Portfolio Optimisation with Linear Regression

Nakarin Kungvalchokchai
ID:2400952

CF969-7-SP/ZU: Machine Learning for Finance
Dr Panagiotis Kanellopoulos

Portfolio Optimisation with Linear Regression

1. Introduction

This report presents a portfolio optimisation analysis using the Capital Asset Pricing Model (CAPM), which explains the relationship between a security's expected return and its risk (CFI Team, n.d.) and quadratic programming. The objective is to determine optimal portfolio weights that minimize risk for a given level of expected return and to analyze the efficient frontier that illustrates the risk-return tradeoff.

2. Methodology

2.1. Data Collection

Historical stock price data was collected for 11 major stocks, balancing growth and defensive stocks in a well-diversified portfolio across sectors, over a 5-year period using the yfinance library. The stocks selected were:

AAPL – Apple Inc.: Technology
AMT – American Tower Corporation: Real Estate
AMZN – Amazon.com, Inc.: E-commerce & Cloud Computing
GOOGL – Alphabet Inc. (Class A): Technology
JPM – JPMorgan Chase & Co.: Financials
LIN – Linde plc: Materials
NEE – NextEra Energy, Inc.: Utilities (Renewable Energy)
PG – Procter & Gamble Co.: Consumer Staples
RTX – RTX Corporation (Raytheon Technologies): Industrials & Defense
UNH – UnitedHealth Group Incorporated: Healthcare
XOM – Exxon Mobil Corporation: Energy

The S&P 500 index (^GSPC) was used as the market benchmark. Daily returns were calculated from the closing prices, and the risk-free rate was set to: $R_{\text{f daily}} = \frac{R_{\text{f annual}}}{252} = \frac{0.02}{252}$ (where 252 is trading days per year).

2.2. CAPM Linear Regression

For each asset, linear regression was performed to estimate the CAPM parameters:

$$R_i - R_f = \alpha_i + \beta_i(R_m - R_f) + \varepsilon_i$$

where:

- R_i is the return of asset i
- R_f is the risk-free rate
- R_m is the market return
- α_i and β_i are the regression coefficients
- ε_i is the error term

The expected return for each asset was calculated using: $\mu_i = R_f + \beta_i(E[R_m] - R_f)$,where E[Rm] was the average market return over the period.

The idiosyncratic risk was estimated as the variance of the regression residuals:

Idiosyncratic Variance = $Var(\varepsilon_i) = Var(Y_i - (\alpha_i + \beta_i X))$,where $Y_i = R_i - R_f$

The regression was implemented using StatsModels library, with:

- Independent variable (X): Daily market excess returns $(R_m - R_f)$
- Dependent variable (Y): Daily asset excess returns $(R_i - R_f)$

2.3. Covariance Matrix Construction

The covariance matrix Σ was constructed using:

- Systematic risk component: $\sigma_{ij} = \beta_i \beta_j \sigma_m^2 \ (for\ i \neq j)$
- Diagonal elements with idiosyncratic risk: $\sigma ii = \beta_i^2 \sigma^2_m + Var(\varepsilon i)$

2.4. Portfolio Optimisation

The portfolio optimisation problem was formulated as a quadratic programming (QP) problem:

Minimise $\frac{1}{2} w^T \Sigma w$

Subject to:
- $w^T \mu = \mu_P$ (target return constraint)
- $w^T 1 = 1$ (weights sum to 1)
- $w \geq 0$ (non-negative weights)

where: w is the vector of portfolio weights , μ is the vector of expected returns, μ_P is the target portfolio return

The CVXOPT library was used to solve this quadratic programming problem.

2.5. Efficient Frontier

The efficient frontier was computed by varying the target return μ_P across 20 equally spaced values between the minimum and maximum asset expected returns. For each target return, the optimisation problem was solved to find the minimum risk portfolio.

3. Results and Analysis

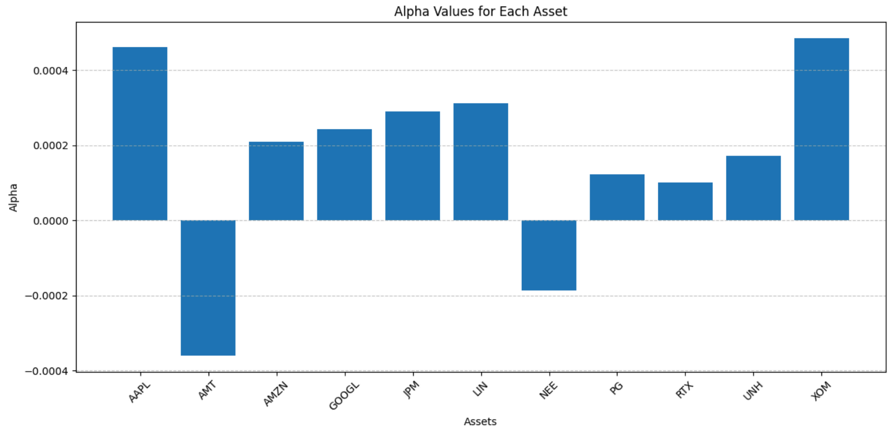
3.1. CAPM Regression Results

The table below presents the results of the CAPM regression for each asset:

Stock	Alpha (Daily)	p-value (α)	Beta	p-value (β)	R^2	Idiosyncratic Variance	Expected Return
AAPL	0.000462	0.186	1.150126	<0.001	0.603043	0.000154	0.000691
AMT	-0.000361	0.414	0.840071	<0.001	0.333505	0.000249	0.000526
AMZN	0.000210	0.772	1.096914	<0.001	0.417415	0.000297	0.000663
GOOGL	0.000243	0.470	1.142018	<0.001	0.546117	0.000192	0.000687
JPM	0.000290	0.376	1.075118	<0.001	0.492148	0.000211	0.000651
LIN	0.000312	0.235	0.902492	<0.001	0.551825	0.000117	0.000559
NEE	-0.000187	0.676	0.799187	<0.001	0.304761	0.000257	0.000504
PG	0.000123	0.688	0.526425	<0.001	0.283544	0.000124	0.000359
RTX	0.000100	0.775	0.932487	<0.001	0.359263	0.000274	0.000575
UNH	0.000172	0.725	0.793674	<0.001	0.320141	0.000236	0.000501
XOM	0.000486	0.229	0.811579	<0.001	0.253209	0.000343	0.000511

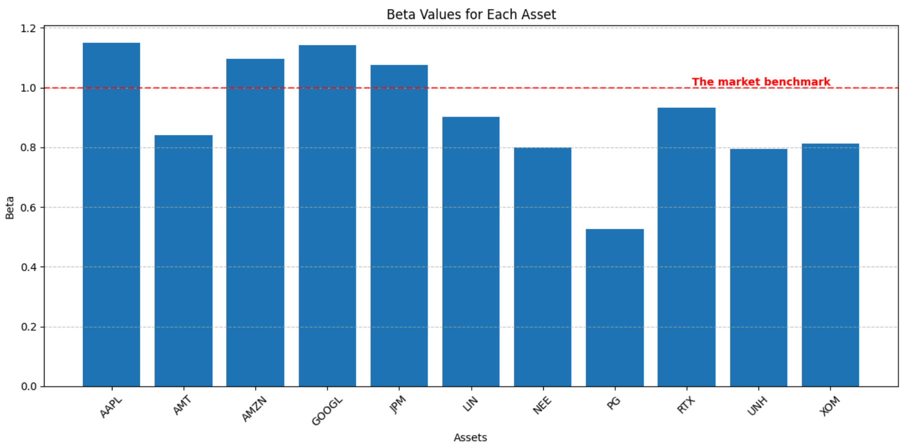
Interpretation of Alpha and Beta and Expected Return

- Alpha (α)



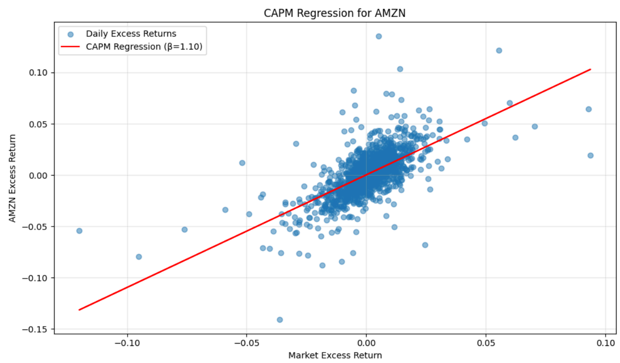
Alpha measures the excess return of a stock relative to its expected return, as estimated using the Capital Asset Pricing Model (CAPM), where risk is represented by beta. A positive alpha suggests that a stock outperforms expectations based on its beta, whereas a negative alpha indicates underperformance (CFI Team, n.d.). For example, stocks such as AAPL (0.000462) and XOM (0.000486) have positive alpha, which indicates that they deliver higher returns than expected given their beta values. On the other hand, stocks such as AMT (-0.000361) and NEE (-0.000187) have negative alpha, meaning their returns are lower than expected based on their market risk. However, since all alpha p-values exceed 0.05, we fail to reject the null hypothesis, suggesting that these excess returns are not statistically significant at the 5% significance level.

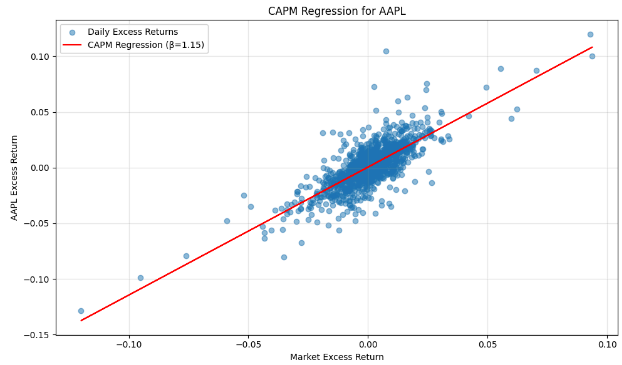
- Beta (β)



Beta quantifies a stock's risk by assessing how its price movements compare to overall market fluctuations. For example, a beta of 1.5 indicates that the stock is 50% more volatile than the market, while a beta of 1 suggests that the stock's expected return is in line with the market's average return. Conversely, a beta of -1 signifies a perfect inverse relationship with the market (CFI Team, n.d.). All stocks have statistically significant beta values ($p < 0.001$), indicating a strong statistical relationship between each stock's returns and the market's returns.

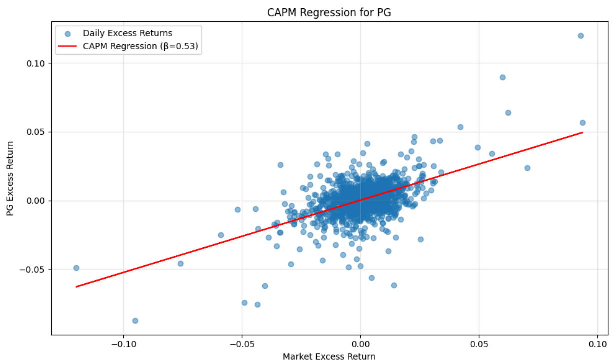
Higher Beta Assets:

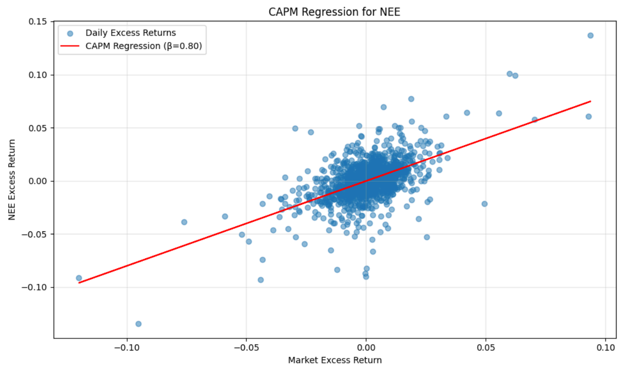




AAPL (Apple), AMZN (Amazon), GOOGL (Google), and JPM (JPMorgan) all have beta values above 1, indicating that they are more volatile than the market. As a result, these stocks have the potential to outperform or underperform the market, depending on market conditions.

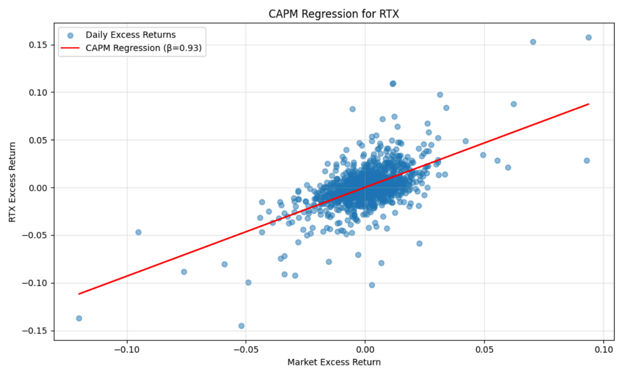
Lower Beta Assets:

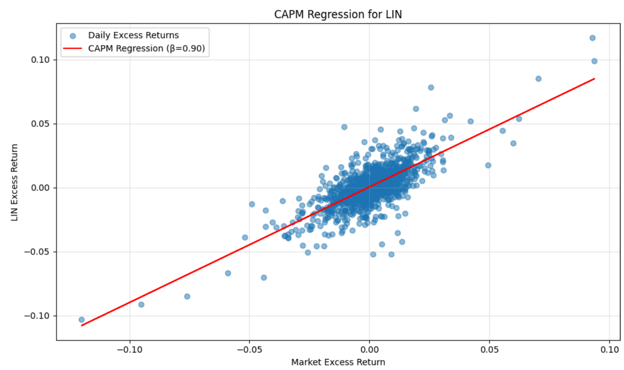




PG (Procter & Gamble) has the lowest beta (below 0.5), which classifies it as a defensive stock due to its low volatility and low sensitivity to market fluctuations. NEE (NextEra Energy) and UNH (UnitedHealth) both have beta values below 1, indicating lower volatility and relatively stable performance compared to the broader market.

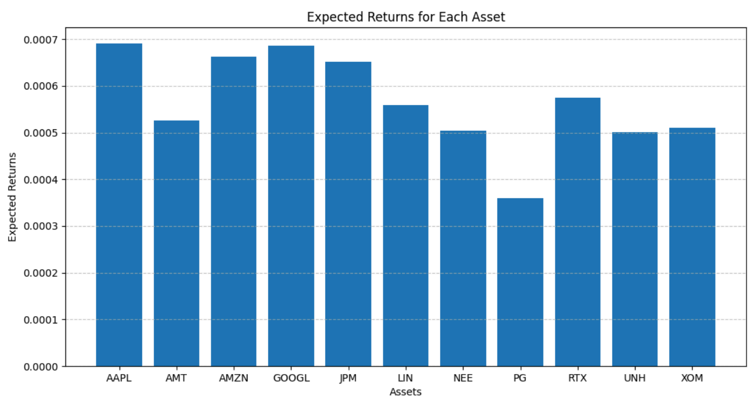
Moderate Beta Assets:





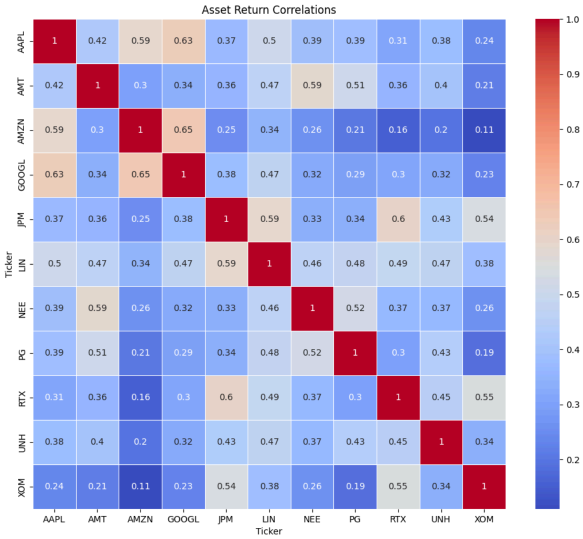
RTX (Raytheon) and LIN (Linde) have beta values slightly below 1, indicating they are slightly responsive to market changes but less volatile than high-beta stocks. These stocks provide a balanced risk-return profile.

- Expected Return



The expected return represents the forecasted return based on the analysis, with higher expected returns being more attractive for investors looking for growth opportunities. Stocks with higher expected returns, such as AAPL (0.000691) and GOOGL (0.000687), indicate strong growth potential. On the other hand, stocks such as PG (0.000359) and NEE (0.000504) have lower expected returns, implying more stable but less growth opportunities.

- Asset return correlations

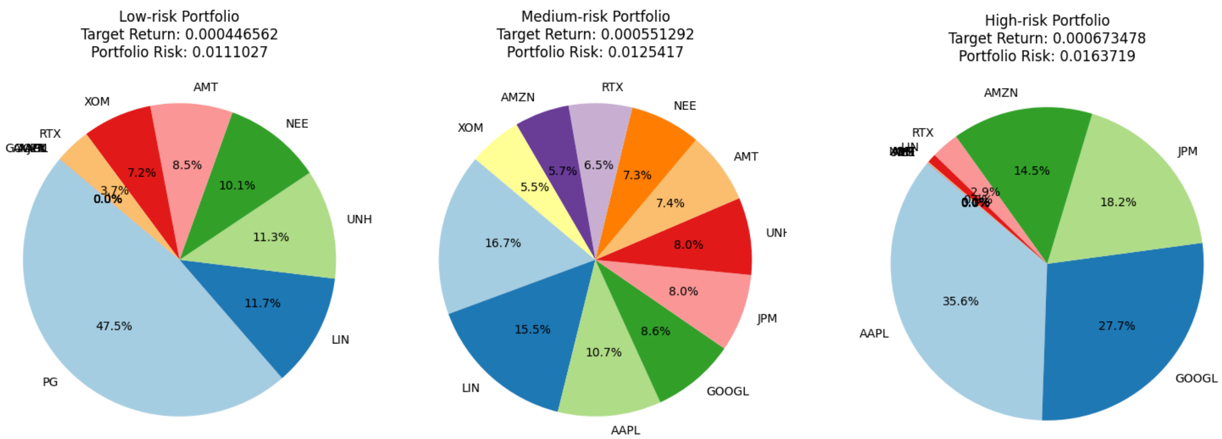


The asset return correlation matrix reveals significant relationships among the 11 assets. High correlations occur between technology companies (AAPL-GOOGL: 0.63, AMZN-GOOGL: 0.65) and between financial and industrial sectors (JPM-RTX: 0.60, JPM-LIN: 0.59). In contrast, minimal correlations appear between technology and energy sectors (AMZN-XOM: 0.11) and technology and aerospace/defense (AMZN-RTX: 0.16). These varying correlation patterns provide substantial diversification opportunities, particularly when combining technology assets with utilities, financial, and energy stocks, which enable effective risk reduction while maintaining expected returns in portfolio construction.

Summary Insights:

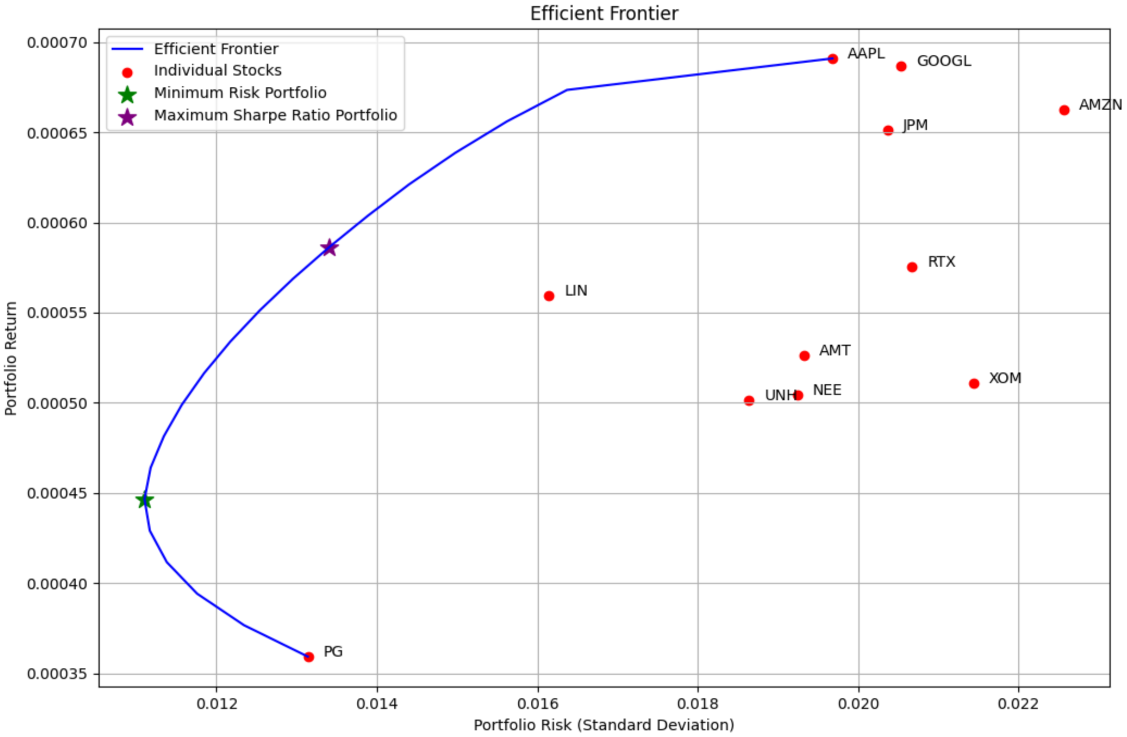
- Growth Stocks (AAPL, AMZN, GOOGL) tend to have higher expected returns but also higher volatility.
- Defensive stocks (PG, NEE) tend to have lower expected returns and offer more stability, as they are less sensitive to market fluctuations.

3.2. Portfolio Optimisation Results



The optimal portfolio weights show significant sensitivity to changes in expected returns and risk levels. In the low-risk portfolio with a low target return, most investments are in stable and low-risk stocks such as PG (47.48%), LIN (11.71%), and UNH (11.31%), showing a preference for safety and minimal risk assets. In the medium-risk portfolio, the allocation becomes more balanced, reducing PG's weight to 16.74% while increasing investments in growth stocks such as AAPL (10.67%) and GOOGL (8.60%). In the high-risk portfolio with a high target return, the focus shifts significantly towards high-return and more volatile stocks, with AAPL (35.58%), GOOGL (27.66%), and JPM (18.22%) receiving the highest allocation, while defensive stocks such as PG are almost removed. This change highlights the tradeoff between risk and return, which higher expected returns require more investment in riskier assets, increasing overall portfolio volatility.

3.3. Efficient Frontier Analysis



The efficient frontier in the plot follows a concave upward shape, which shows the relationship between risk and return in portfolio optimisation. This curve represents the optimal portfolios that offer the highest expected return for a given level of risk. At lower risk levels, the frontier is steeper, indicating that small increases in risk can lead to significant improvements in expected returns, making moderate diversification beneficial. However, at higher risk levels, the curve flattens, indicating additional risk does not correspondingly increase returns, highlighting the challenges of aggressive investment strategies. The minimum risk portfolio, positioned at the lower-left of the frontier, consists of low-volatility stocks such as PG, LIN, and UNH, making it suitable for conservative investors. In contrast, the maximum sharpe ratio portfolio, located near the middle of the curve, offers the highest risk-adjusted returns, balancing risk and return effectively. High-risk portfolios, found at the right end of the frontier, are dominated by high-beta stocks such as AAPL, AMZN, and GOOGL, providing high expected returns but also higher volatility. The shape of the frontier emphasizes the importance of diversification in reducing risk while maintaining returns. Investors can use this information to design portfolios that meet their risk tolerance, whether prioritizing stability, balanced growth, or higher returns. Overall, the efficient frontier highlights how strategic asset allocation can optimize investment performance, reinforcing the benefits of diversification in risk management.

4. Discussion

The optimal portfolio weights are sensitive to changes in expected return and risk. A slight increase in an asset's expected return can significantly impact its weight in the optimal portfolio.

The selection of 20 equally spaced target return values for the efficient frontier provides a comprehensive view of the risk-return tradeoff. These points span from the minimum-risk portfolio to the maximum-return portfolio, allowing investors to select a portfolio that matches their risk tolerance.

5. Conclusion

The portfolio optimisation analysis demonstrates the integration of the Capital Asset Pricing Model for asset return and risk estimation alongside quadratic programming techniques for portfolio weight determination. The efficient frontier clearly illustrates diversification advantages in risk mitigation while maintaining expected returns.

The findings indicate that optimal allocation across the 11 selected assets can significantly reduce risk exposure compared to individual asset investments. The shape of the efficient frontier and the changing asset weights along it provide valuable insights for investors with different risk preferences.

Reference

CFI Team. (n.d.). *Capital Asset Pricing Model (CAPM).* Corporate Finance Institute. https://corporatefinanceinstitute.com/resources/valuation/what-is-capm-formula/

CFI Team. (n.d.). *Alpha.* Corporate Finance Institute. https://corporatefinanceinstitute.com/resources/career-map/sell-side/capital-markets/alpha/