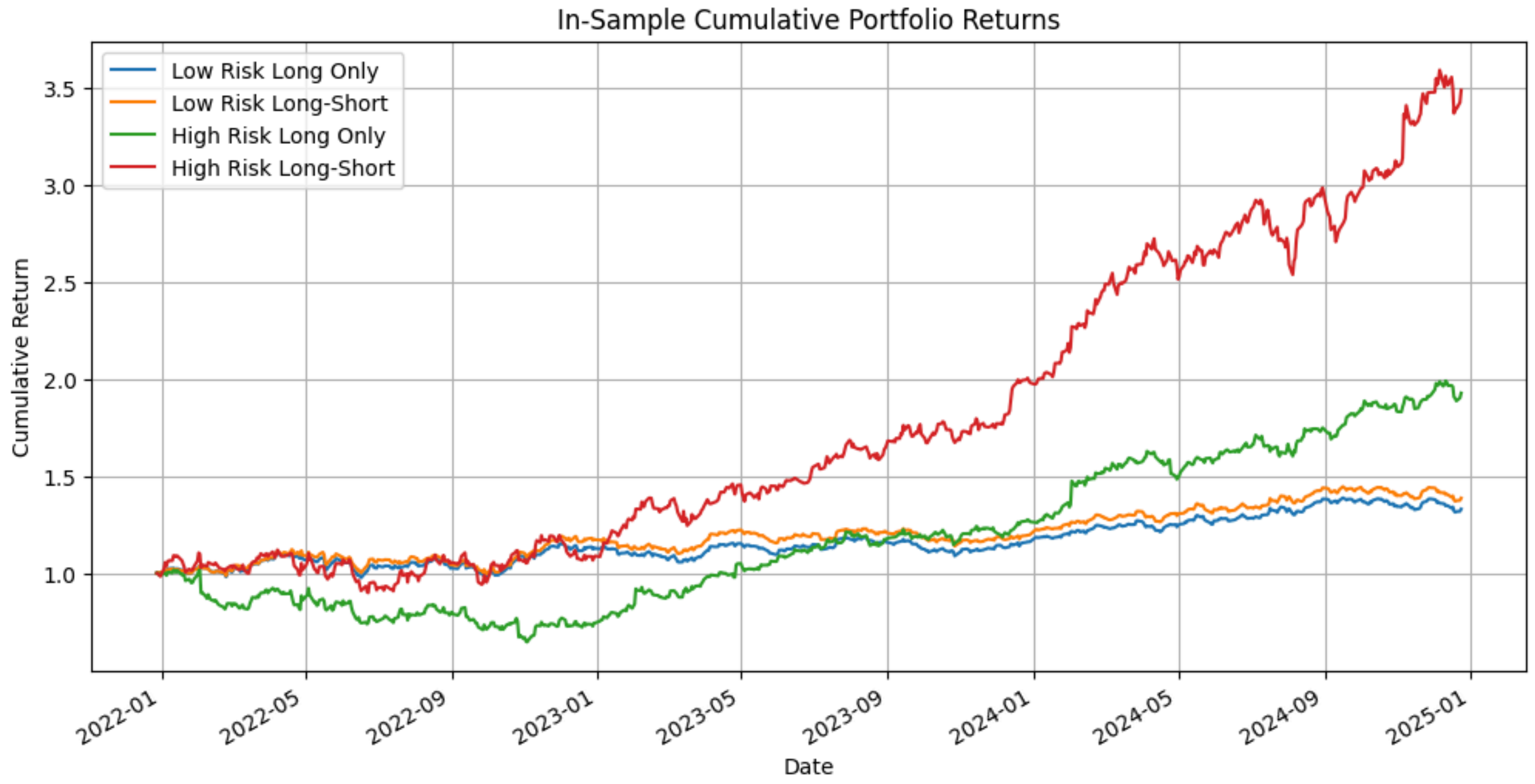



```
Annual Return Projections for Low Risk Long-Short (Based on $100 investment):
Expected Annual Return: 12.02%
Annual Volatility: 16.68%
Sharpe Ratio: 0.57
Expected Value after 1 year: $112.02
95% Confidence Interval: $81.32 - $156.39
```

```
Annual Return Projections for High Risk Long-Short (Based on $100 investment):
Expected Annual Return: 63.06%
Annual Volatility: 25.00%
Sharpe Ratio: 2.42
Expected Value after 1 year: $163.06
95% Confidence Interval: $115.10 - $306.68
```

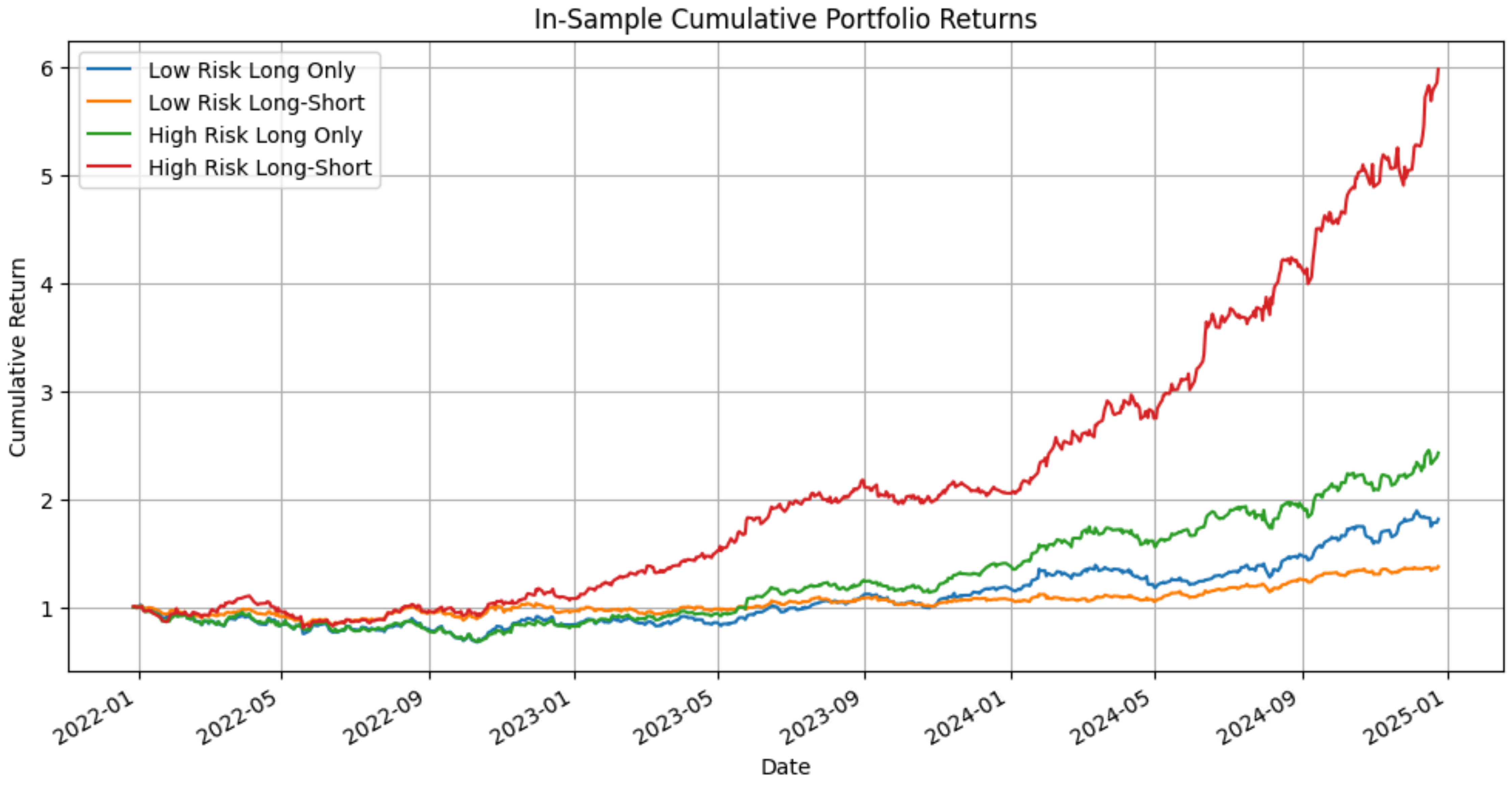
Portfolio Construction: In-sample characteristic

Basket 1




```
Annual Return Projections for Low Risk Long Only (Based on $100 investment):
Expected Annual Return: 10.30%
Annual Volatility: 11.90%
Sharpe Ratio: 0.66
Expected Value after 1 year: $110.30
95% Confidence Interval: $87.79 - $139.96

Annual Return Projections for Low Risk Long-Short (Based on $100 investment):
Expected Annual Return: 11.67%
Annual Volatility: 11.70%
Sharpe Ratio: 0.78
Expected Value after 1 year: $111.67
95% Confidence Interval: $89.34 - $141.35

Annual Return Projections for High Risk Long Only (Based on $100 investment):
Expected Annual Return: 25.05%
Annual Volatility: 24.66%
Sharpe Ratio: 0.91
Expected Value after 1 year: $125.05
95% Confidence Interval: $79.23 - $208.30

Annual Return Projections for High Risk Long-Short (Based on $100 investment):
Expected Annual Return: 44.96%
Annual Volatility: 25.00%
Sharpe Ratio: 1.70
Expected Value after 1 year: $144.96
95% Confidence Interval: $96.04 - $255.89
```

Basket 2

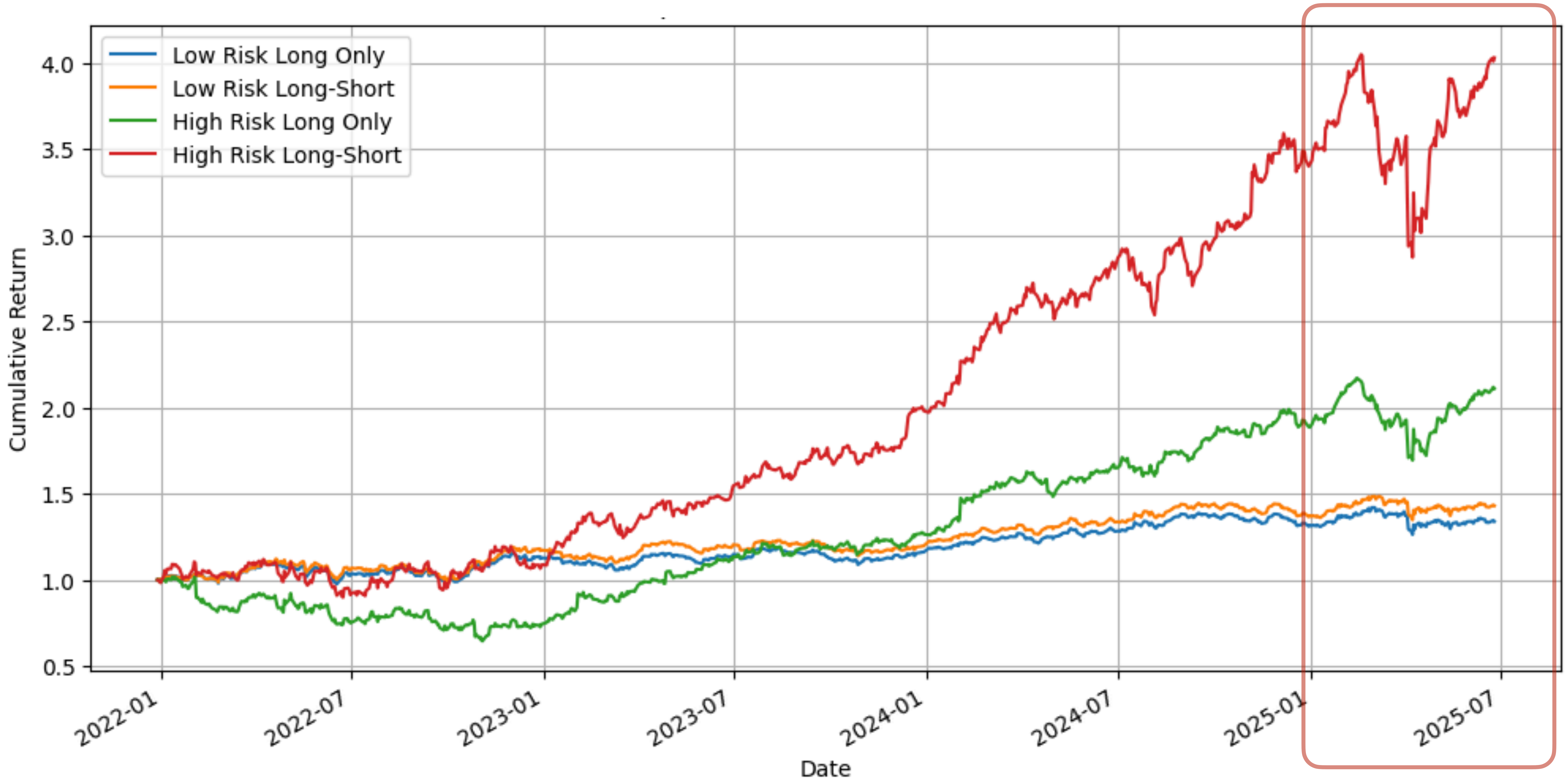



```
Annual Return Projections for Low Risk Long Only (Based on $100 investment):
Expected Annual Return: 23.13%
Annual Volatility: 25.22%
Sharpe Ratio: 0.82
Expected Value after 1 year: $123.13
95% Confidence Interval: $76.87 - $206.61

Annual Return Projections for Low Risk Long-Short (Based on $100 investment):
Expected Annual Return: 12.02%
Annual Volatility: 16.68%
Sharpe Ratio: 0.57
Expected Value after 1 year: $112.02
95% Confidence Interval: $81.32 - $156.39

Annual Return Projections for High Risk Long Only (Based on $100 investment):
Expected Annual Return: 32.82%
Annual Volatility: 25.00%
Sharpe Ratio: 1.21
Expected Value after 1 year: $132.82
95% Confidence Interval: $85.06 - $226.63

Annual Return Projections for High Risk Long-Short (Based on $100 investment):
Expected Annual Return: 63.06%
Annual Volatility: 25.00%
Sharpe Ratio: 2.42
Expected Value after 1 year: $163.06
95% Confidence Interval: $115.10 - $306.68
```

Portfolio Construction: Backtesting

Basket 1

Backtesting period




```
Annual Return Projections for High Risk Long-Short (Based on $100 investment):
Expected Annual Return: 44.96%
Annual Volatility: 25.00%
Sharpe Ratio: 1.70
Expected Value after 1 year: $144.96
95% Confidence Interval: $96.04 - $255.89
```

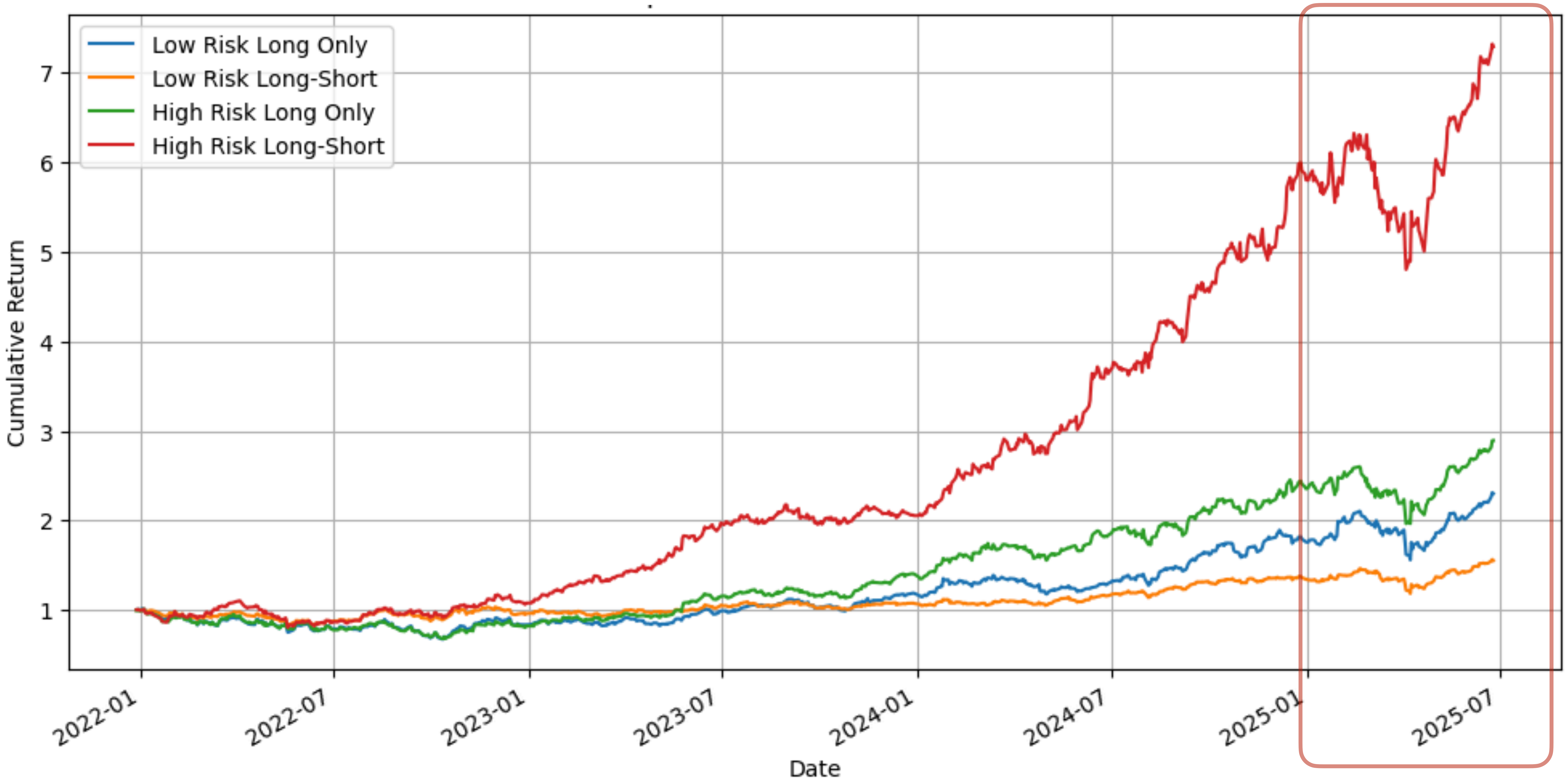
```
Backtest Results for High Risk Long-Short:
Cumulative Return: 15.70%
Annualized Return: 37.56%
Annualized Volatility: 38.70%
Sharpe Ratio: 0.91
```

Decreased Sharpe ratio

Portfolio Construction: Backtesting

Basket 2

Backtesting period



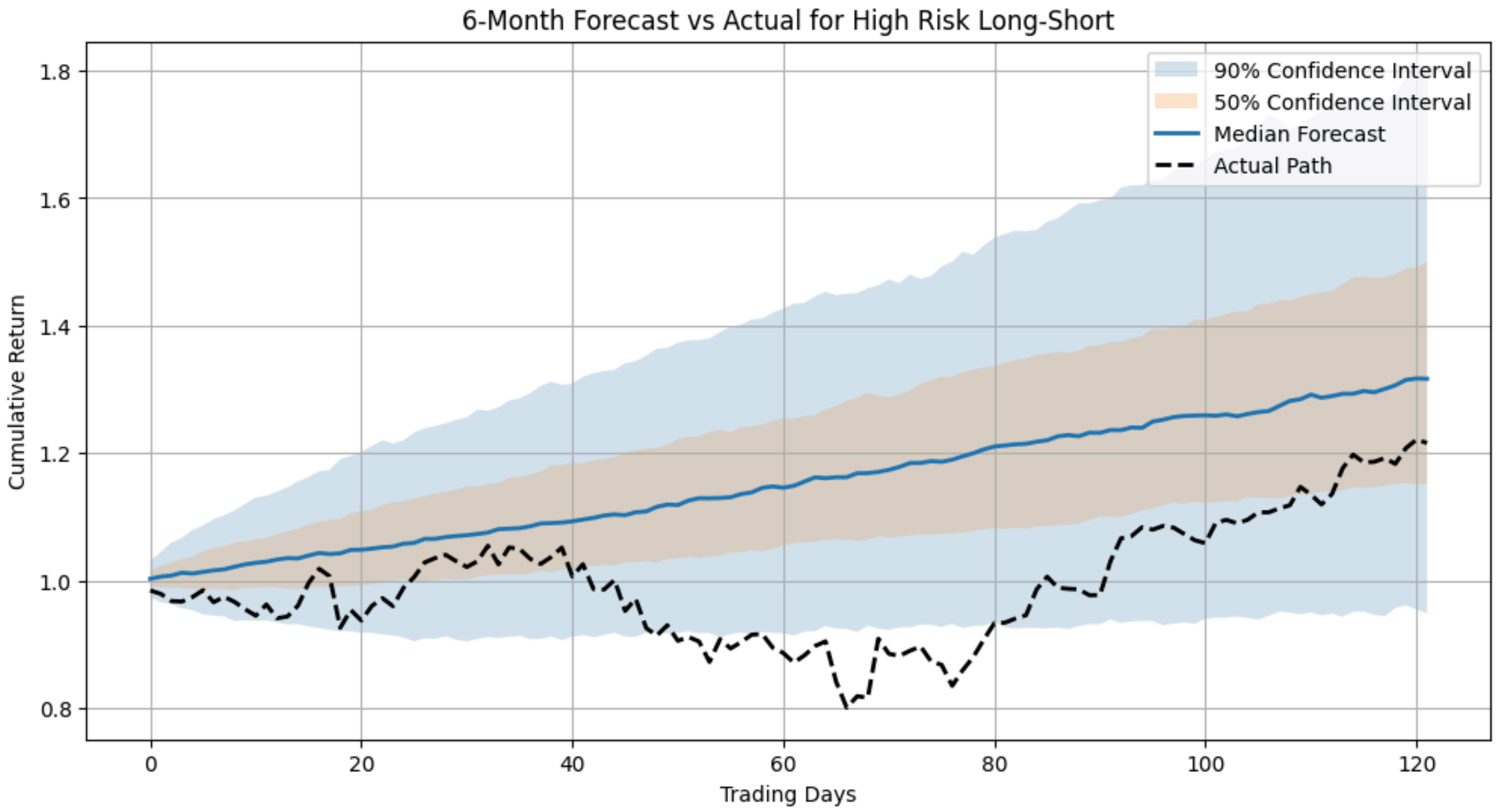

```
Annual Return Projections for High Risk Long-Short (Based on $100 investment):
Expected Annual Return: 63.06%
Annual Volatility: 25.00%
Sharpe Ratio: 2.42
Expected Value after 1 year: $163.06
95% Confidence Interval: $115.10 - $306.68
```

```
Backtest Results for High Risk Long-Short:
Cumulative Return: 21.61%
Annualized Return: 47.77%
Annualized Volatility: 38.49%
Sharpe Ratio: 1.2
```

Decreased Sharpe ratio

Portfolio Construction: Forecasting

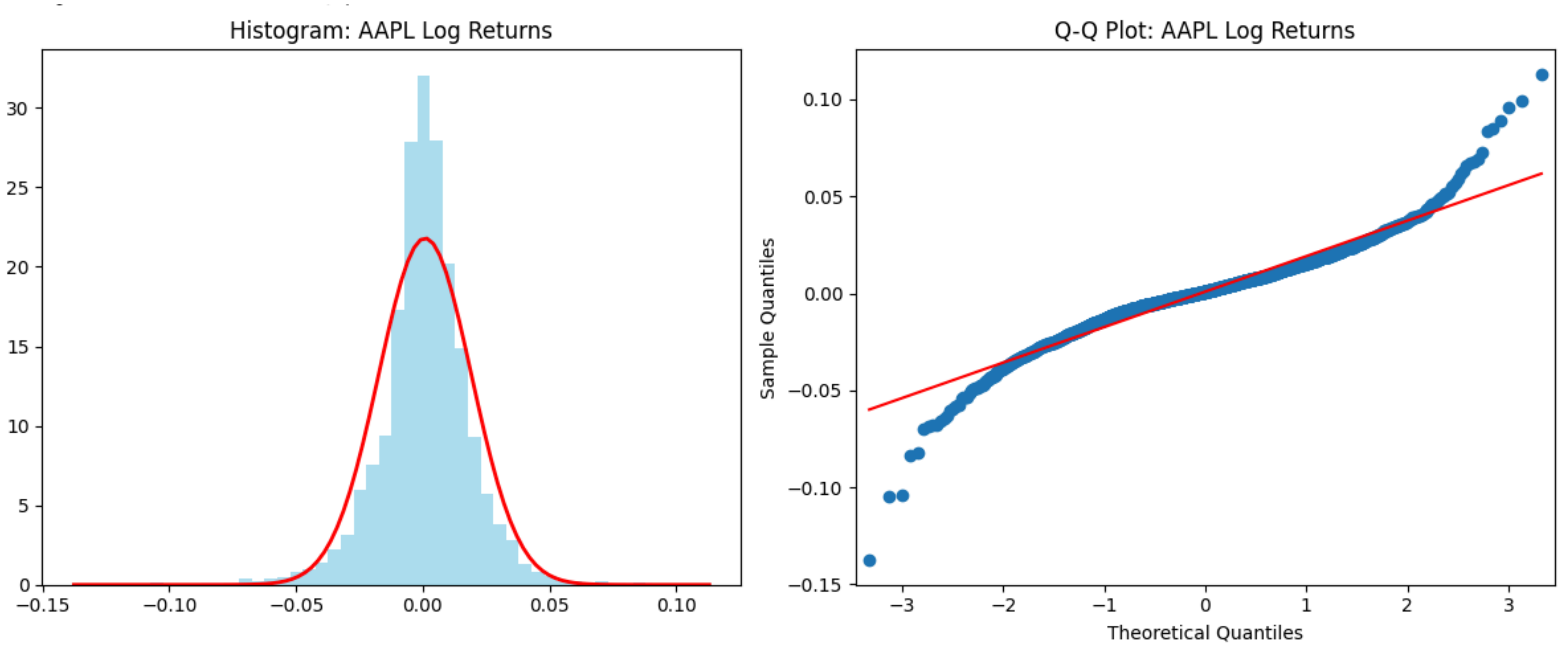
Forecasting Basket 2 (tech-heavy) high risk long short portfolio



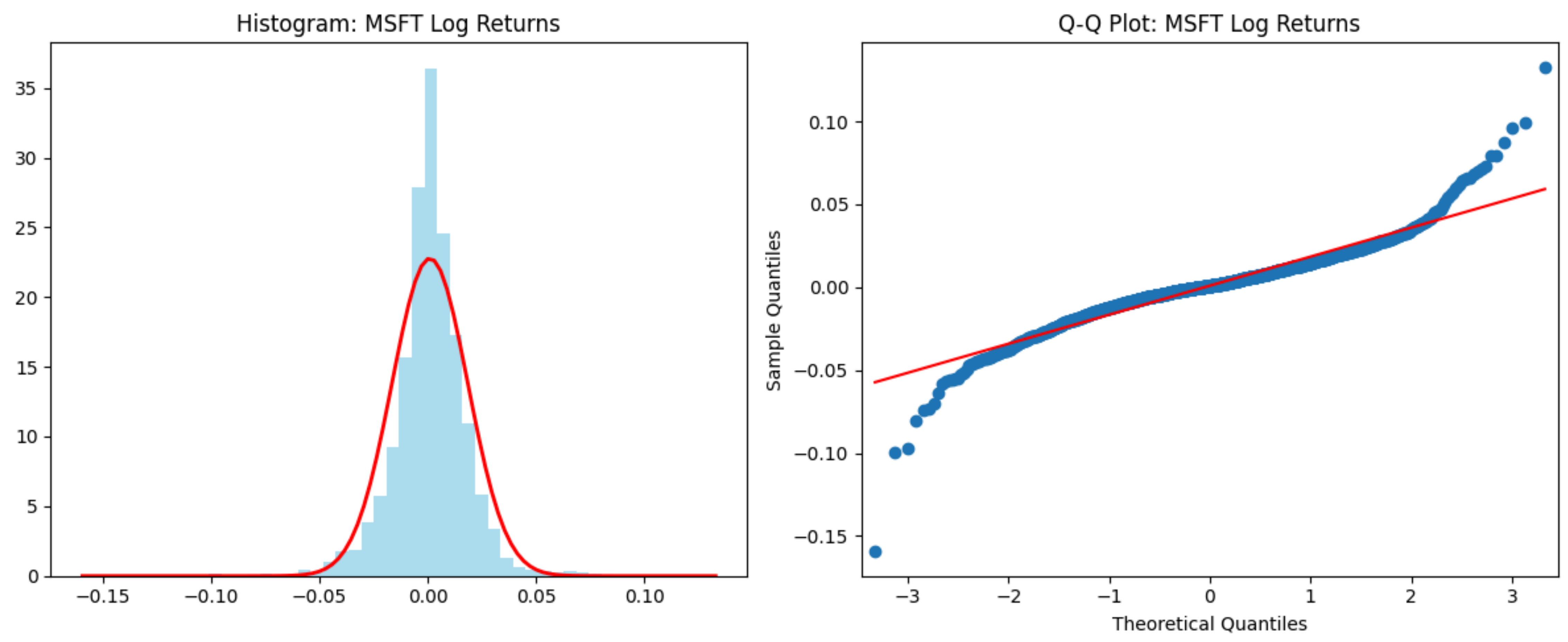
Project-2
Normality Testing

When is financial data normally distributed?

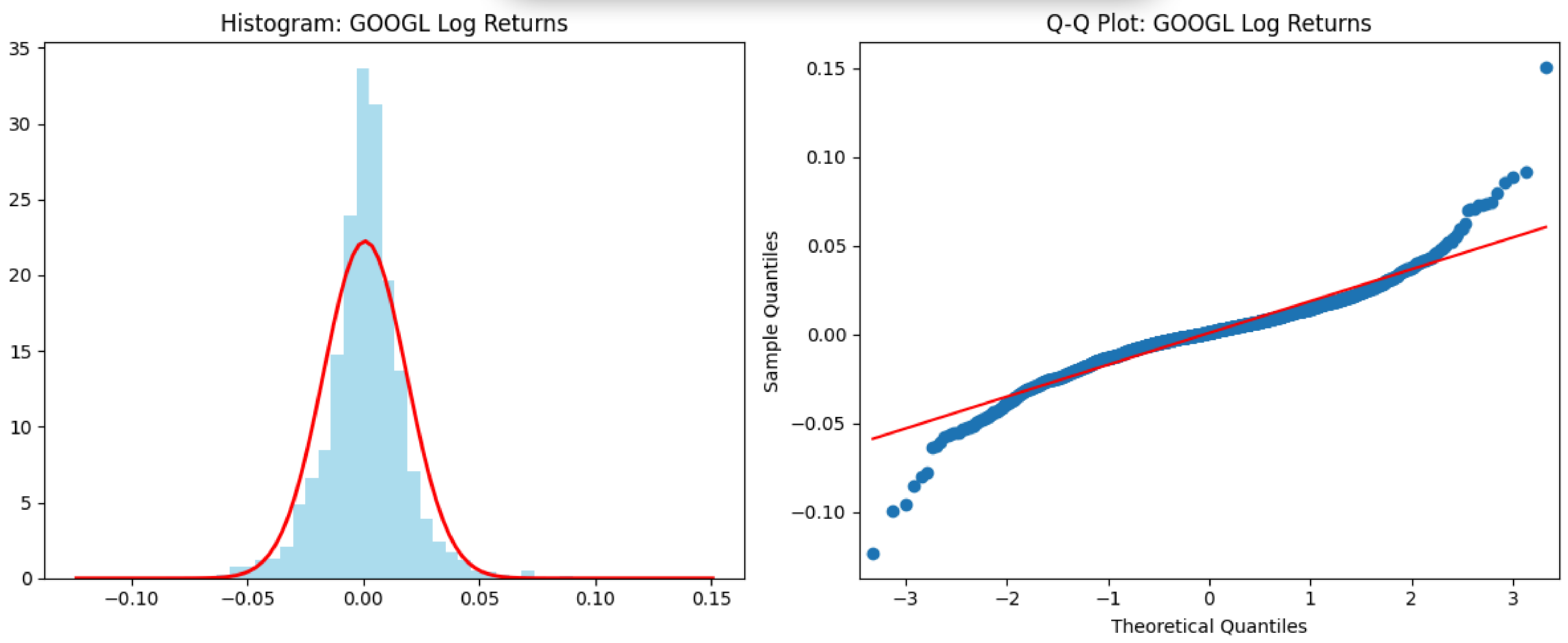
Historical log return of a stock/index data shows strong evidence against normality



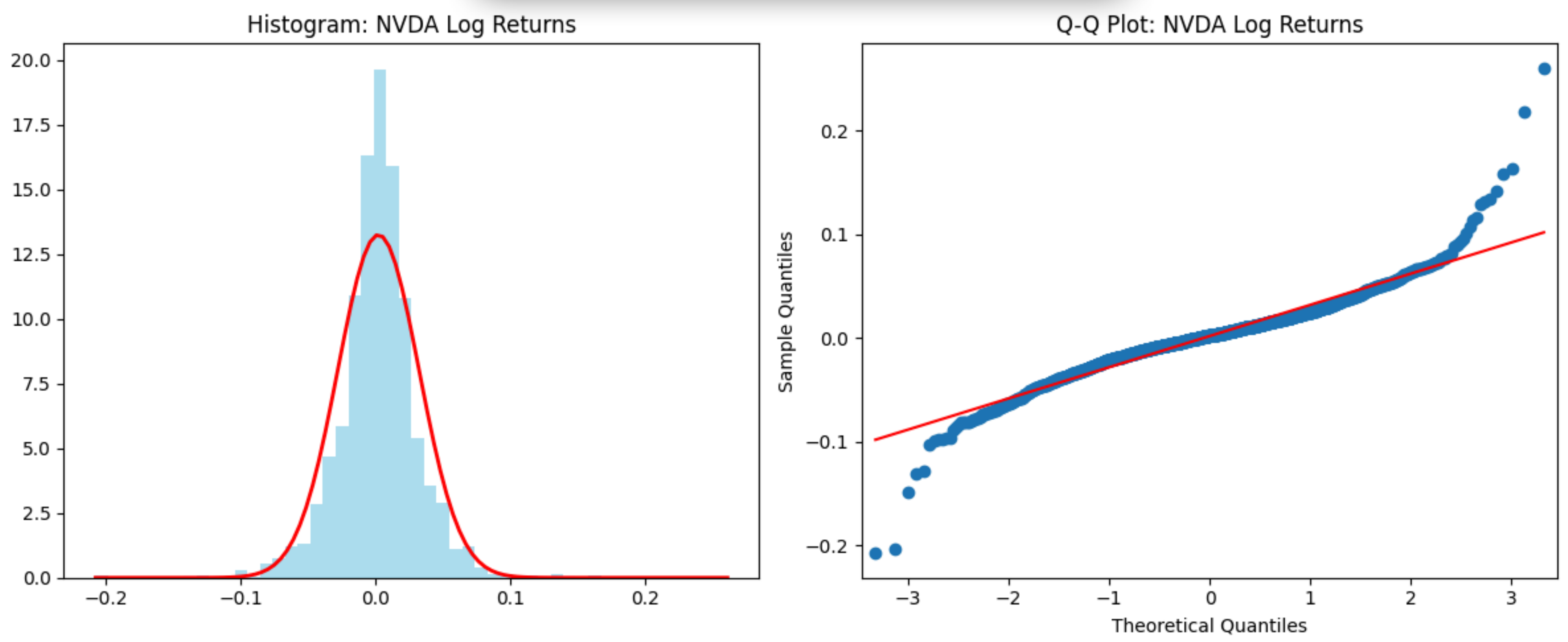

```
--- AAPL ---
Shapiro-Wilk: stat=0.9416, p=0.0000
Anderson-Darling: stat=25.5381
D'Agostino-Pearson: stat=290.0381, p=0.0000
Kolmogorov-Smirnov: stat=0.0788, p=0.0000
```




```
--- MSFT ---
Shapiro-Wilk: stat=0.9249, p=0.0000
Anderson-Darling: stat=30.0568
D'Agostino-Pearson: stat=362.7032, p=0.0000
Kolmogorov-Smirnov: stat=0.0804, p=0.0000
```



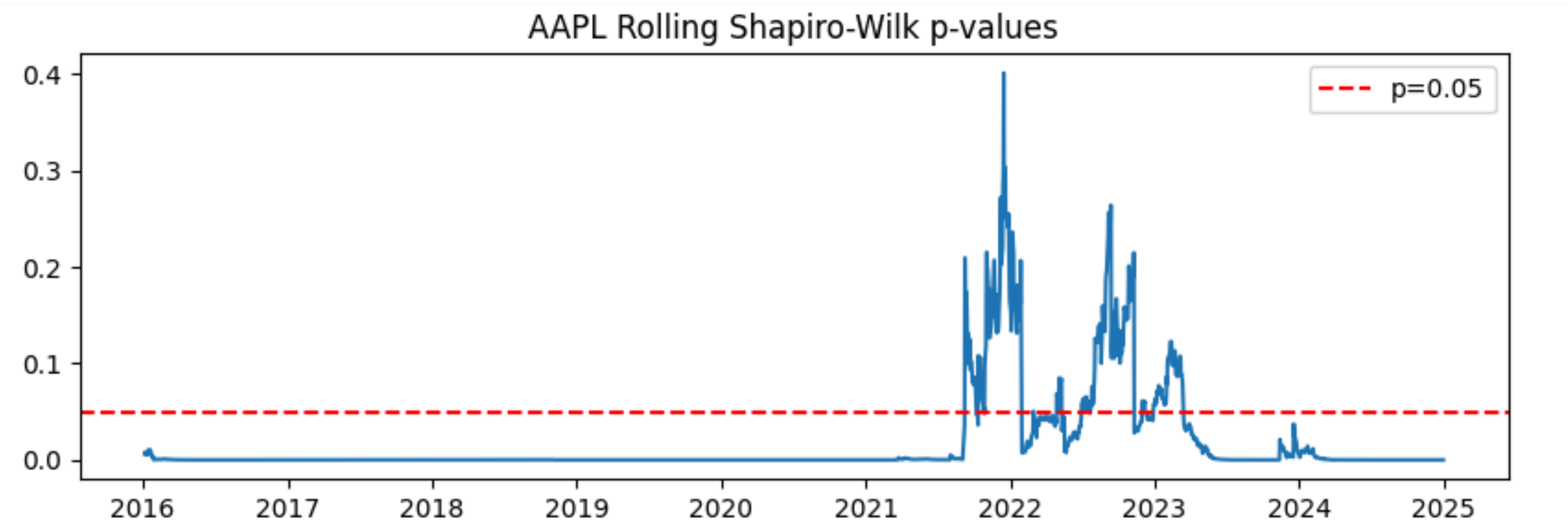

```
--- GOOGL ---
Shapiro-Wilk: stat=0.9340, p=0.0000
Anderson-Darling: stat=29.5181
D'Agostino-Pearson: stat=302.9749, p=0.0000
Kolmogorov-Smirnov: stat=0.0789, p=0.0000
```

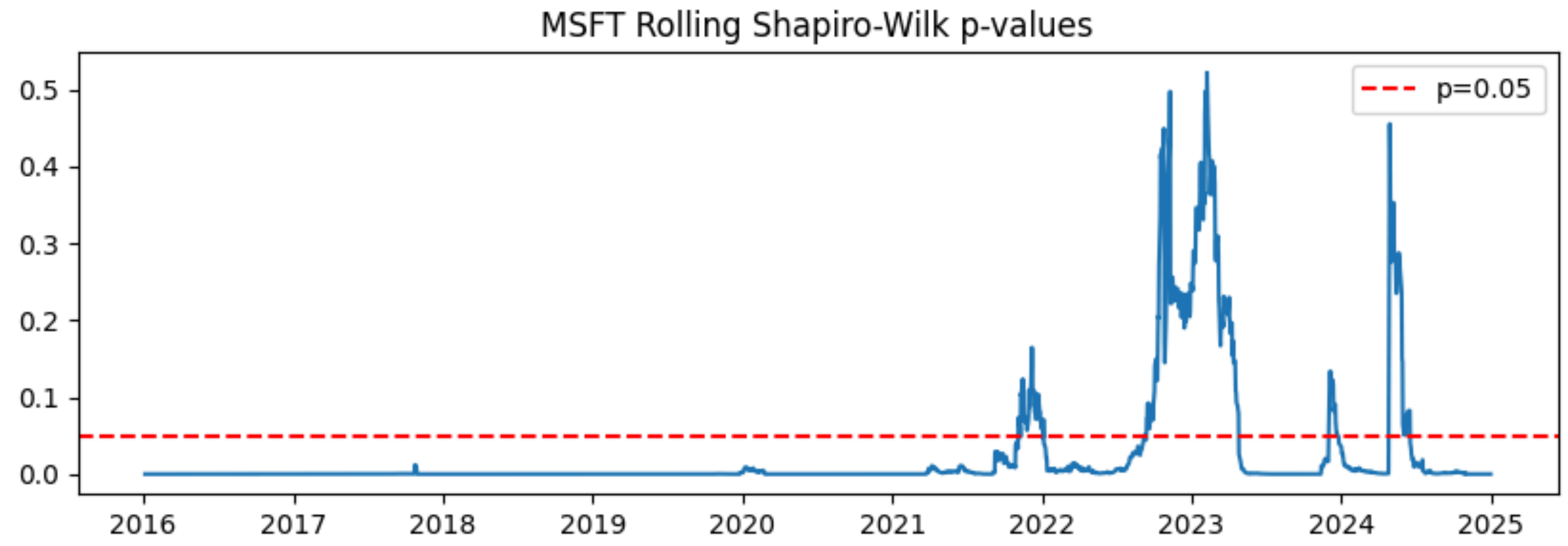


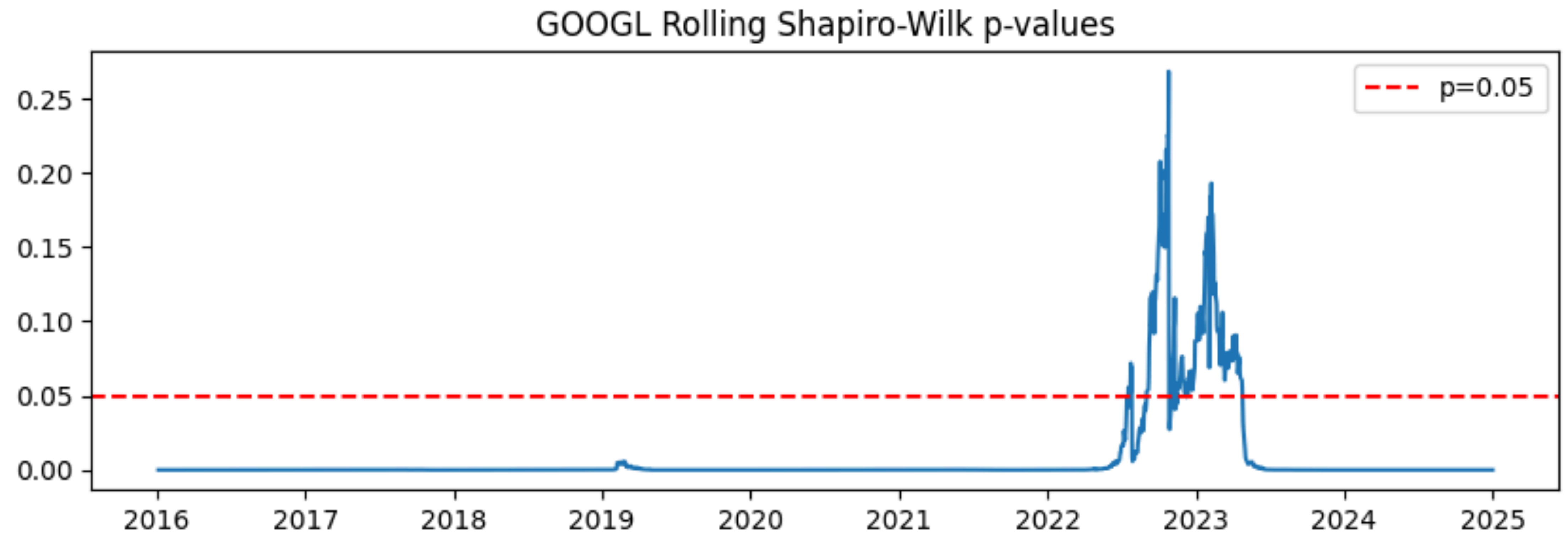

```
--- NVDA ---
Shapiro-Wilk: stat=0.9330, p=0.0000
Anderson-Darling: stat=25.0793
D'Agostino-Pearson: stat=357.3382, p=0.0000
Kolmogorov-Smirnov: stat=0.0696, p=0.0000
```

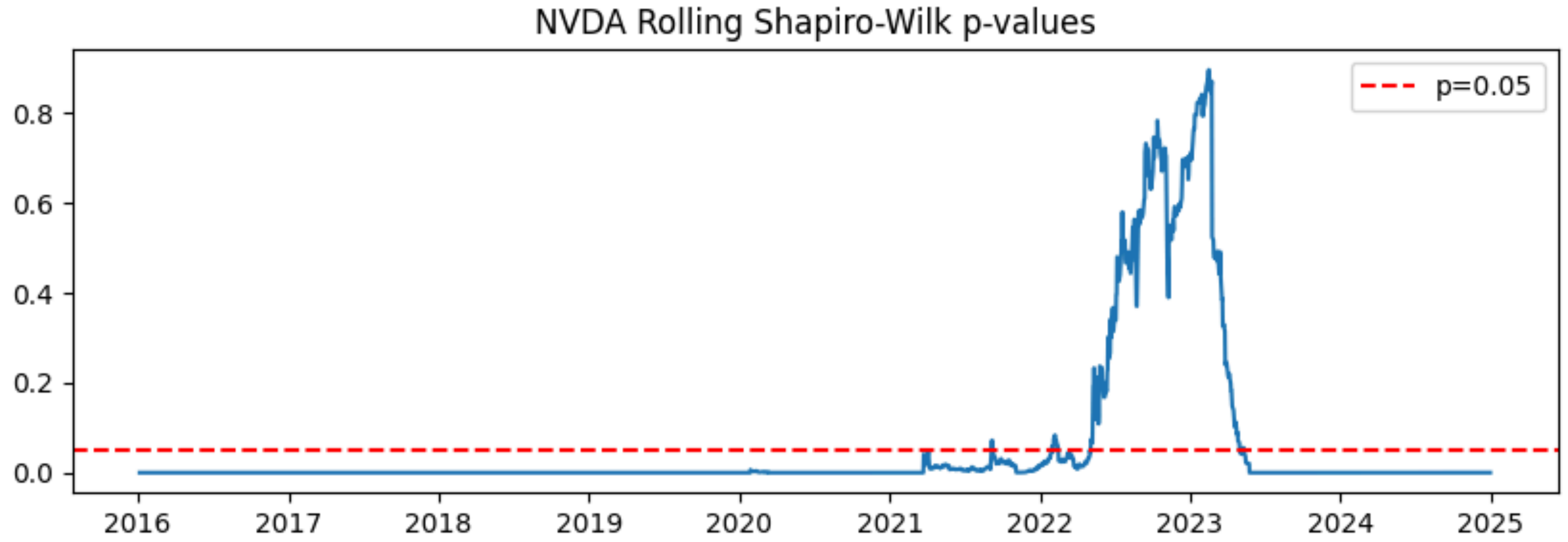
But are period of times when the log-returns have some evidence of normal distribution?

We consider the daily log-return data of stocks: ['AAPL', 'MSFT', 'GOOGL', 'AMZN', 'META', 'TSLA', 'JPM', 'NVDA', 'SPY', ‘QQQ'] from 2015-01-01 to 2025-01-01 and perform of a Rolling Shapiro-Wilk p-values with window of 252 days









There are one-year periods where Shapiro-Wilk p-value $p>0.05$

Therefore cannot reject that the data is normal. However need more tests …

From the Rolling Shapiro-Wilk p-value test get the following periods ranked by period length

GOOGL: 2022-12-12 to 2023-04-25 (~90+ days)

Shapiro-Wilk: stat=0.9800, p=0.1687
Anderson-Darling: stat=0.3585
D'Agostino-Pearson: stat=5.3203, p=0.0699
Kolmogorov-Smirnov: stat=0.0605, p=0.8684

AAPL: 2022-12-30 to 2023-03-15 (~50+ days)

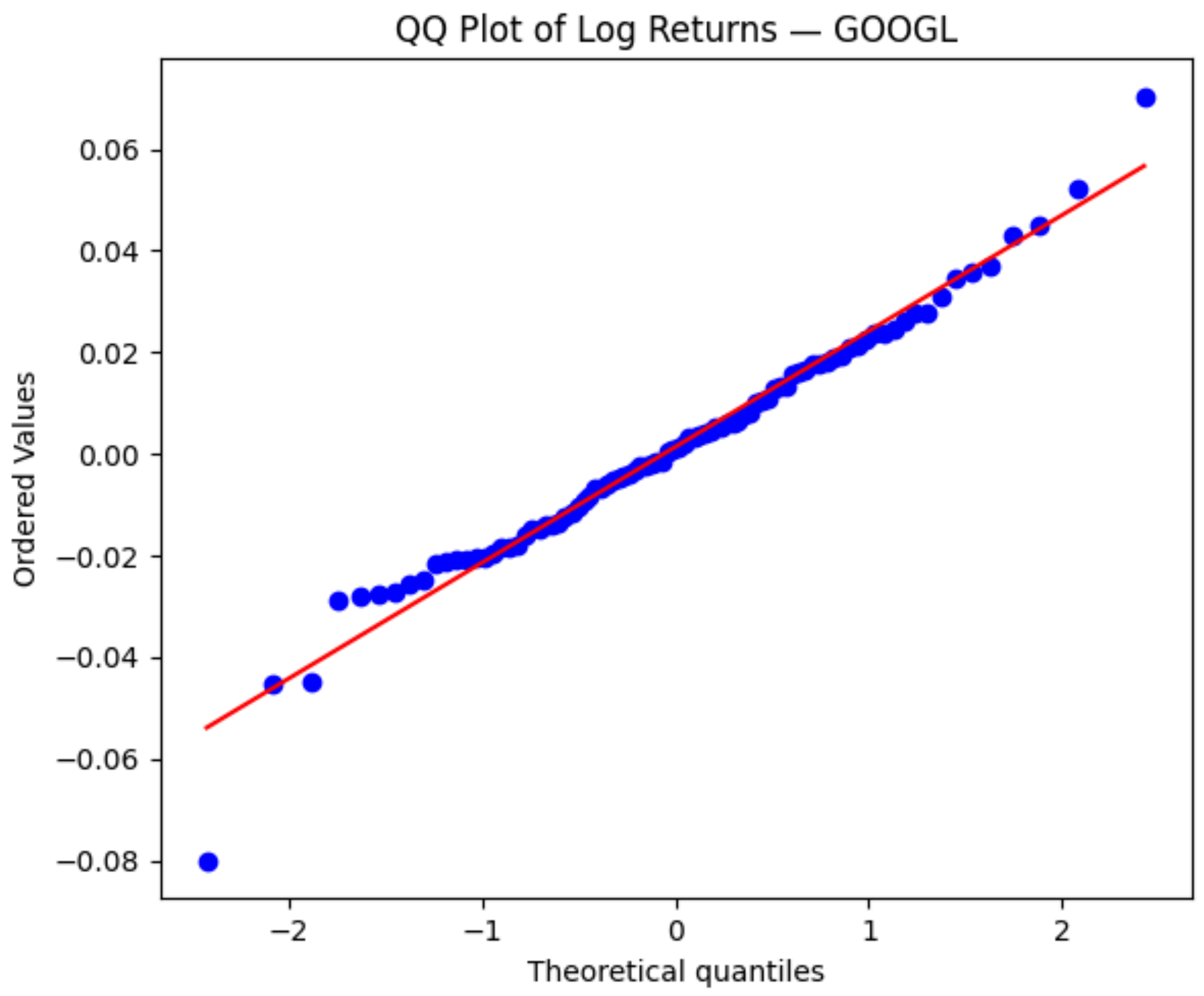
Shapiro-Wilk: stat=0.9863, p=0.8177
Anderson-Darling: stat=0.2424
D'Agostino-Pearson: stat=0.2361, p=0.8887
Kolmogorov-Smirnov: stat=0.0872, p=0.8011

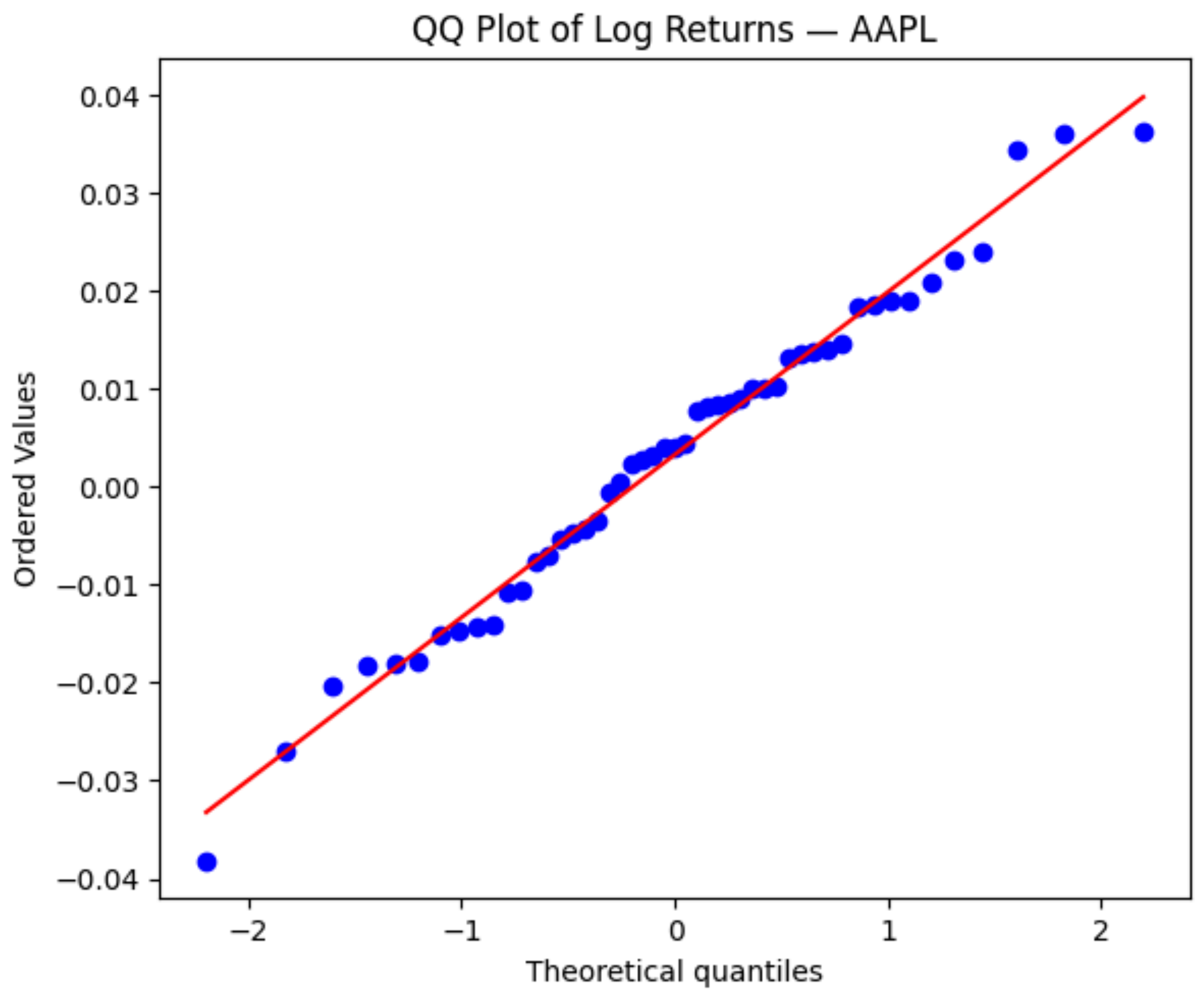
TSLA: 2022-11-07 to 2023-01-03 (~40 days)

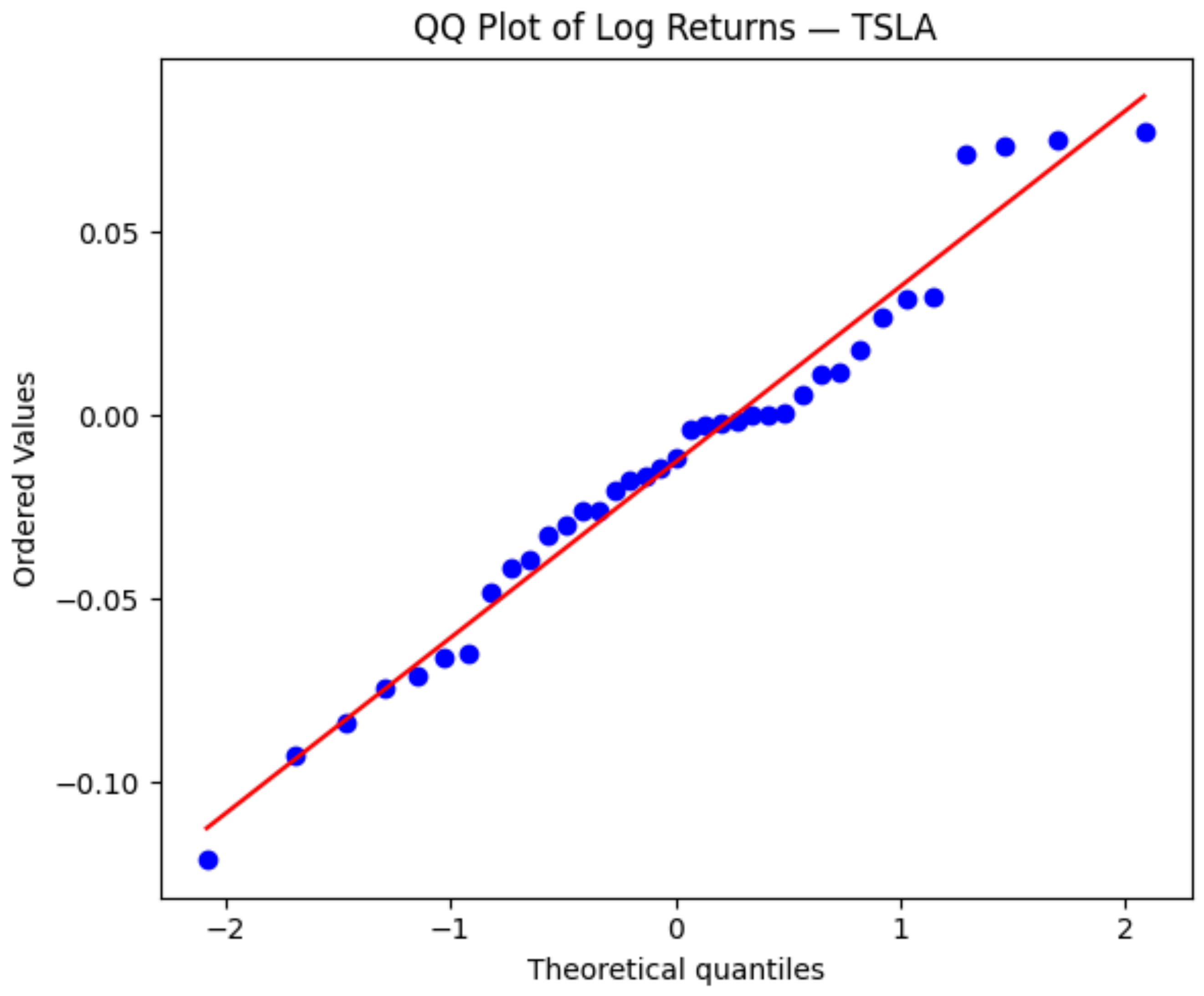
Shapiro-Wilk: stat=0.9732, p=0.4691
Anderson-Darling: stat=0.3494
D'Agostino-Pearson: stat=0.2776, p=0.8704
Kolmogorov-Smirnov: stat=0.0803, p=0.9456

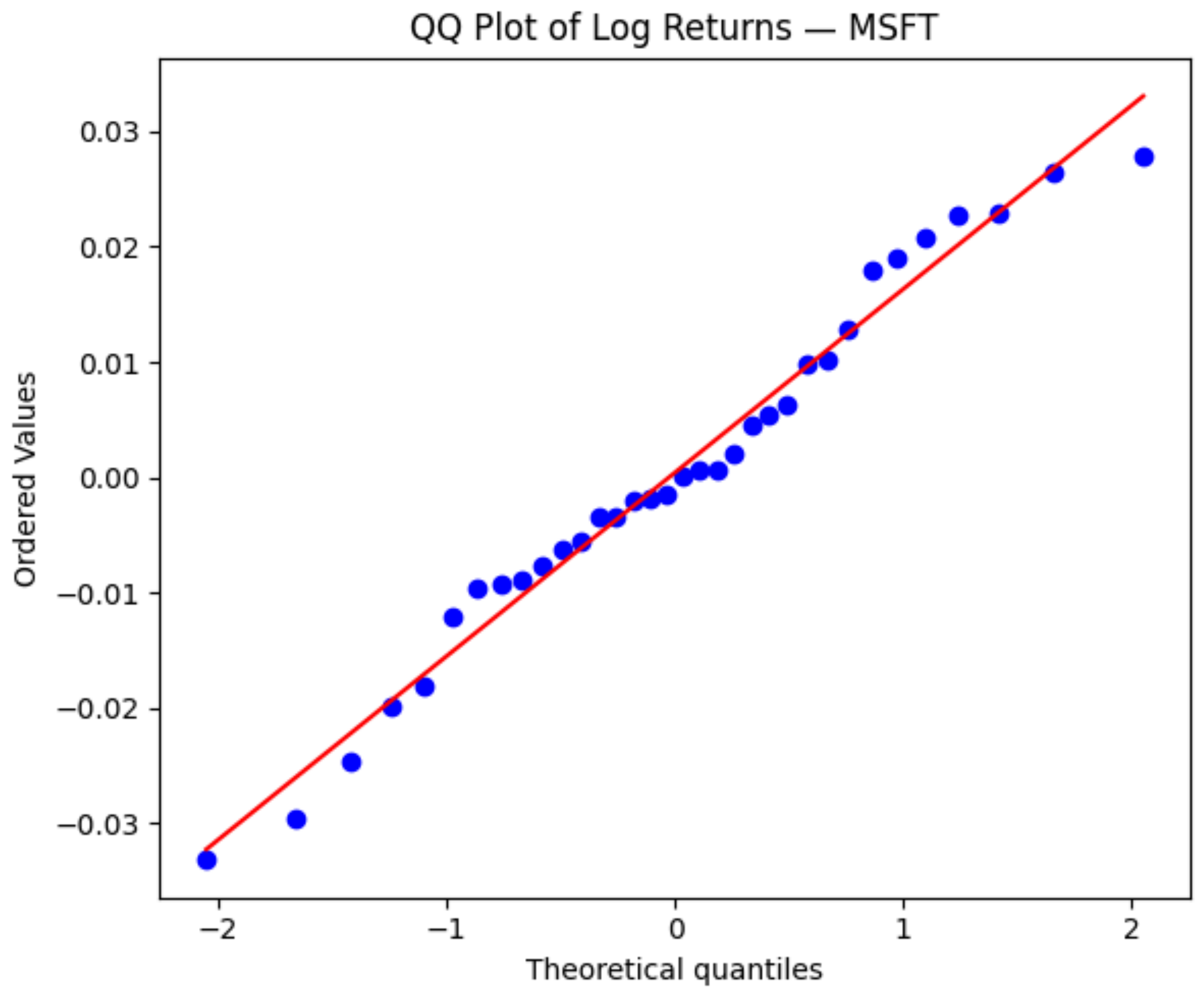
MSFT: 2021-11-11 to 2022-01-03 (~35 days)

Shapiro-Wilk: stat=0.9752, p=0.5842
Anderson-Darling: stat=0.3011
D'Agostino-Pearson: stat=0.1215, p=0.9411
Kolmogorov-Smirnov: stat=0.0879, p=0.9207

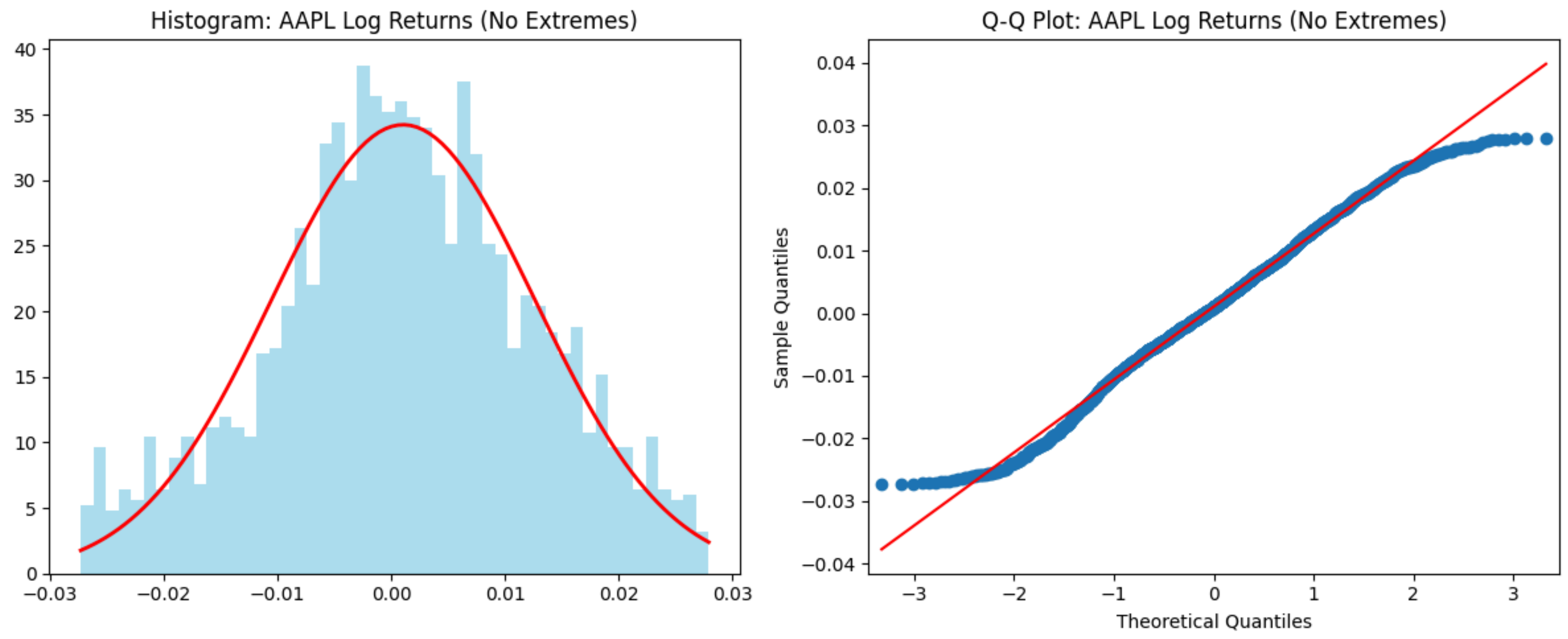




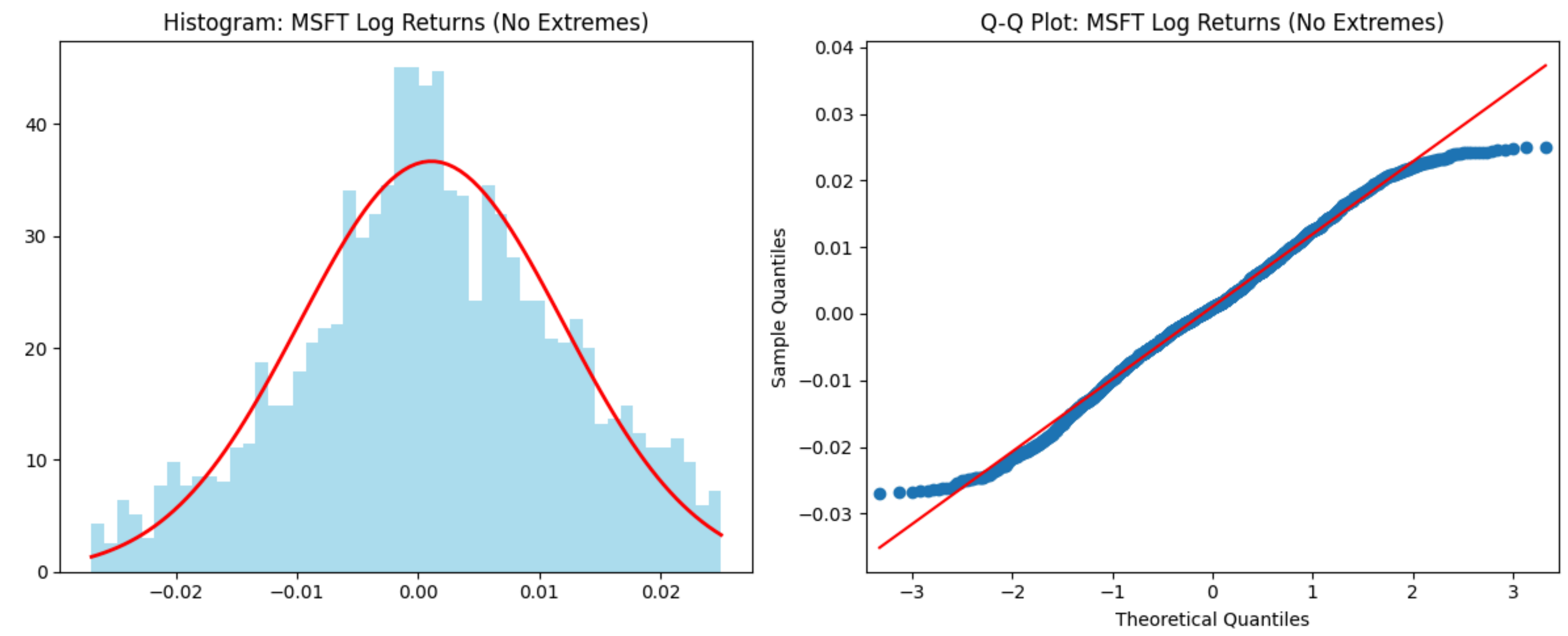




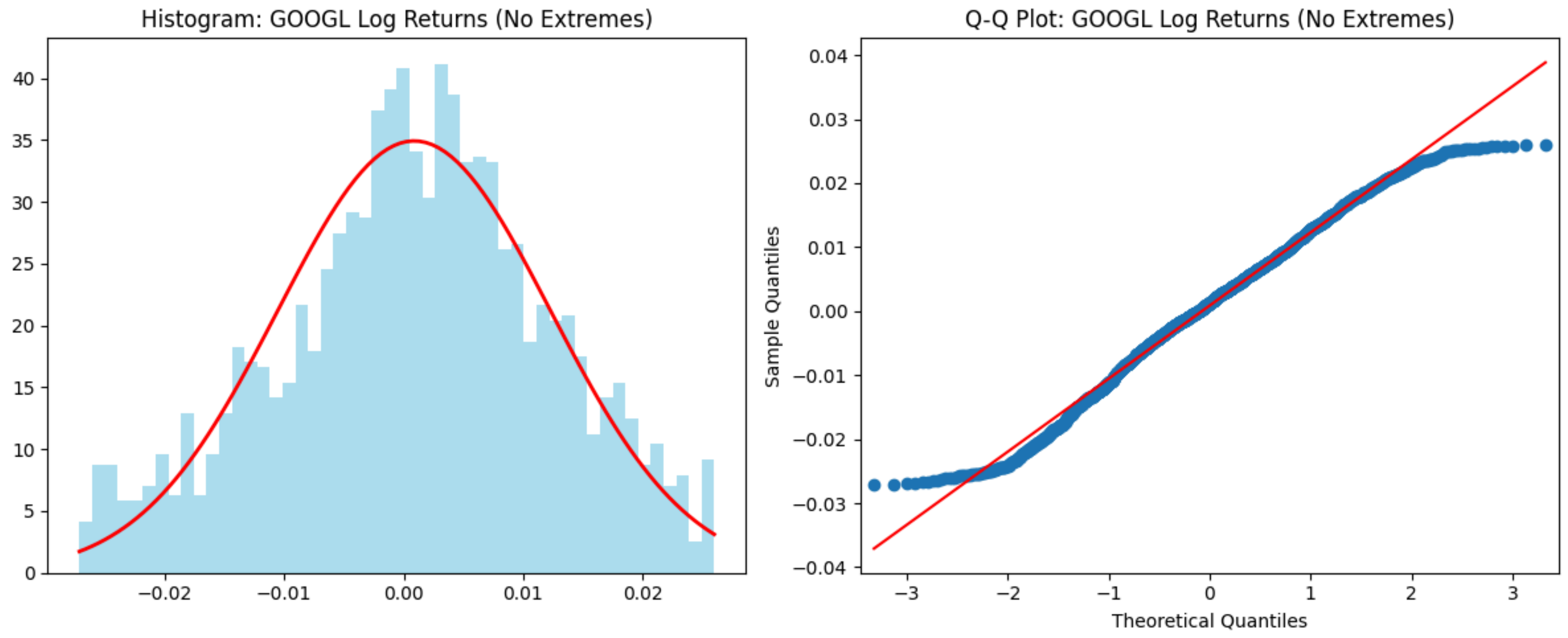
Furthermore removing extremal return data creates a distribution with evidence of being normal



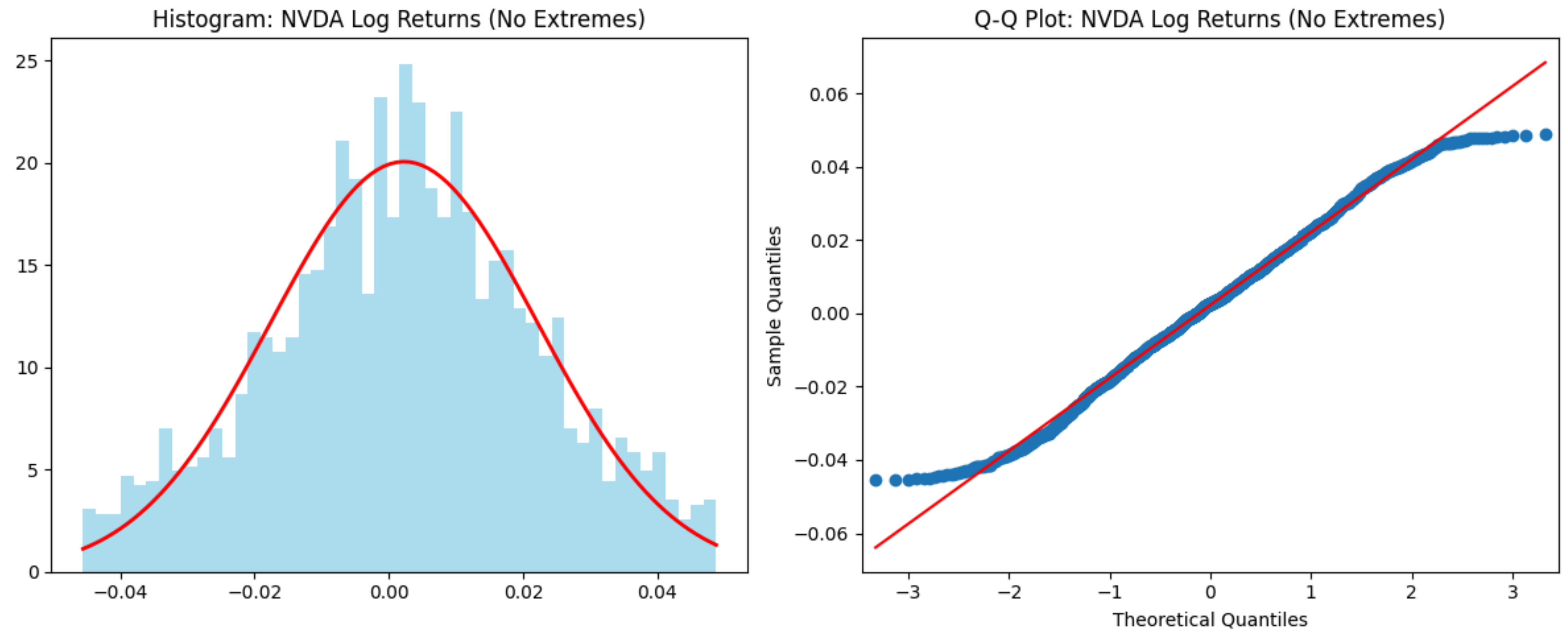

```
--- AAPL (extremes removed) ---
Shapiro-Wilk: stat=0.9928, p=0.0000
Anderson-Darling: stat=1.9460
D'Agostino-Pearson: stat=17.9135, p=0.0001
Kolmogorov-Smirnov: stat=0.0210, p=0.2660
```




```
--- MSFT (extremes removed) ---
Shapiro-Wilk: stat=0.9924, p=0.0000
Anderson-Darling: stat=2.3144
D'Agostino-Pearson: stat=18.4311, p=0.0001
Kolmogorov-Smirnov: stat=0.0230, p=0.1785
```




```
--- GOOGL (extremes removed) ---
Shapiro-Wilk: stat=0.9908, p=0.0000
Anderson-Darling: stat=3.1972
D'Agostino-Pearson: stat=30.0170, p=0.0000
Kolmogorov-Smirnov: stat=0.0290, p=0.0435
```




```
--- NVDA (extremes removed) ---
Shapiro-Wilk: stat=0.9942, p=0.0000
Anderson-Darling: stat=1.1597
D'Agostino-Pearson: stat=18.1879, p=0.0001
Kolmogorov-Smirnov: stat=0.0152, p=0.6680
```

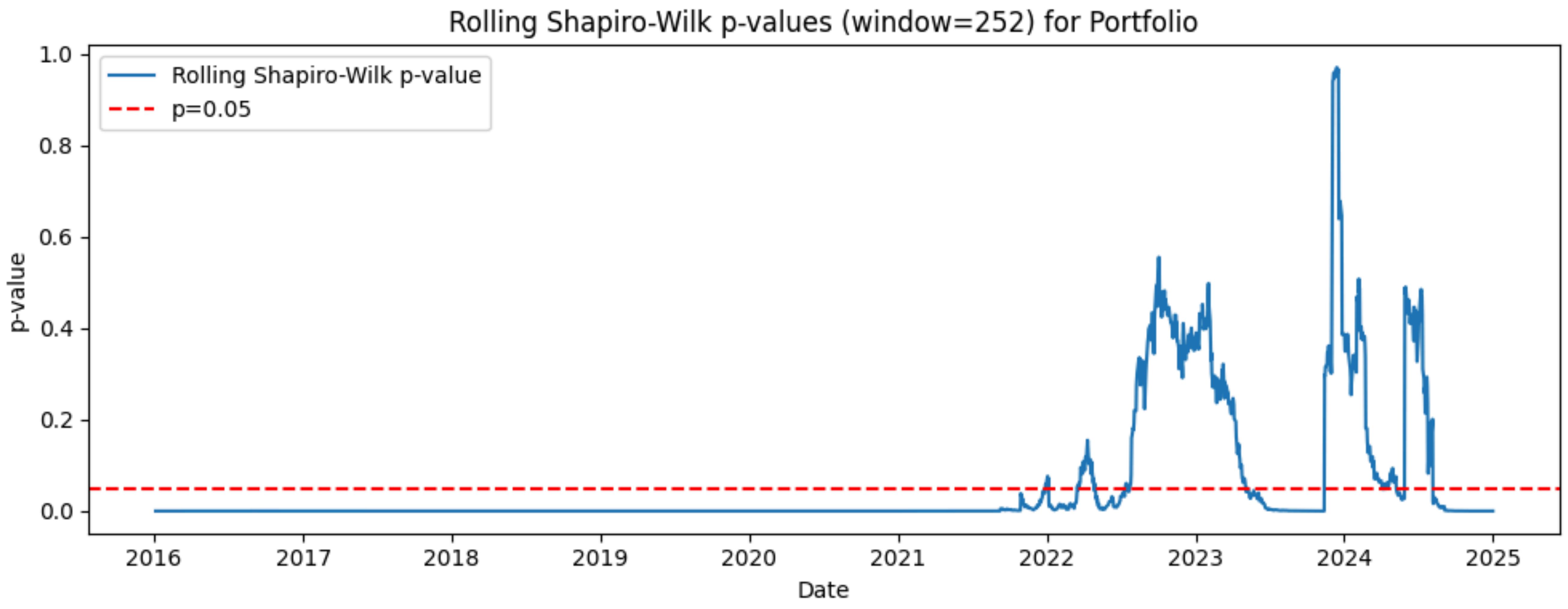
A portfolio of stocks with historical log return data that is normally distributed

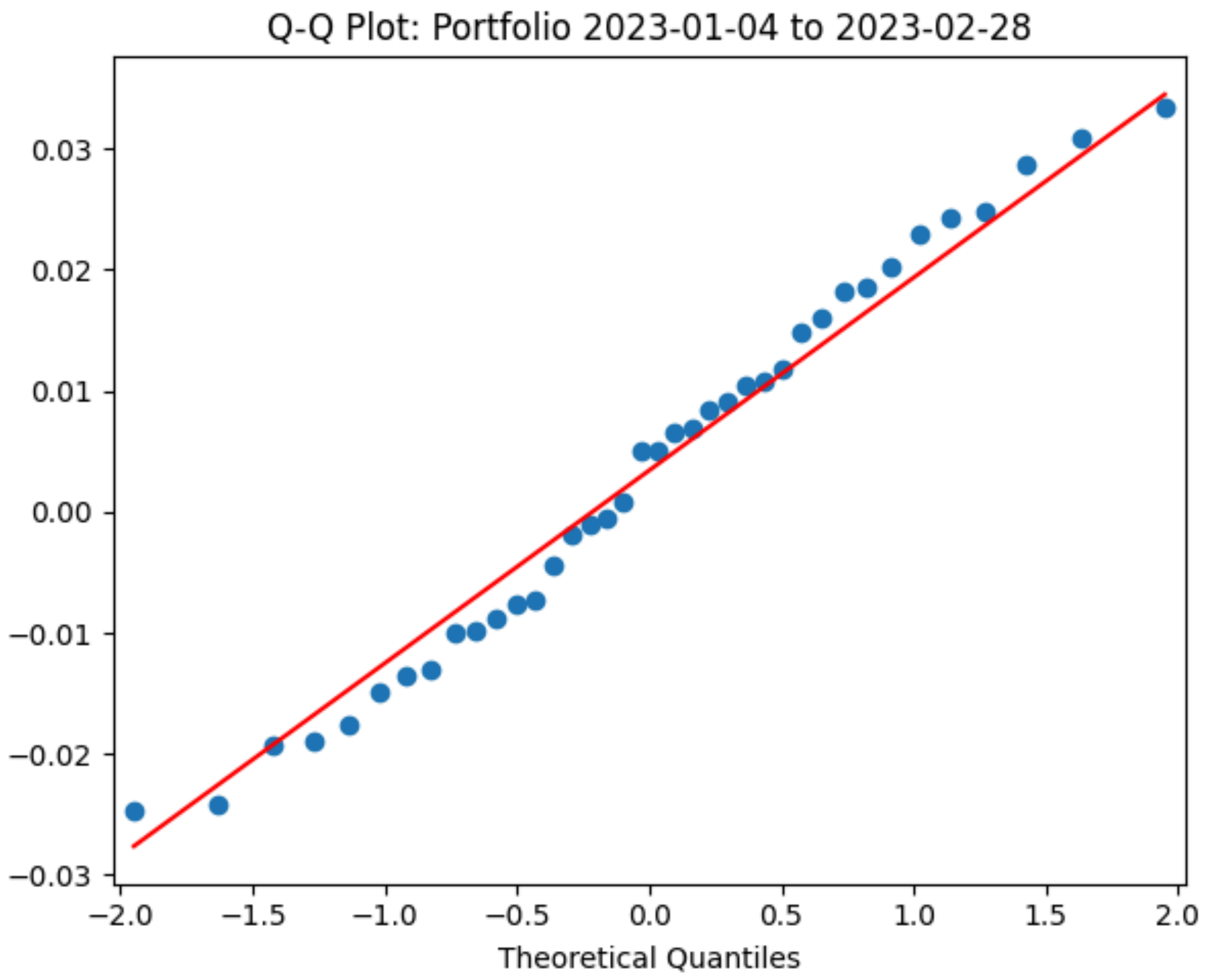
Based on the Shapiro-Wilk p-value test, we searched for the longest window containing at least 5 stocks that passes the test. We then construct an equal weight portfolio off these stocks

Equal weight portfolio:

```
Largest window: 2023-01-04 to 2023-02-28 (56 days)
Stocks: JPM, SPY, GOOGL, NVDA, AAPL, MSFT, QQQ
```

Running a rolling Shapiro-Wilk p-value on the portfolio shows more periods of normality



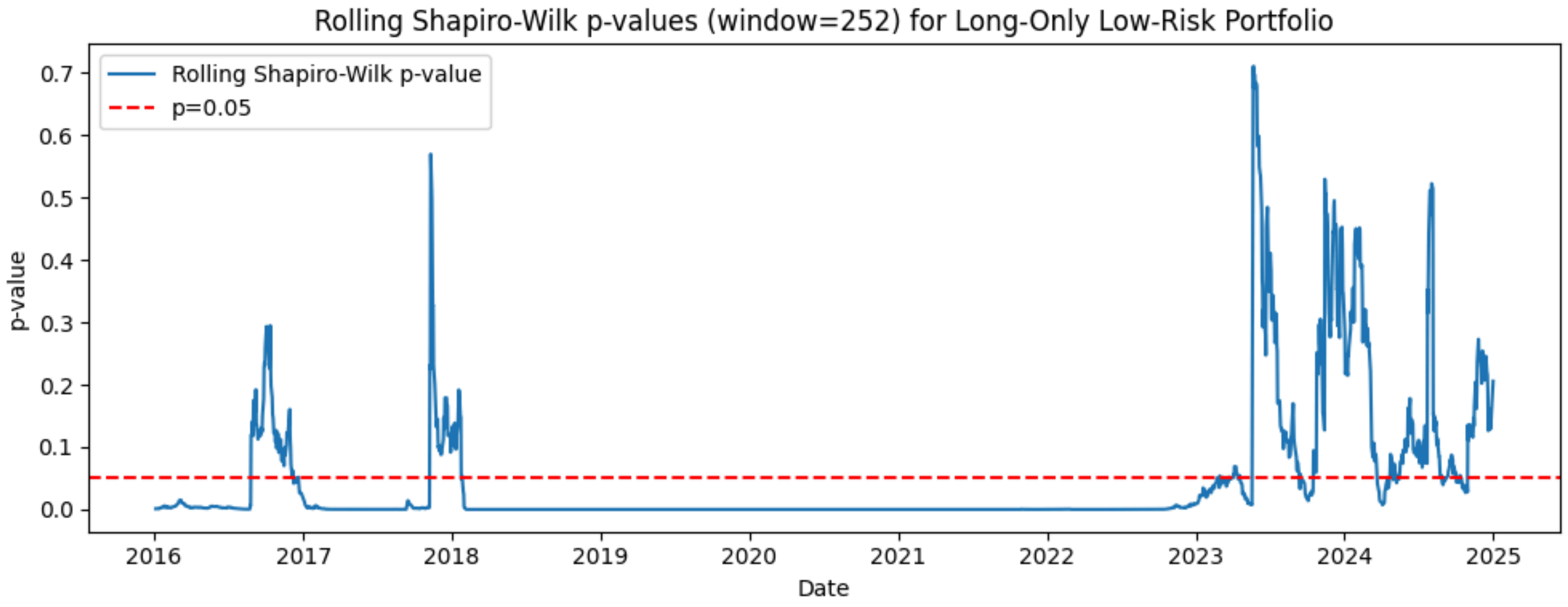


Normality statistics for the given period:

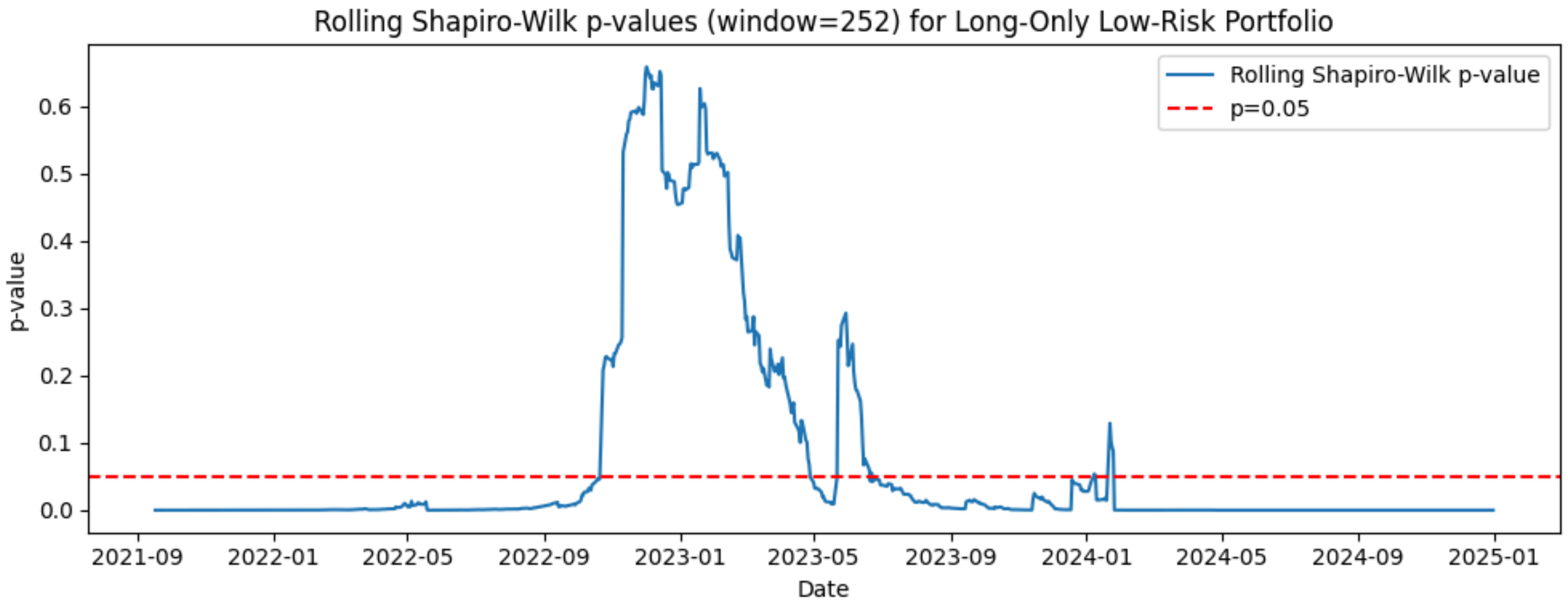
```
Normality tests for equal-weighted portfolio from 2023-01-04 to 2023-02-28:
  Shapiro-Wilk: stat=0.9720, p=0.4473
  Anderson-Darling: stat=0.2662
  D'Agostino-Pearson: stat=3.6089, p=0.1646
  Kolmogorov-Smirnov: stat=0.0904, p=0.8877
```

Normality of portfolios created in Project 1

Long-only low risk portfolio from basket-1



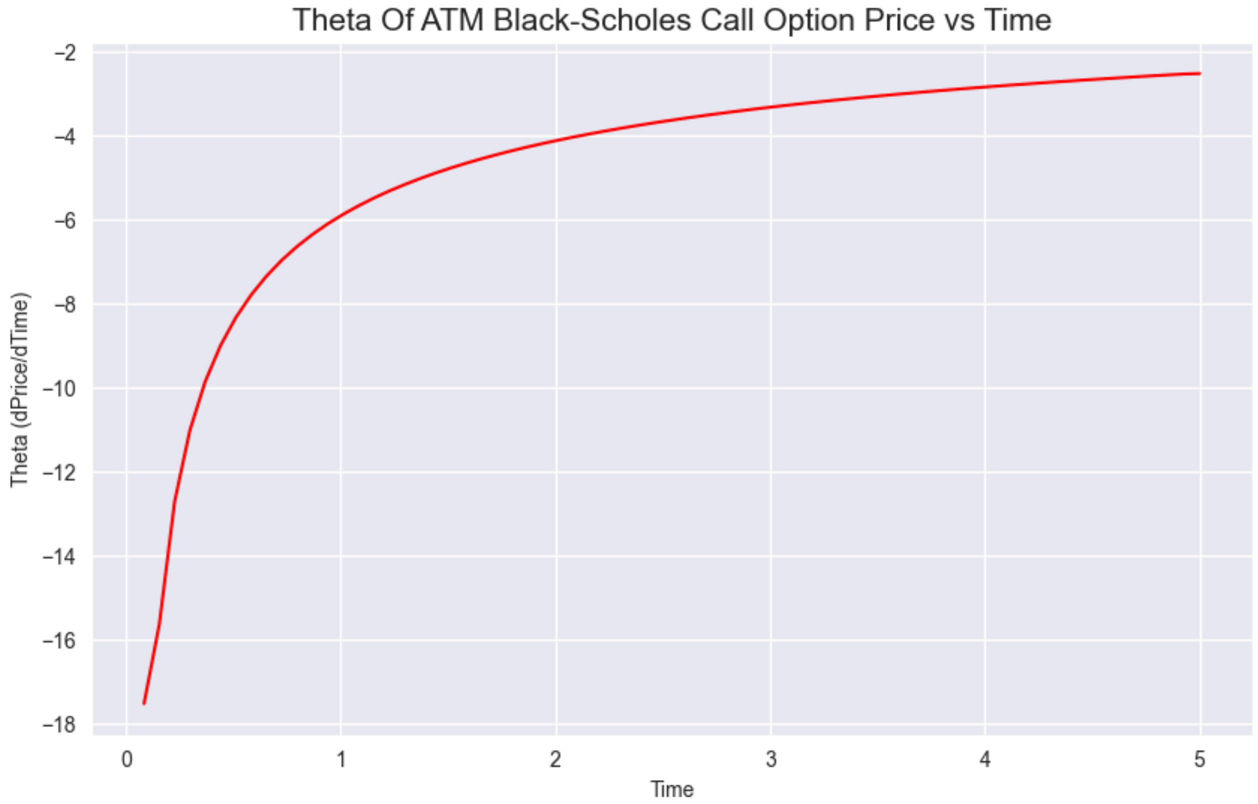
Long-only low-risk portfolio from basket-2



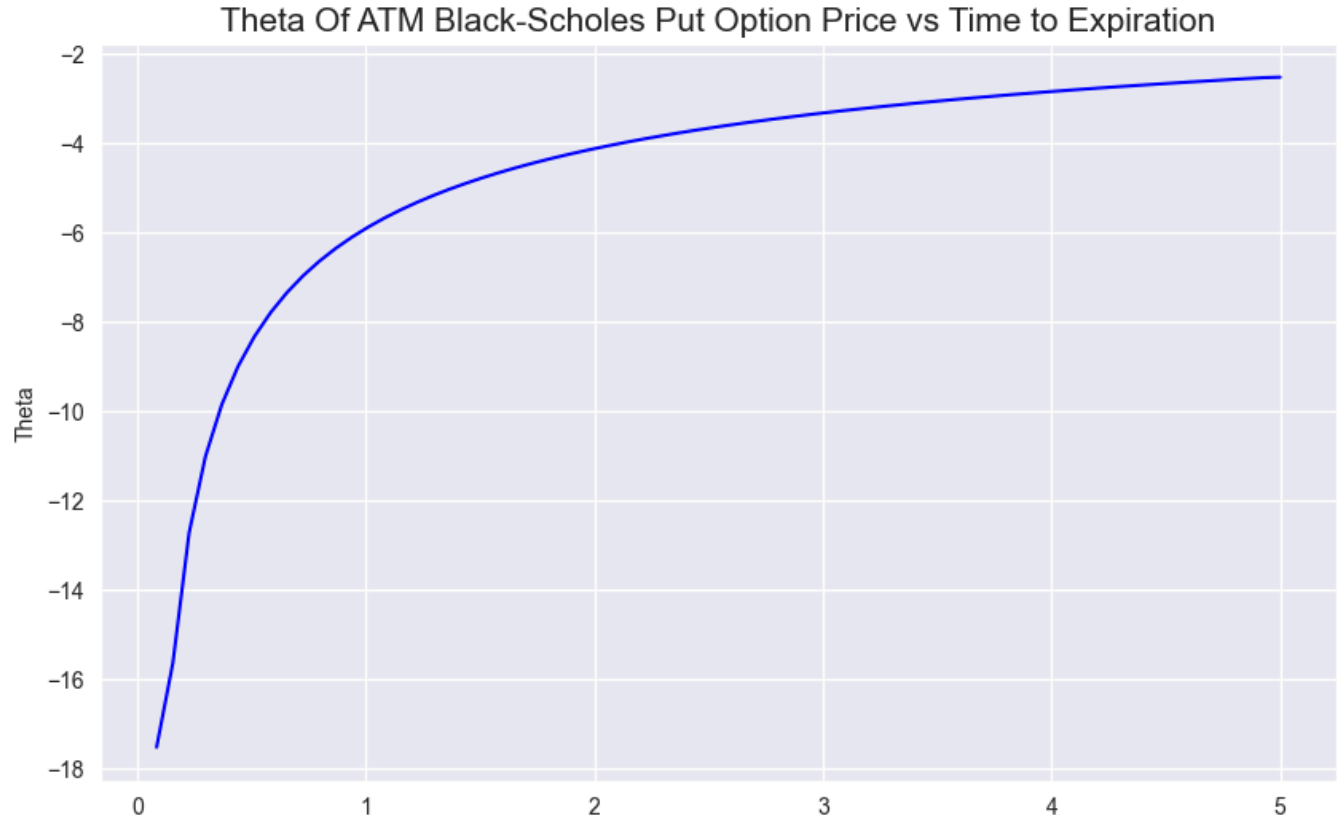
Project-3
BS as function of time

Black Scholes Δ and Θ as a function of time to expiry

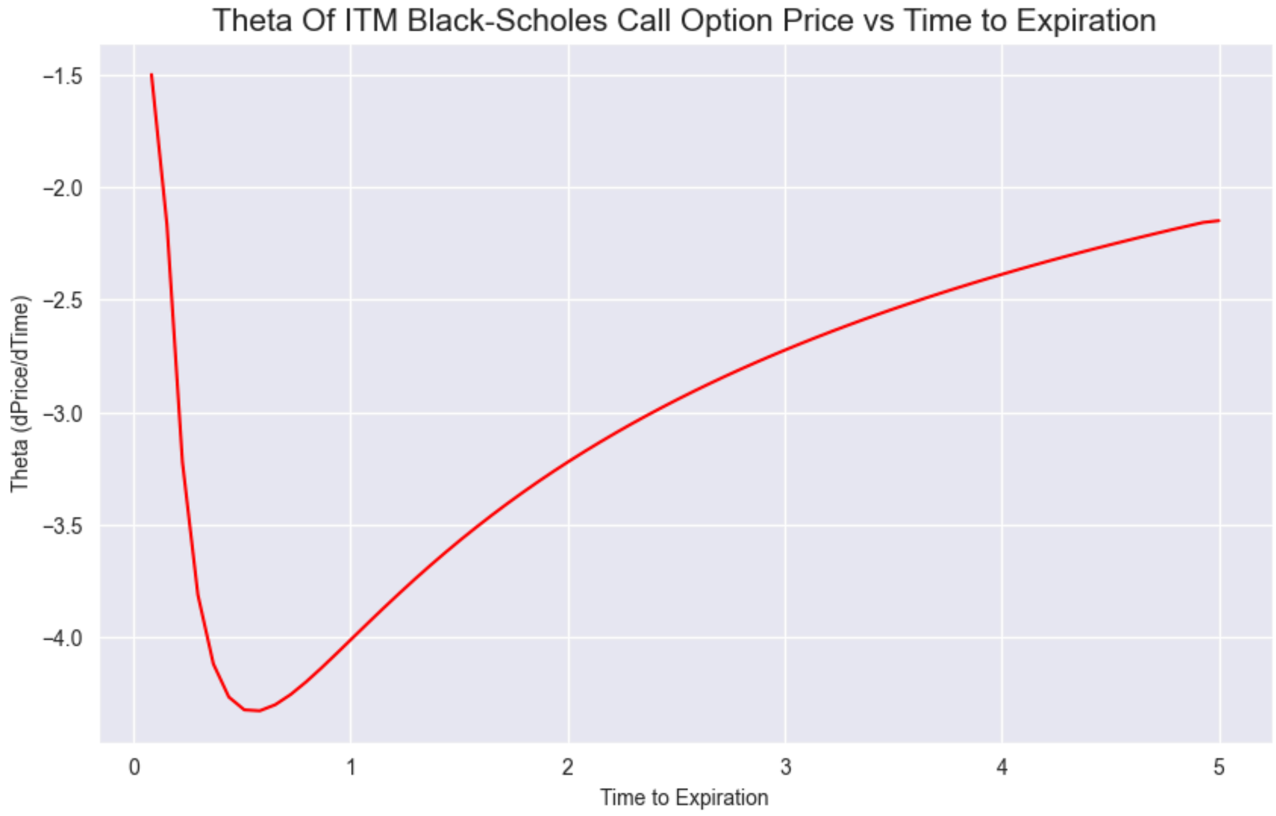
Variation of Black-Scholes Θ with time



$S_0 = 100\ ,\ K = 100\ ,\ \sigma = 0.3$

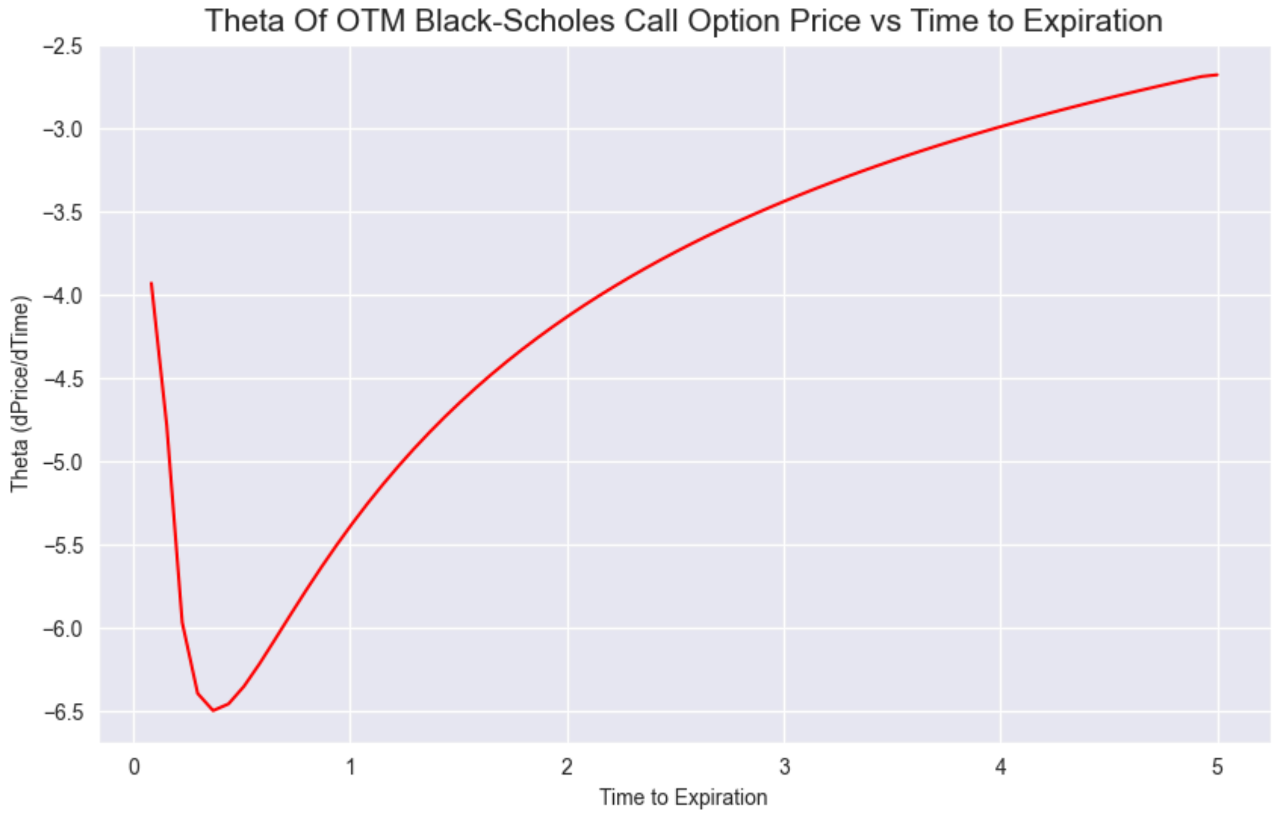


$S_0 = 100\ ,\ K = 100\ ,\ \sigma = 0.3$

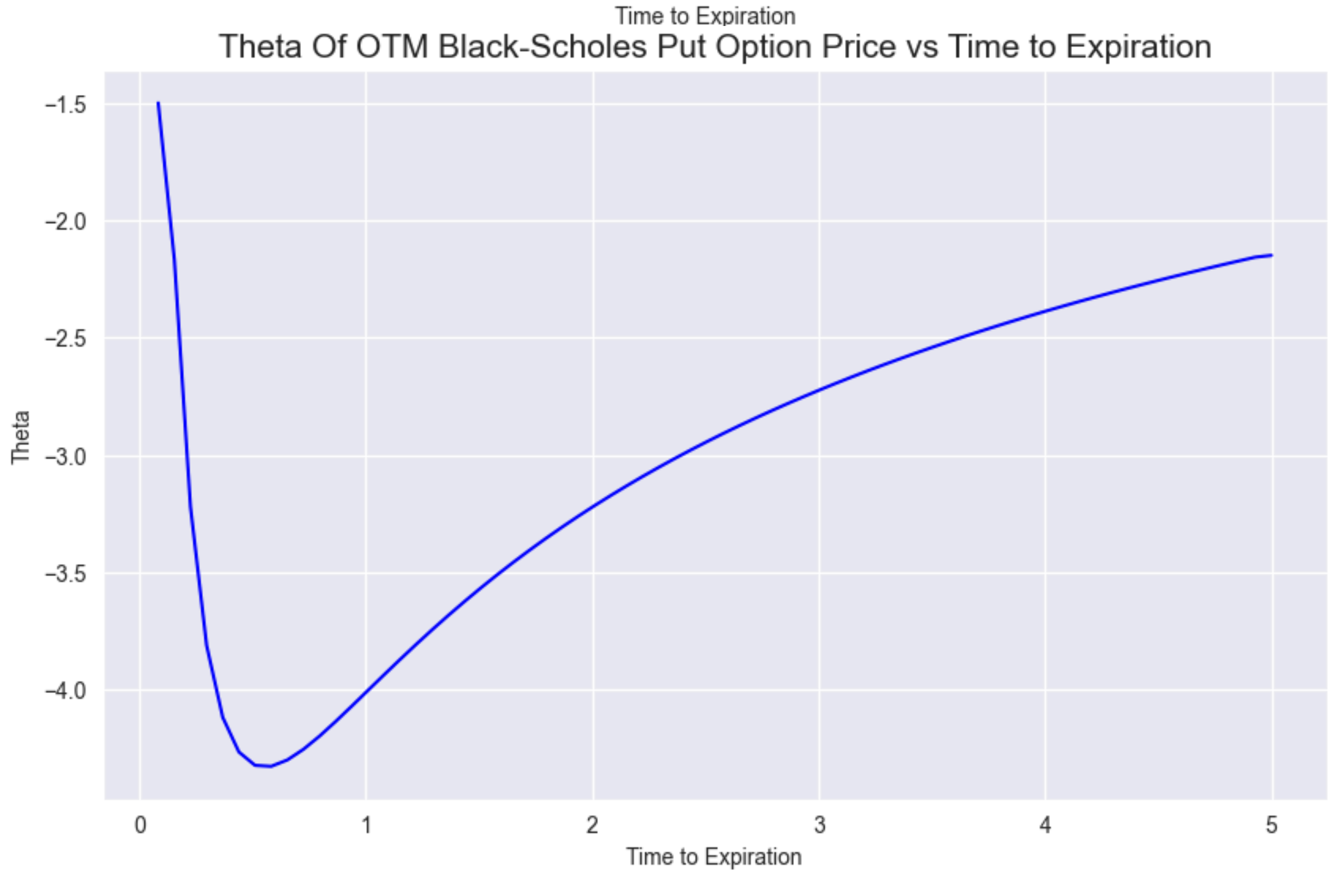


$S_0 = 100\ ,\ K = 80\ ,\ \sigma = 0.3$

$S_0 = 100\ ,\ K = 120\ ,\ \sigma = 0.3$



$S_0 = 100\ ,\ K = 120\ ,\ \sigma = 0.3$



$S_0 = 100\ ,\ K = 80\ ,\ \sigma = 0.3$

Variation of Black-Scholes Δ with time

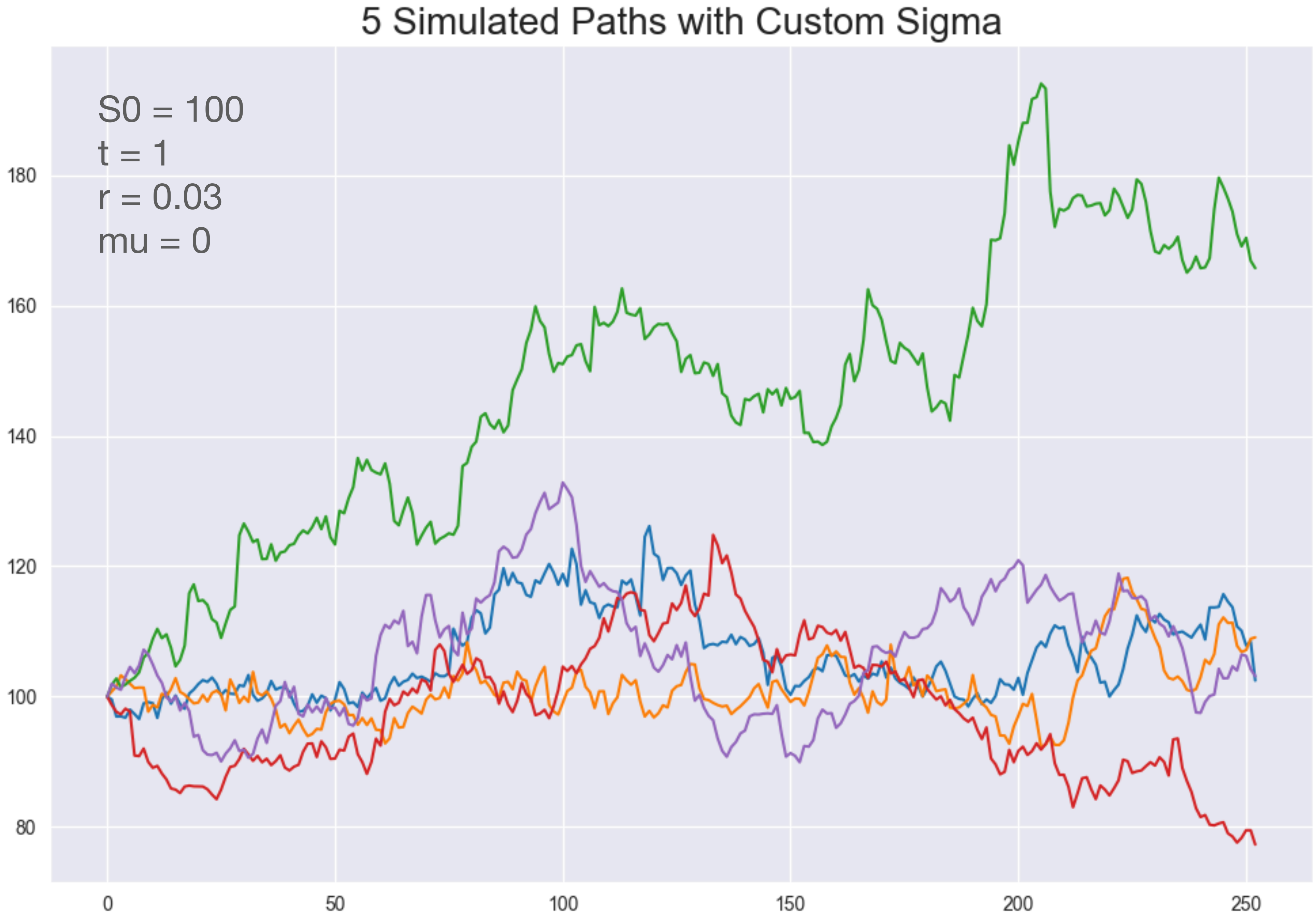
Black-Scholes Call

Black-Scholes Put

Project-4
Hedging in stochastic volatility

MC pricing of a European Call Option on a custom stock

We consider a custom stock model



$$S_T = S_0 \exp\left(\sum_{k=1}^{T} \Delta \log S_k\right)$$

where:

$$\Delta \ln S_j = \left(\mu + r - \frac{1}{2}\sigma_j^2\right)\Delta t + \sigma_j \sqrt{\Delta t} \cdot Z_j$$

$Z_i \sim \mathcal{N}(0,1)$

$\sigma_{i,j} \sim \text{Discrete}(\{0.2, 0.3, 0.45\}, \{0.5, 0.3, 0.2\})$

$\Delta t = T/N_{\text{steps}} \quad (N_{\text{steps}}: 252)$

The volatility of the stock is drawn from the above distribution

Value of the European call at $t = 0$: $V = e^{-rT}\mathbb{E}_{\mathbb{Q}}[\max(S_T - K)]$

MC Result for K=100, and 100000 simulations without Δ hedging: 13.12 $\pm$ 0.06

Δ Hedging

To protect from huge losses, the seller maintains a portfolio of options and underlying stock that need to be rebalanced regularly

In an n-step hedging process

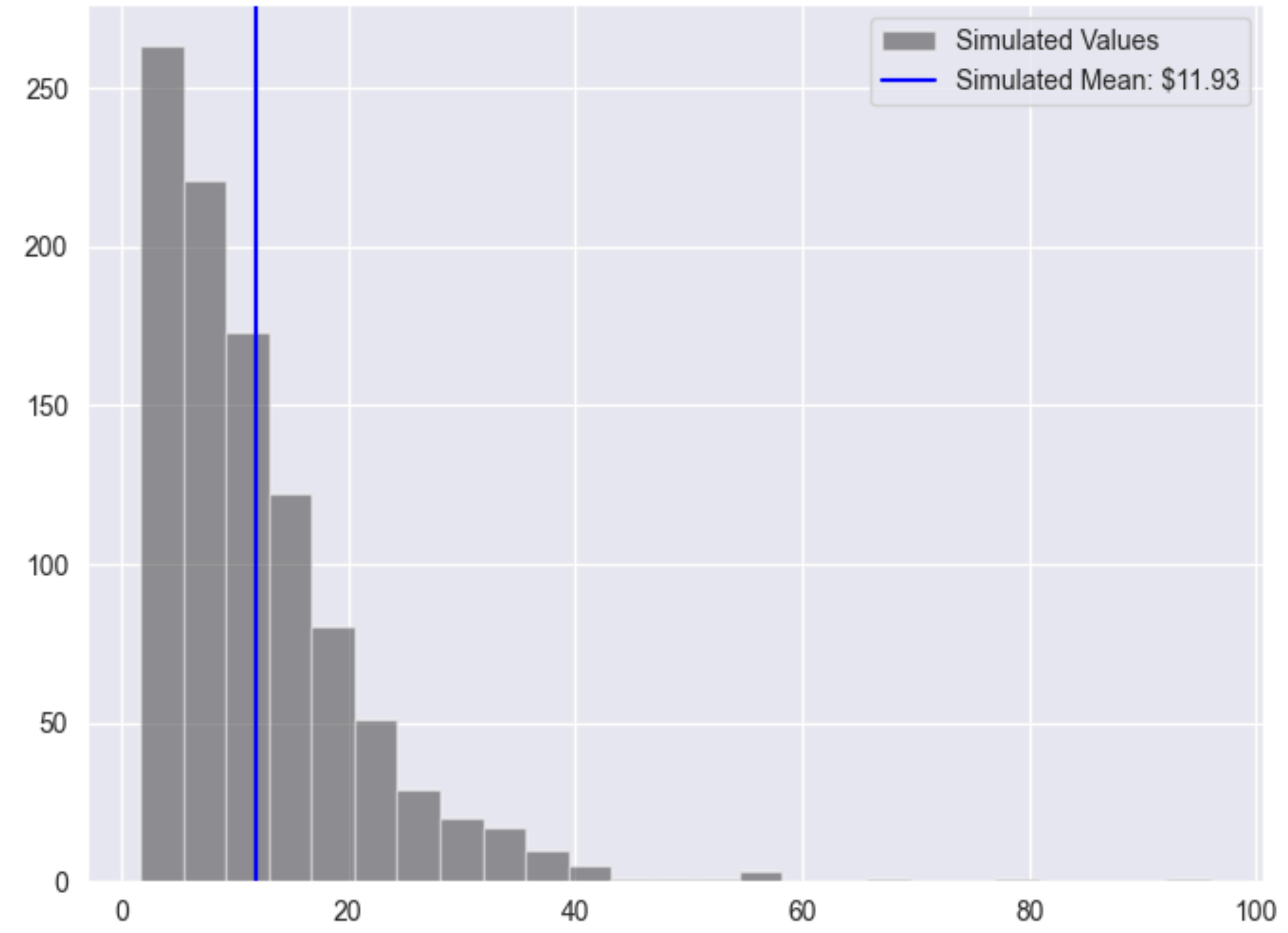
$$P = \text{premium} - e^{-rT}C_T + \sum_{i=i}^{n} e^{-rt_i}\{S_{t_i} - e^{r(t_i - t_{i-1})}S_{t_{i-1}}\}\Delta_{t_{i-1}} \qquad \Delta_{t_i} = \left.\frac{\partial C}{\partial S}\right|_{t=t_i}$$

In models with non-constant volatility we compute Δ numerically

$$\Delta_{C_0} \approx \frac{C_0(S_0+\varepsilon) - C_0(S_0-\varepsilon)}{2\varepsilon}$$

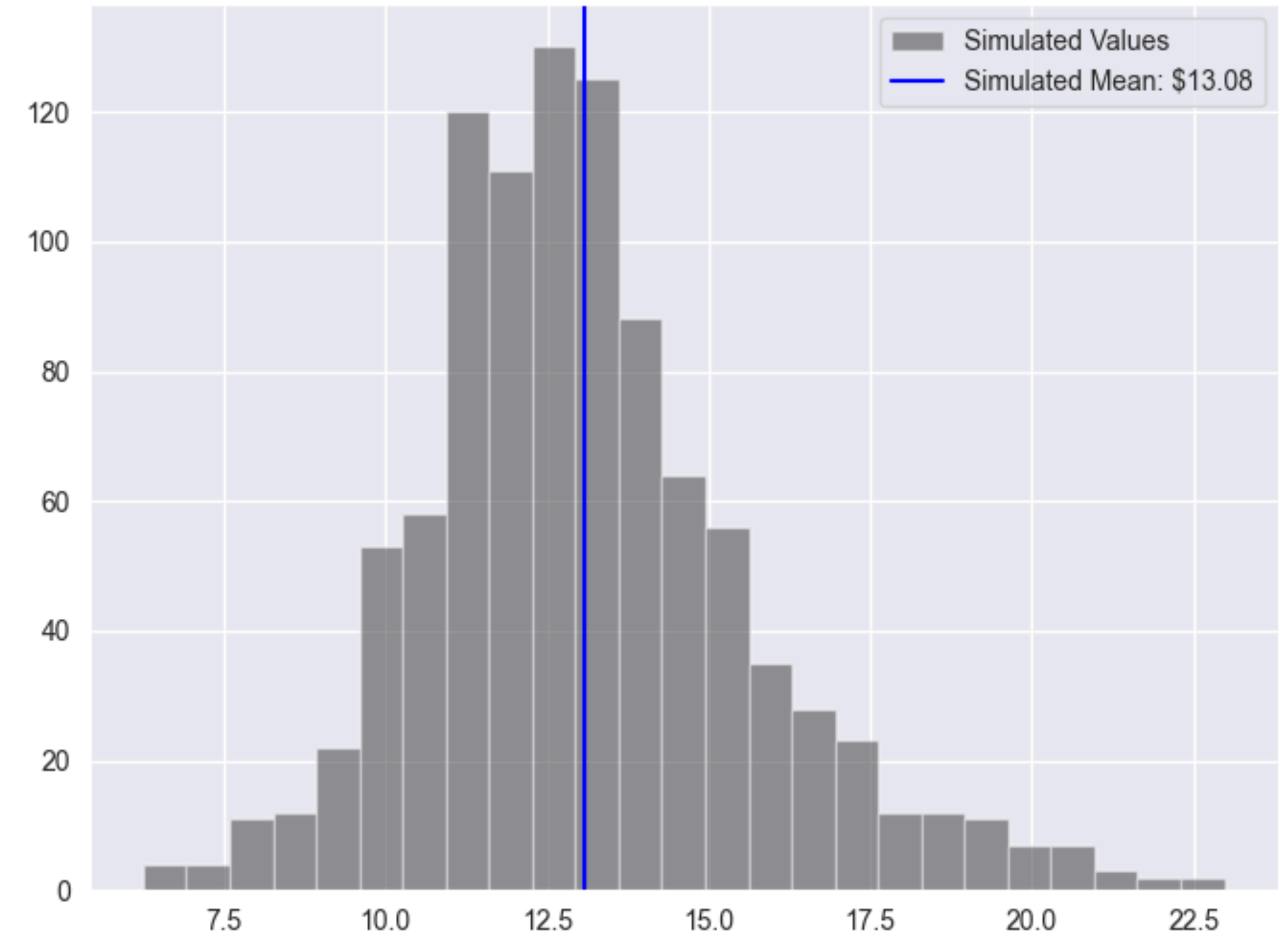
MC simulation of the European Call with increasing number of Δ hedges

Distribution of simulated Black-Scholes values with 1000 simulations and 1 control variants



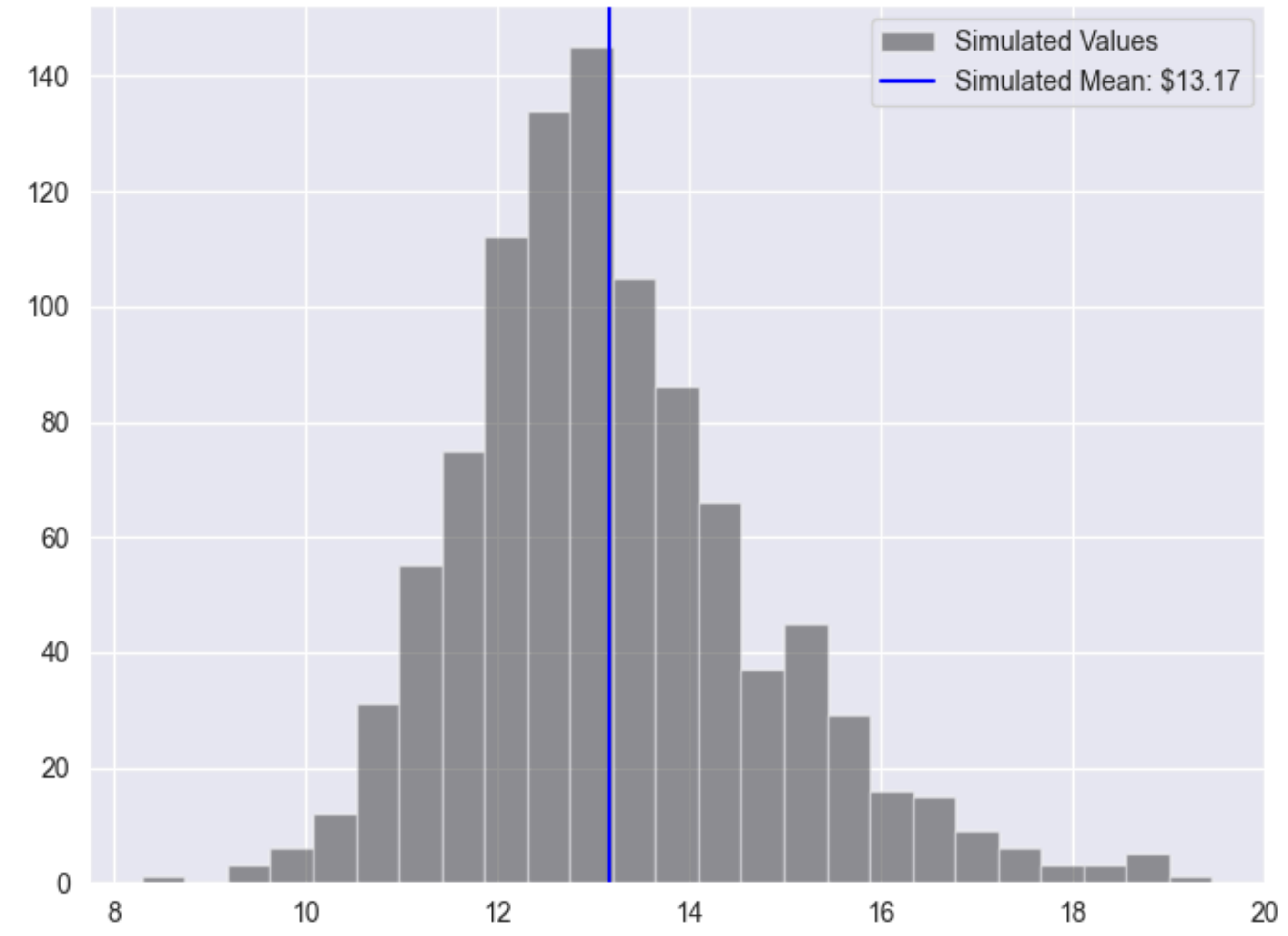

```
European Call Option Price (Monte Carlo, custom sigma, per contract): 11.93
Standard Error on Call Price (per contract): 0.3010
```

Distribution of simulated Black-Scholes values with 1000 simulations and 25 control variants



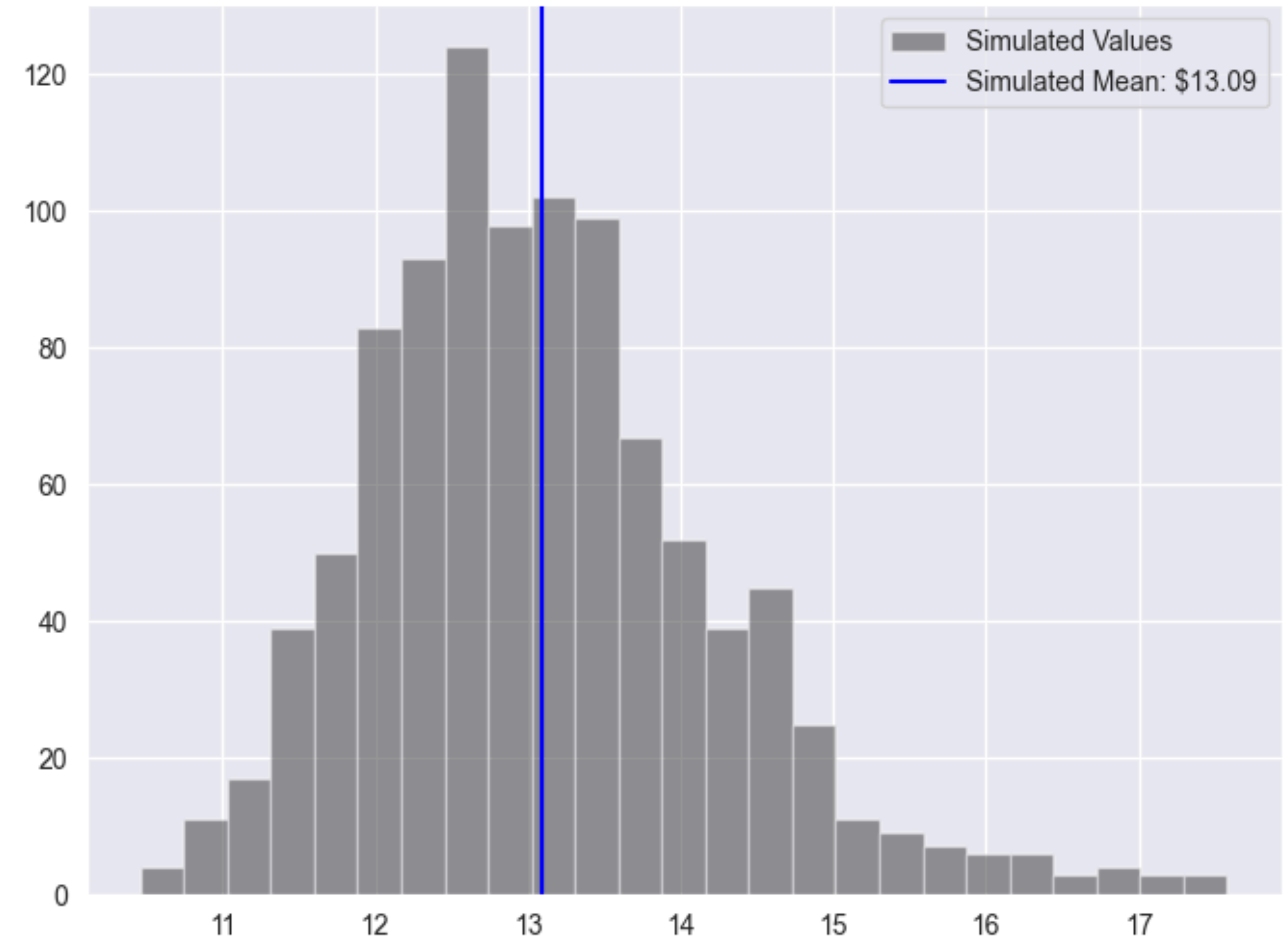

```
European Call Option Price (Monte Carlo, custom sigma, per contract): 13.08
Standard Error on Call Price (per contract): 0.0803
```

Distribution of simulated Black-Scholes values with 1000 simulations and 100 control variants




```
European Call Option Price (Monte Carlo, custom sigma, per contract): 13.17
Standard Error on Call Price (per contract): 0.0497
```

Distribution of simulated Black-Scholes values with 1000 simulations and 252 control variants




```
European Call Option Price (Monte Carlo, custom sigma, per contract): 13.09
Standard Error on Call Price (per contract): 0.0360
```

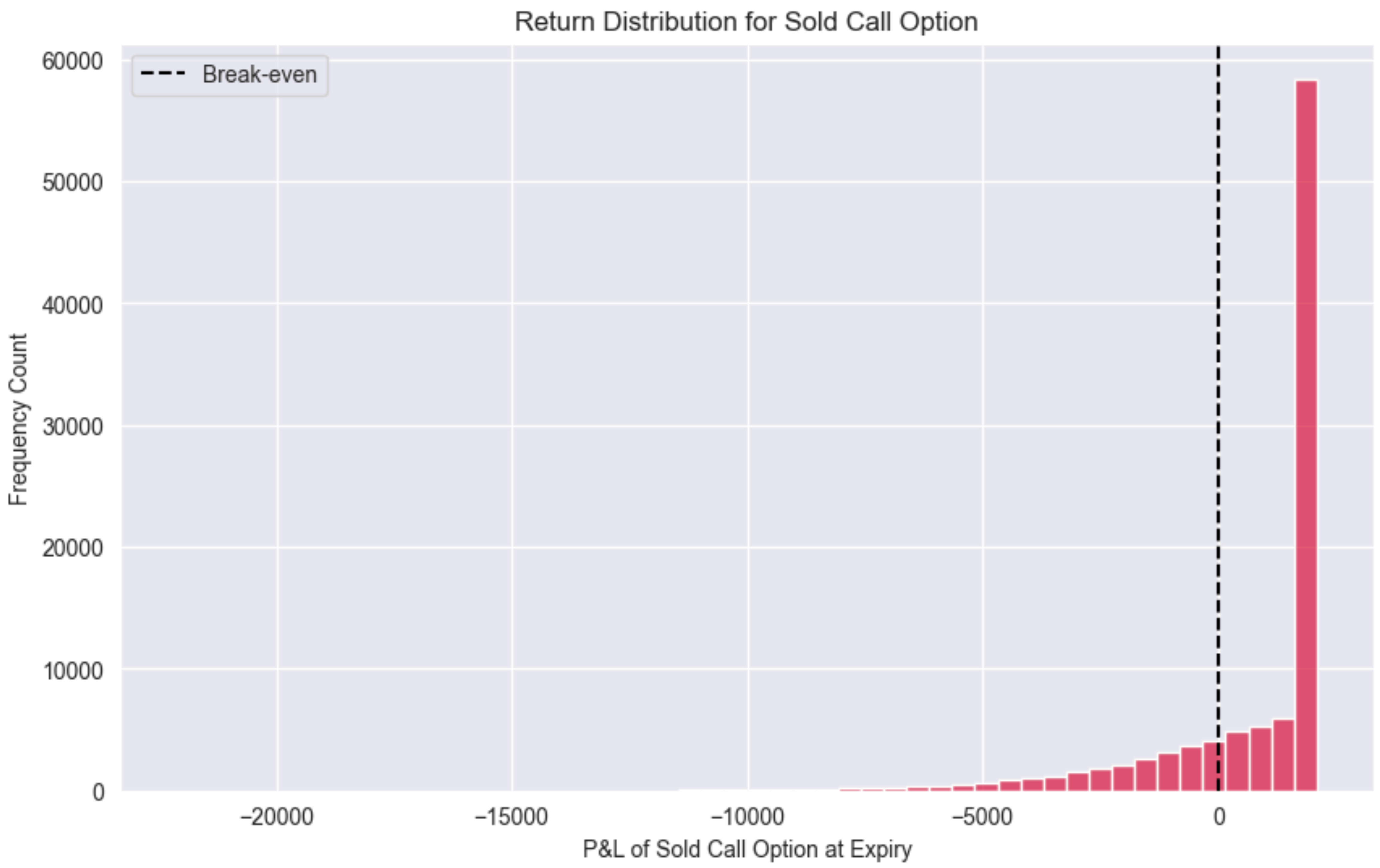
Standard error reduces as the number of hedges increase

PL distribution of seller of the European Call with increasing number of Δ hedges

Number of sold contracts: 100

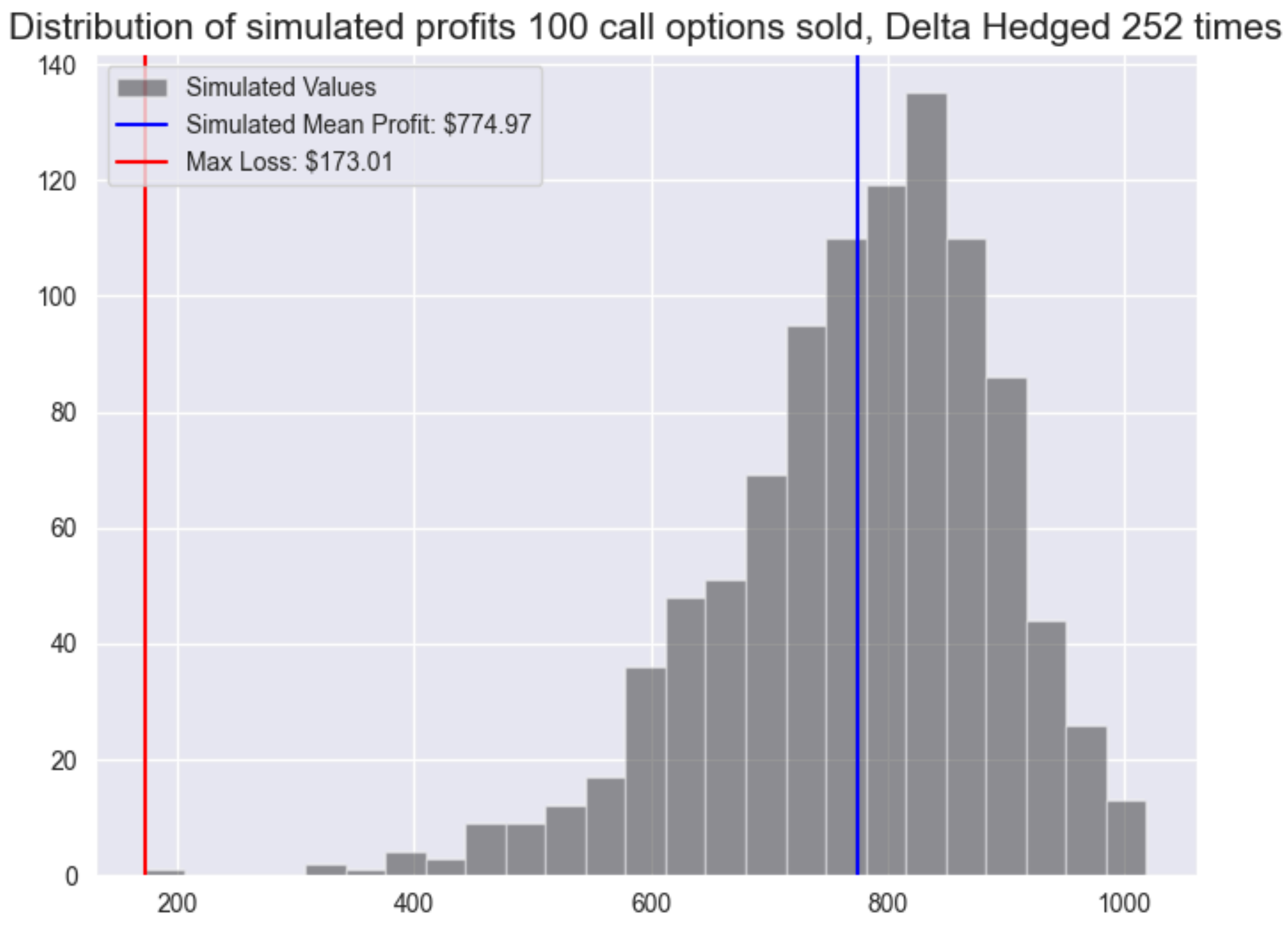
Premium per contract: `black_scholes_call(S0, K, t, r, bs_sigma=0.5)= 20.96`

P&L distribution without Δ hedging




```
Expected Profits (sold 100 calls): 784.49
Max Loss (sold 100 calls): -22104.63
Max Profit (sold 100 calls): 2096.15
```

P&L distribution with Δ hedging




```
Expected Profits (sold 100 calls): 774.97 ± 3.71
Max Loss (sold 100 calls): 173.01
Max Profit (sold 100 calls): 1018.18
```

MC pricing of a European Call Option in the GARCH(1,1) model

In GARCH(1,1) model, returns follow the following process:

$$r_t = \sigma_t \epsilon_t \,, \qquad \epsilon_t \sim \mathcal{N}(0,1)$$

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2$$

We start the simulation by setting $\sigma_0 = \omega/(1-\alpha-\beta)$ the unconditional variance if $\omega \neq 0$ and $\alpha + \beta < 1$ else use $\sigma_0 = 0.01$

The variance are then used to model the price process as follows:

$$S_T = S_0 \exp\left(\sum_{t=1}^{T} \Delta \log S_t\right)$$

where:

$$\Delta \ln S_t = \left(\mu + r - \frac{1}{2}\sigma_t^2\right) \Delta t + \sigma_t \sqrt{\Delta t} \cdot Z_t$$

That we use for modelling the option price: $V = e^{-rT}\mathbb{E}_{\mathbb{Q}}[\max(S_T - K)]$

Simulation

Model Parameters and Variables

$$\omega = 0, \ \alpha = 0.1, \ \beta = 0.88$$

$$S_0 = 100, \ K = 100, \ T = 1, \ r = 0.05, \ \mu = 0$$

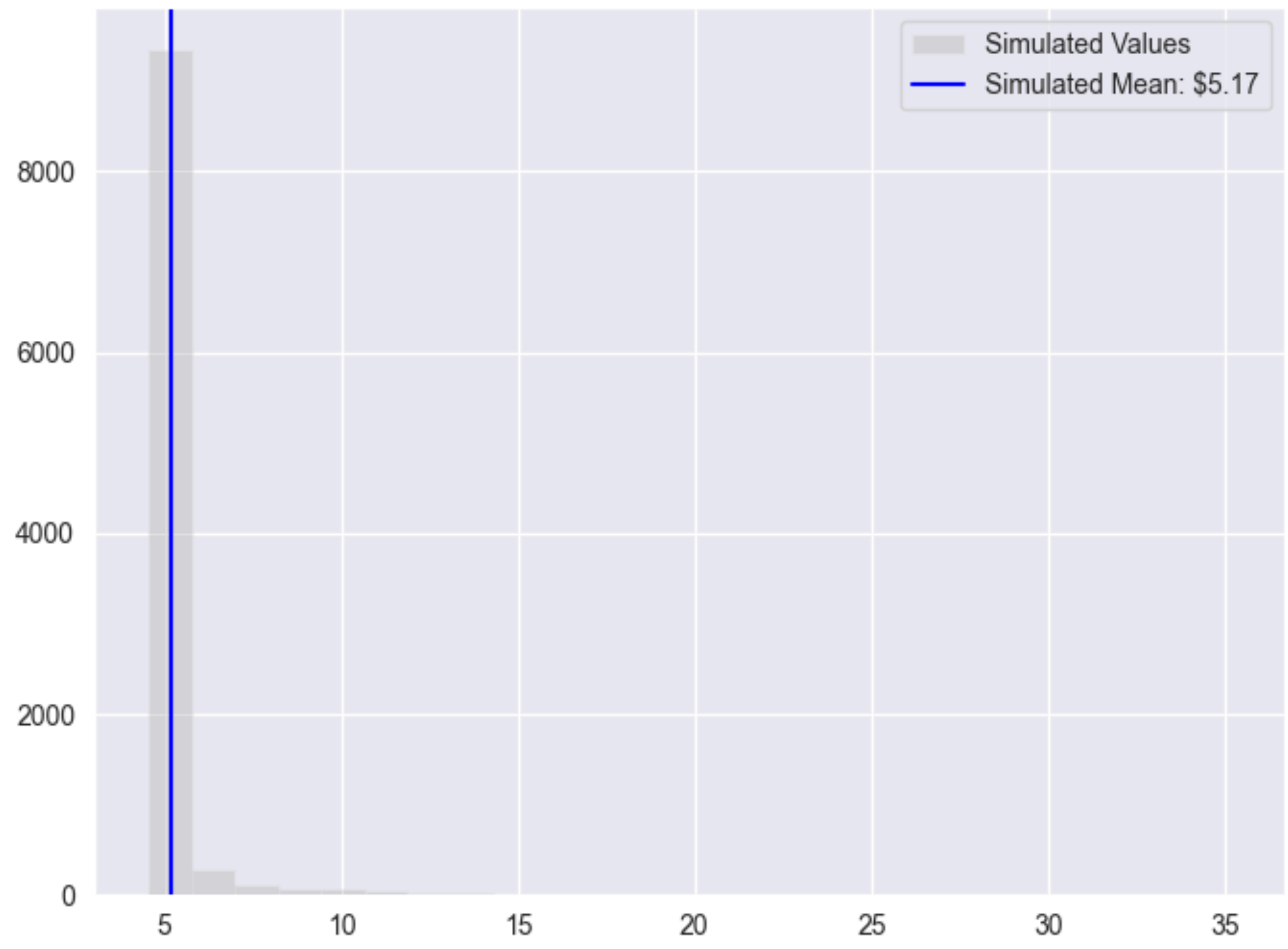
MC Result with 10,000 simulations without Δ hedging: \$5.16 ± \$0.0395

Close to Black-Scholes price of \$5.07 with average GARCH volatility 0.04

To numerically compute Δ we compute the empirical distribution of GARCH(1,1) volatilities

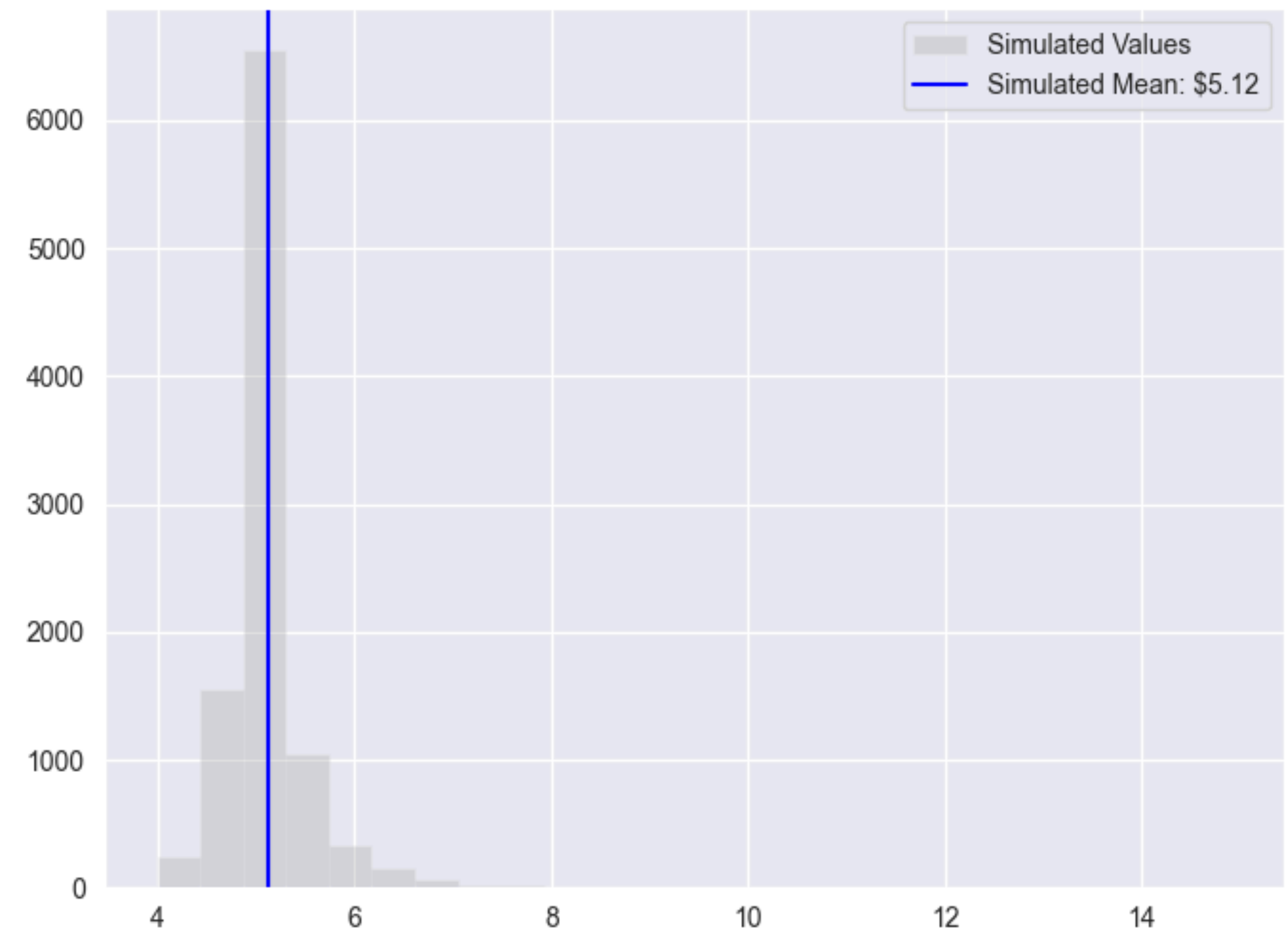
Δ hedging in GARCH(1,1) model

Distribution of simulated Black-Scholes values with 10000 simulations and 1 control variants



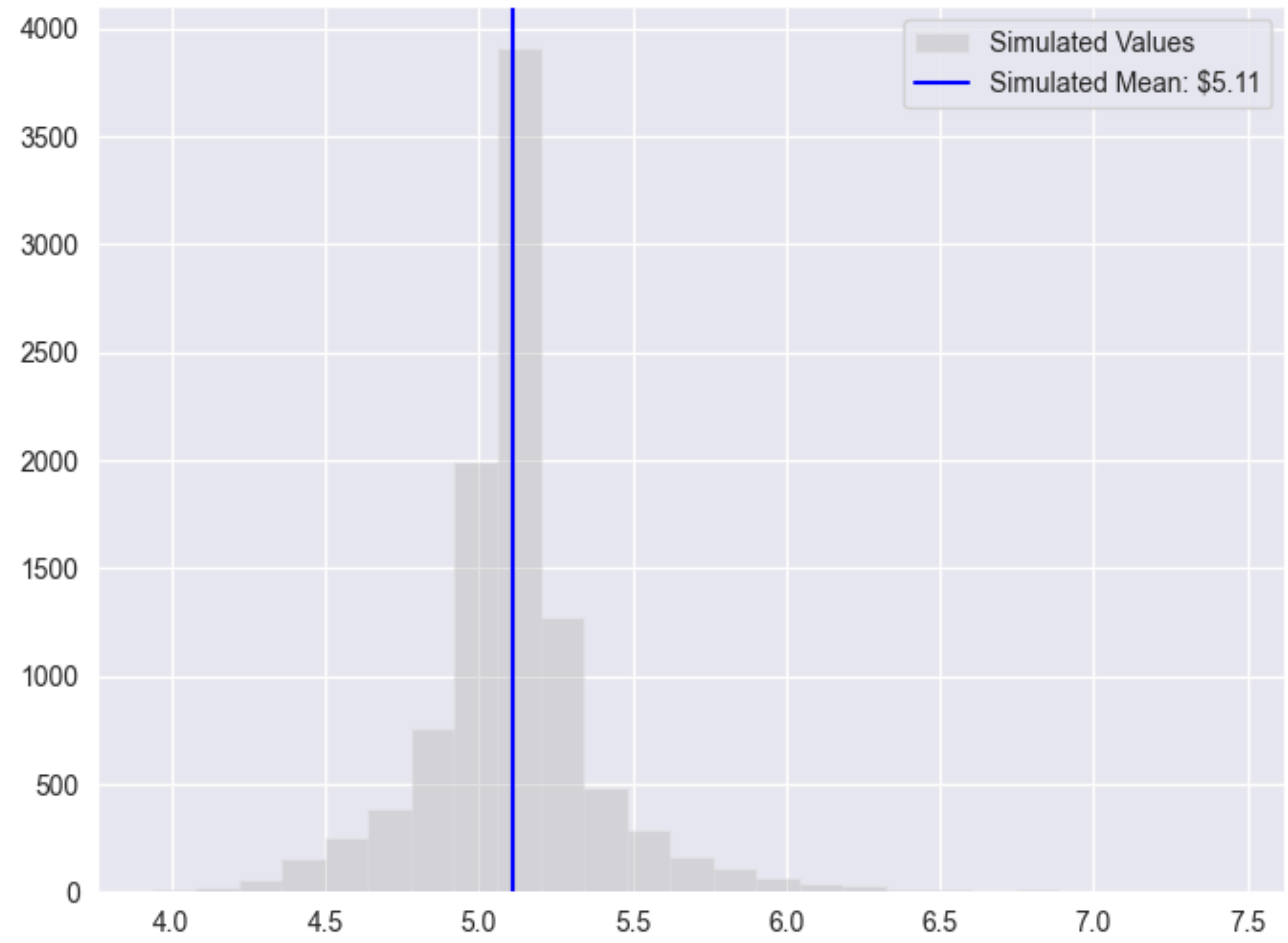

```
European Call Option Price (Monte Carlo, custom sigma, per contract): 5.17
Standard Error on Call Price (per contract): 0.0431
```

Distribution of simulated Black-Scholes values with 10000 simulations and 25 control variants



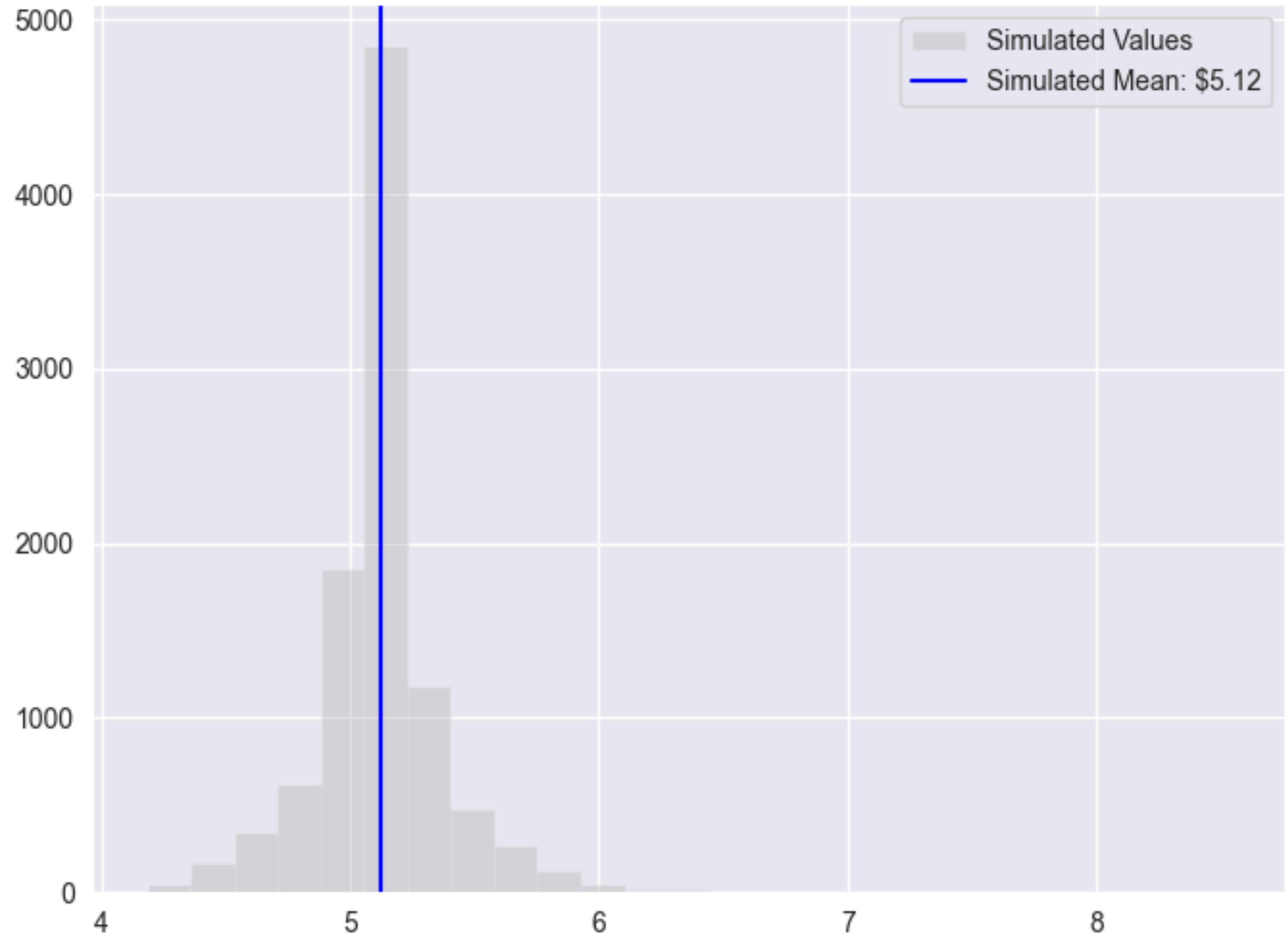

```
European Call Option Price (Monte Carlo, custom sigma, per contract): 5.12
Standard Error on Call Price (per contract): 0.0144
```

Distribution of simulated Black-Scholes values with 10000 simulations and 100 control variants




```
European Call Option Price (Monte Carlo, custom sigma, per contract): 5.11
Standard Error on Call Price (per contract): 0.0089
```

Distribution of simulated Black-Scholes values with 10000 simulations and 252 control variants



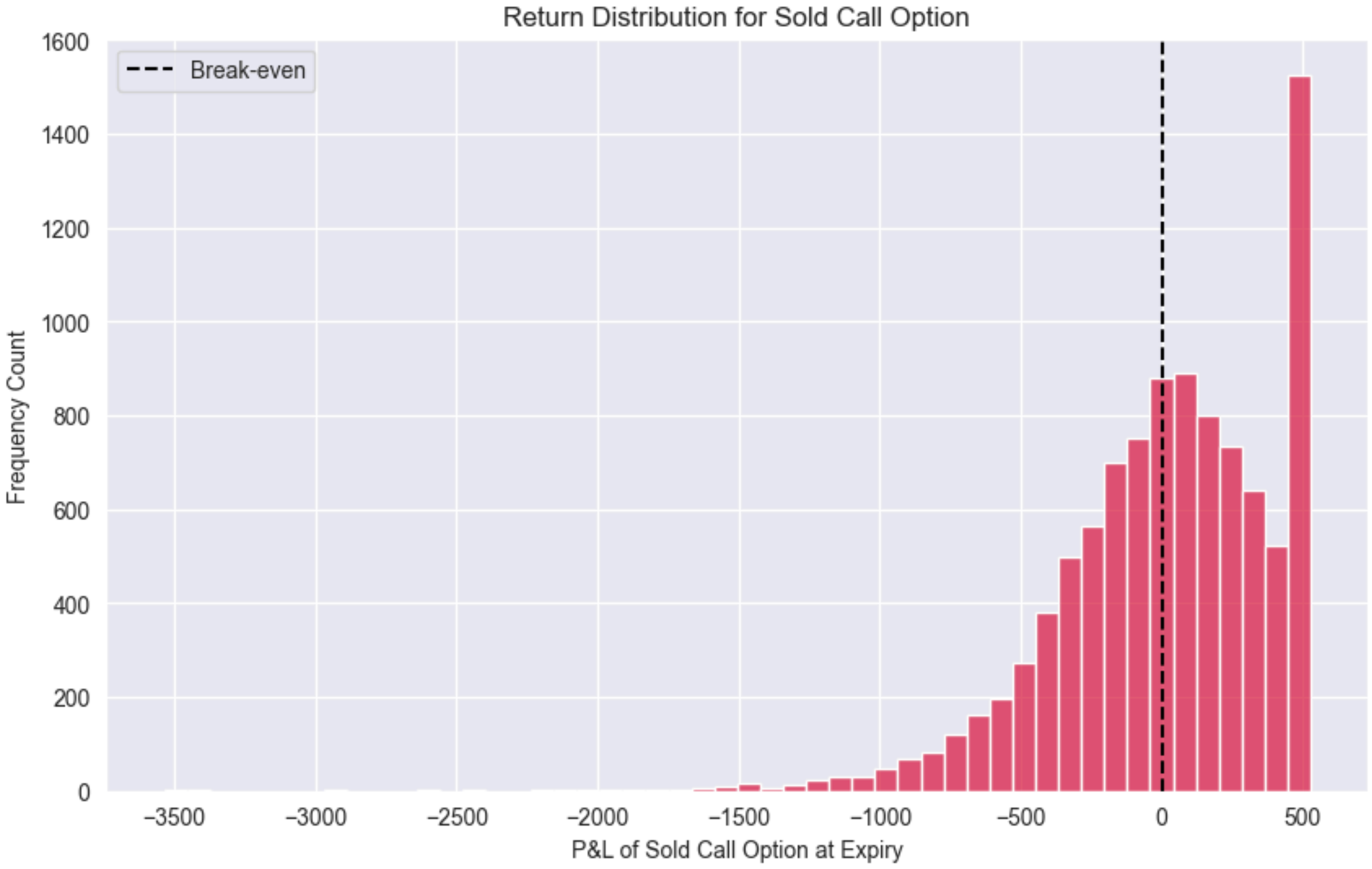

```
European Call Option Price (Monte Carlo, custom sigma, per contract): 5.12
Standard Error on Call Price (per contract): 0.0080
```

PL distribution of seller of the European Call in GARCH(1,1)

Number of sold contracts: 100

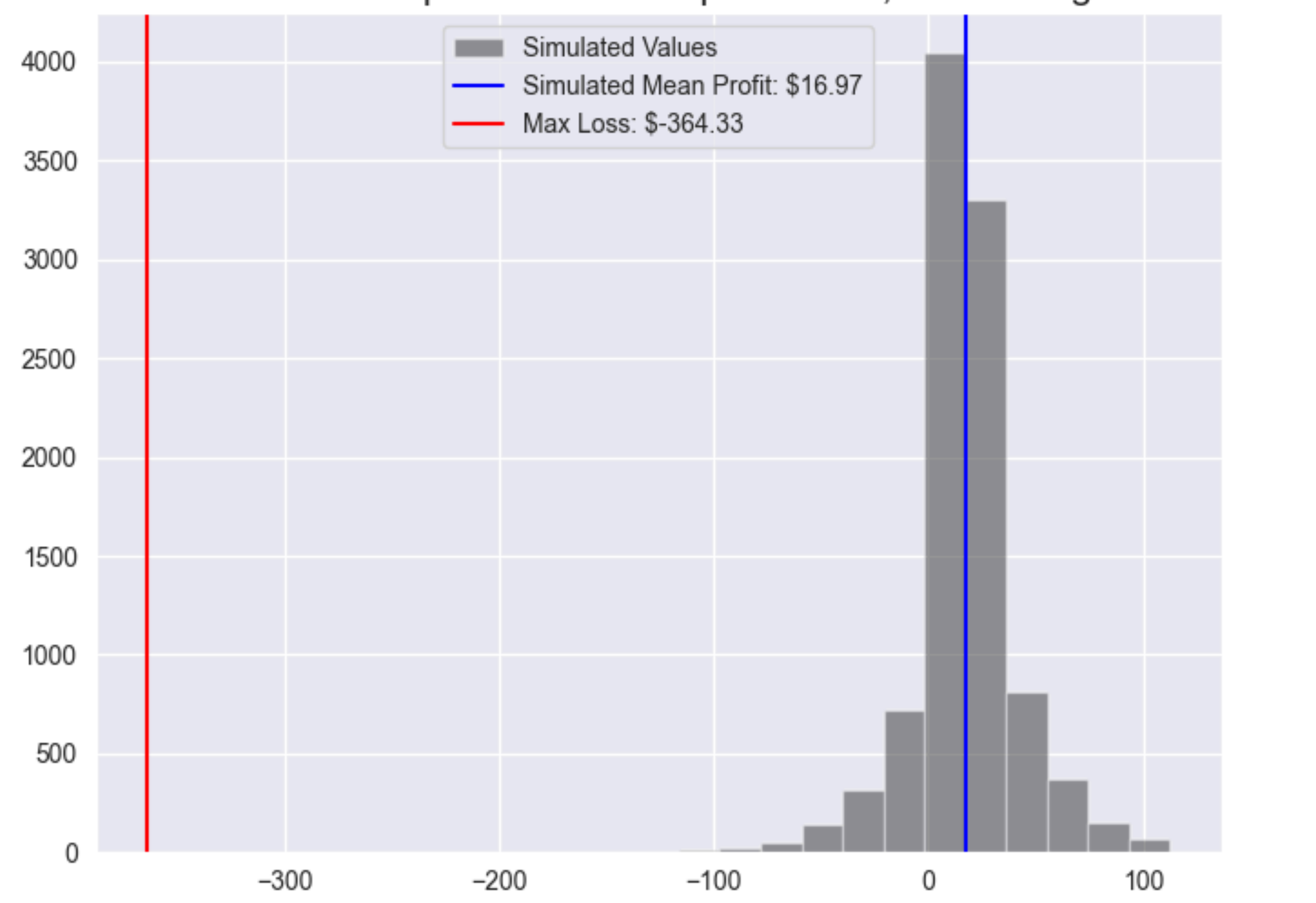
Premium per contract: `black_scholes_price(S0, K, T, r, avg_volatility+0.01)= 5.28`

P&L distribution without Δ hedging




```
Expected Profits (sold 100 calls): 12.02
Max Loss (sold 100 calls): -3538.80
Max Profit (sold 100 calls): 527.92
```

P&L distribution with Δ hedging




```
Expected Profits (sold 100 calls): 16.97 ± 0.25
Max Loss (sold 100 calls): -364.33
Max Profit (sold 100 calls): 112.47
```

Acknowledgements

We'd like to thanks Thomas Polstra, Amalya Lehmann
and Roman Holowinsky @ the Erdos Institute