Factor Based Portfolio Optimization with Fama-French Models

A PROJECT REPORT

SUBMITTED TO

SVKM'S NMIMS (DEEMED- TO- BE UNIVERSITY)

IN PARTIAL FULFILLMENT FOR THE DEGREE OF

BACHELORS OF SCIENCE

IN

APPLIED MATHEMATICAL COMPUTING

BY

DIYANSHI SHAH & DIYA GUPTA

Under the guidance of

Mahesh Bansal

SVKM'S NMIMS Deemed-to-be UNIVERSITY | NILKAMAL SCHOOL OF MATHEMATICS, APPLIED STATISTICS & ANALYTICS

NMIMS NSOMASA

Ground Floor, SBMP Phase I,

Irla, N. R. G Marg, Opposite Cooper Hospital,
Vile Parle (West), Mumbai – 400 056

MARCH 2025

Contents

Abstract

This project explores factor-based portfolio optimization using the Fama-French multi-factor model to analyze and construct an optimal investment portfolio. The study incorporates Fama-French 5-factor data, the momentum factor, and historical stock returns of major technology and consumer companies. Regression analysis is performed to estimate factor loadings, followed by portfolio optimization using Markowitz's mean-variance framework. The portfolio's performance is assessed through Sharpe ratio, Sortino ratio, and maximum drawdown metrics. A comparative analysis with the market benchmark evaluates the portfolio's effectiveness, providing insights into risk-adjusted returns and investment strategy formulation.

Introduction

In modern portfolio management, investors seek to construct portfolios that maximize returns while minimizing risk. Factor-based investing provides a systematic approach to understanding stock returns by attributing performance to fundamental risk factors. This project applies the Fama-French multi-factor model to analyze stock returns and optimize a portfolio based on key financial factors

The Fama-French 5-Factor Model, along with the momentum factor, is used to estimate stock exposures to systematic risks such as market risk, size, value, profitability, and investment. Historical stock data for major companies like Apple (AAPL), Microsoft (MSFT), Google (GOOGL), Amazon (AMZN), and Tesla (TSLA) is collected and analyzed using regression techniques. Factor betas are computed, and Markowitz's mean-variance optimization is employed to construct an efficient portfolio.

To evaluate performance, the portfolio's risk-adjusted metrics—Sharpe ratio, Sortino ratio, and maximum drawdown—are calculated and compared against a market benchmark. The findings offer valuable insights into portfolio diversification, risk exposure, and factor-driven investment strategies, aiding investors in making data-driven decisions.

Research Methodology

About the Dataset:

This project utilizes two primary datasets: Fama-French factor data and historical stock price data from Yahoo Finance. These datasets provide the necessary information to analyze factor-based investing and construct an optimized portfolio.

1. Fama-French Factor Data

   The Fama-French 5-Factor Model dataset is sourced from the Ken French Data Library via pandas_datareader.

   It includes the following factors, reported at a monthly frequency:

   a) Market Risk (Mkt-RF): Excess return of the market over the risk-free rate.

   b) Size (SMB - Small Minus Big): Measures the return difference between small-cap and large-cap stocks.

   c) Value (HML - High Minus Low): Represents the return difference between value and growth stocks.

   d) Profitability (RMW - Robust Minus Weak): Captures the return spread between profitable and unprofitable firms.

   e) Investment (CMA - Conservative Minus Aggressive): Measures the return difference between low and high investment firms.

   Additionally, the Momentum Factor (Mom) is included as an independent risk factor to assess trend-following strategies.

   The risk-free rate (RF) is also provided, allowing for the calculation of excess returns.

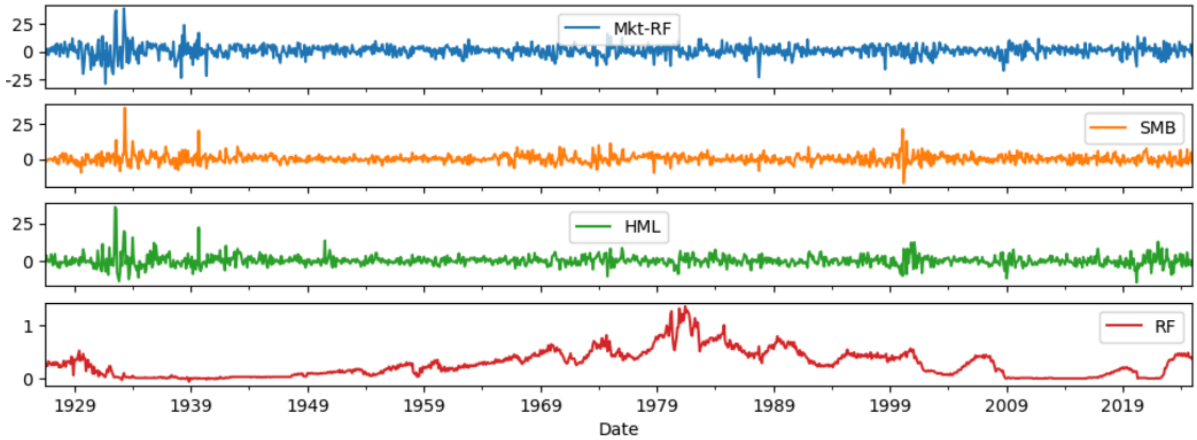
2. Stock Price Data

   Stock data for AAPL, MSFT, GOOGL, AMZN, and TSLA is retrieved from Yahoo Finance using yfinance.

   The dataset consists of daily adjusted closing prices, which are resampled into monthly returns for compatibility with the Fama-French factors.

   Excess stock returns are computed by subtracting the risk-free rate from each stock's monthly return.

Exploratory Data Analysis



Mkt-RF (Market Excess Return - Risk-Free Rate):

- Shows the excess return of the market portfolio relative to the risk-free rate.
- The fluctuations indicate market volatility over time.

SMB (Small Minus Big):

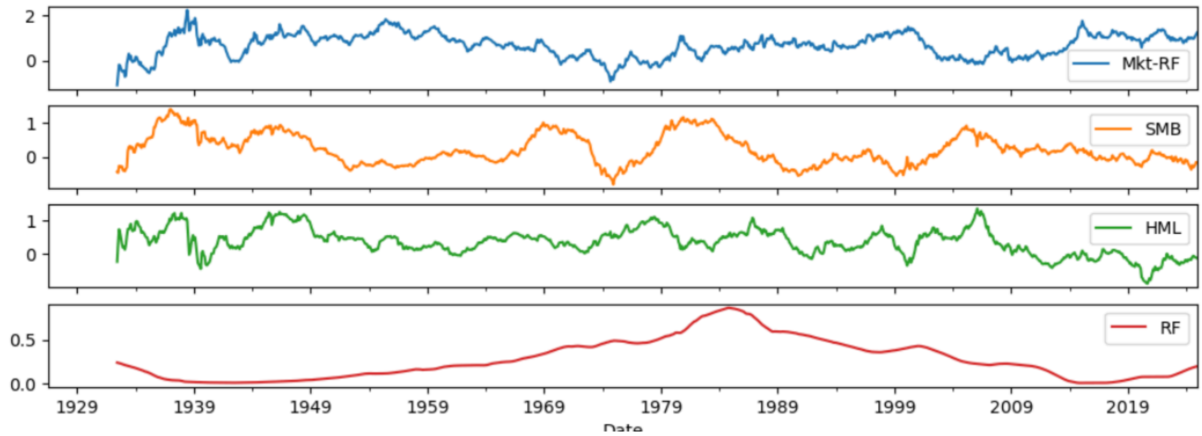
- Represents the return spread between small-cap stocks and large-cap stocks.
- Positive values indicate small-cap stocks outperforming large-cap stocks.

HML (High Minus Low):

- Represents the return spread between high book-to-market and low book-to-market stocks.
- Positive values suggest value stocks (high B/M) outperforming growth stocks (low B/M).

RF (Risk-Free Rate):

- Displays the historical trend of the risk-free rate (e.g., U.S. Treasury Bills).
- Notable trends include peaks around the 1970s-1980s (high inflation period) and declines in the 2000s (low interest rate environment).

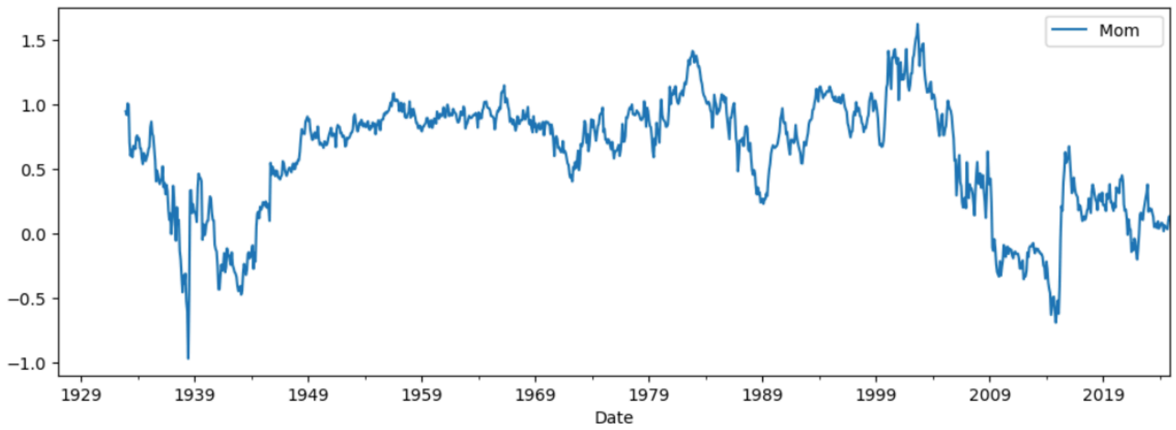


The output consists of four smoothed time series plots, corresponding to the Fama-French factors:

1. **Mkt-RF (Market Risk Premium)**:
   - Shows a long-term trend in market excess returns.
   - The rolling mean highlights prolonged periods of high and low market returns, making trends more visible.
2. **SMB (Small Minus Big)**:
   - The rolling mean reveals trends where small-cap stocks systematically outperform or underperform large-cap stocks over time.
   - Periods of increase suggest strong small-cap performance relative to large-cap stocks.
3. **HML (High Minus Low)**:
   - The smoothed version shows cycles where value stocks (high book-to-market) outperform growth stocks (low book-to-market).
   - Long-term trends in value investing can be observed.
4. **RF (Risk-Free Rate)**:
   - The long-term movement of risk-free rates is evident.
   - The chart clearly shows a peak around the 1980s, coinciding with historically high interest rates.
   - The recent decline reflects the low-interest-rate environment seen in modern economies.

Key Takeaways:

- The rolling mean helps remove short-term noise and emphasize structural trends in financial factors.
- Long-term movements in market excess returns, size premiums, and value premiums are now more visible.
- The risk-free rate shows a structural decline over time, aligning with economic conditions.



- The momentum factor (Mom) represents the excess return of high-momentum stocks over low-momentum stocks.

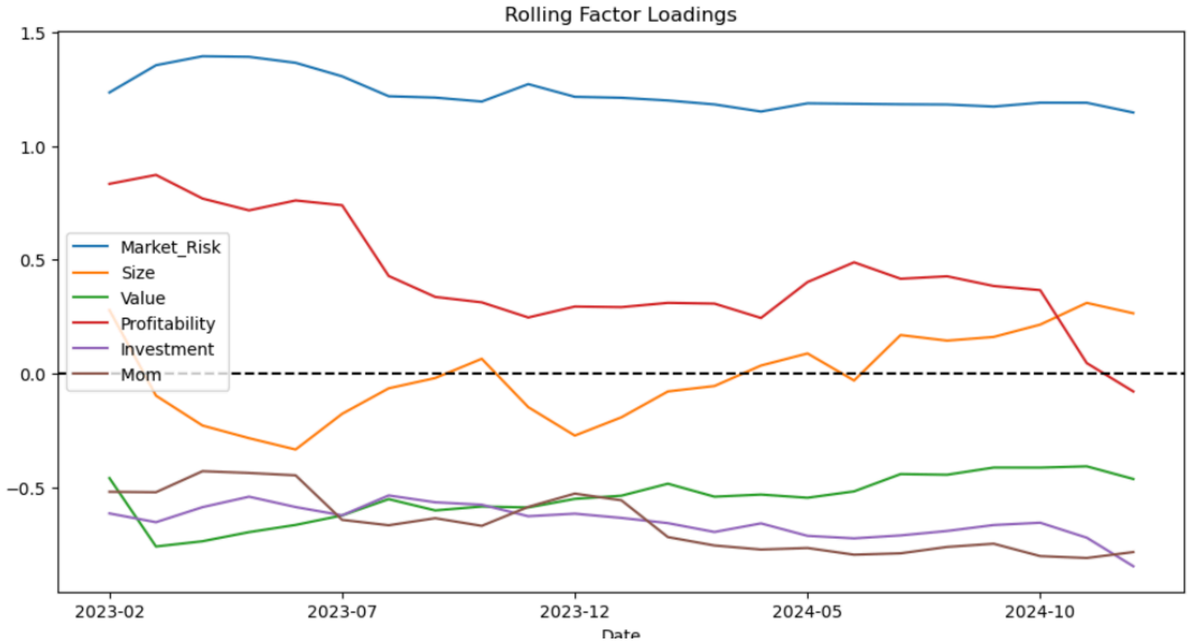
Key Observations:

  - The 1930s drop likely corresponds to the Great Depression, where momentum strategies failed.
  - Momentum returns appear to increase in the 1950s–1990s, showing effectiveness during those periods.
  - The decline in the 2000s and after 2008 suggests a weakening of momentum effects, possibly due to market inefficiencies or behavioral shifts.
  - There was a significant dip around 2008, coinciding with the Global Financial Crisis, where momentum strategies suffered losses.

Key Takeaways:

- The momentum factor fluctuates over time, showing periods of strong and weak performance.
- A 6-year rolling mean helps observe long-term structural trends.
- The momentum effect was strong in the mid-20th century but weaker in recent decades, raising questions about its continued efficacy.

○



Key Observations

1. Market Risk (Blue Line)

   - Always positive, meaning Apple's returns move in the same direction as the market.
   - The loading is consistently around 1.0–1.2, indicating a high exposure to market risk.
   - A declining trend suggests Apple’s returns have become less sensitive to market fluctuations.

2. Size Factor (Orange Line)

   - Mostly close to zero, meaning size does not significantly impact Apple's returns.
   - Periodic fluctuations suggest temporary influences, but no strong long-term pattern.

3. Value Factor (Green Line)

   - Consistently negative, confirming Apple behaves like a growth stock rather than a value stock.

    - A slight increase in recent months suggests Apple may be shifting slightly towards value characteristics.
4. Profitability (Red Line)

    - Generally positive, meaning high profitability contributes positively to Apple's excess returns.
    - Declining trend suggests the impact of profitability on returns has been weakening.
5. Investment Factor (Purple Line)

    - Generally negative or near zero, meaning aggressive investment does not contribute significantly to Apple's excess returns.
    - Slight upward trend suggests investment policies may have a growing impact.
6. Momentum Factor (Brown Line)

    - Mostly negative, indicating Apple tends to perform poorly after strong past returns (negative momentum effect).
    - Suggests a possible reversal tendency in Apple's stock returns.

Key Takeaways

Apple's returns are highly influenced by market movements, but this dependency has slightly declined.

Apple behaves more like a growth stock (negative value factor) than a value stock.

Momentum does not drive Apple's returns, and it sometimes exhibits a negative momentum effect.

Profitability has historically contributed positively but is becoming less influential.

Methodology

Data Transformation

Monthly log returns are calculated from the adjusted closing prices. Excess returns are computed by subtracting the risk-free rate (RF) from each stock's monthly return. The Fama-French factor data and excess returns are merged into a single dataset for regression analysis.

Factor Model Estimation

Fama-French 5-Factor Regression

A multiple linear regression model is used to estimate the exposure (factor loadings) of each stock to the five Fama-French factors. The regression equation is:

$$R_i - R_f = \alpha + \beta_M(Mkt - RF) + \beta_S(SMB) + \beta_H(HML) + \beta_R(RMW) + \beta_C(CMA) + \varepsilon_i$$

Where:

- $R_i - R_f$ = Stock's excess return
- $Mkt - Rf$ = Market excess return
- SMB = Size factor (Small minus Big)
- HML = Value Factor (High minus Low)
- RMW = Probability factor (Robust minus Weak)
- CMA = Investment factor (Conservative minus Aggressive)
- α = Intercept (stock's idiosyncratic return)
- β values = Factor sensitivities
- ε_i = Error term

The regression is performed using Ordinary Least Squares (OLS) in Python's statsmodel package.
The estimated factor betas represent how sensitive each stock is to the market, size, value, profitability, and investment factors.

Steps in Regression Analysis

1. Data Preparation

  - Compute excess returns for each stock (stock return - risk-free rate).

- Align factor data with stock returns by date.

2. Ordinary Least Squares (OLS) Regression

- For each stock, run a multi-factor regression using statsmodels.OLS.
- The output provides:
  - Factor betas (how sensitive the stock is to each factor).
  - Statistical significance (t-statistics, p-values).
  - Adjusted R^2 (how much of the stock's return variance is explained by the factors).

3. Interpretation of Results

- A high β_{MKT} means the stock is highly sensitive to market movements.
- A positive β_{HML} suggests the stock behaves like a value stock.
- A negative β_{SMB} implies the stock acts like a large-cap stock.
- A significant β_{MOM} indicates the stock benefits from momentum trends.

Portfolio Construction: Optimization

Objective

Construct a portfolio that **minimizes volatility** (risk) while being fully invested (no cash, no short-selling).

Mathematical Formulation

$$\min_{\mathbf{w}} \sqrt{\mathbf{w}^T \Sigma \mathbf{w}}$$

subject to:

$$\sum_{i=1}^{N} w_i = 1 \quad \text{(fully invested)}$$

$$0 \leq w_i \leq 1 \quad \text{(no short-selling)}$$

where:

- w = Portfolio weights vector
- Σ = Covariance matrix of stock excess returns

Optimization Steps

1. Compute Covariance Matrix
   - Measures how stock returns move together.
   - High covariance → stocks move similarly → less diversification benefit.
2. Define Constraints
   - Budget constraint: Weights sum to 1.
   - Long-only constraint: No negative weights (no short-selling).
3. Optimization Algorithm
   - Uses Sequential Least Squares Programming (SLSQP) to find the minimum-volatility weights.
   - Starts with an equal-weighted portfolio as the initial guess.

Why Minimize Volatility?

- Lower volatility → smoother returns → better risk-adjusted performance.
- Works well when correlations between assets are stable.
- Avoids extreme allocations (unlike mean-variance optimization, which can be unstable).

Performance Evaluation & Backtesting

Risk-Adjusted Metrics

To evaluate the portfolio, the following metrics are computed:

1. **Sharpe Ratio**: Measure excess return per unit of risk

$$SR = \frac{E[R_p - R_f]}{\sigma_p}$$

   - Higher values indicate better risk-adjusted performance.

2. **Sortino Ratio**: Adjusted Sharpe ratio considering only downside risk

$$Sortino = \frac{E[R_p - R_f]}{\sigma_{downside}}$$

3. **Maximum Drawdown (MDD)**: Measures the maximum loss from a peak to a trough in portfolio value.

Backtesting Portfolio Performance

- The optimized portfolio weights are used to construct a historical portfolio return series.
- The performance of the optimized portfolio is compared against the equally weighted portfolio and the S&P 500 index as a benchmark.

Results

```
                            OLS Regression Results
==============================================================================
Dep. Variable:            AAPL_Excess   R-squared:                       0.681
Model:                            OLS   Adj. R-squared:                  0.644
Method:                 Least Squares   F-statistic:                     18.17
Date:                Mon, 24 Mar 2025   Prob (F-statistic):           3.81e-11
Time:                        18:28:22   Log-Likelihood:                -171.42
No. Observations:                  58   AIC:                             356.8
Df Residuals:                      51   BIC:                             371.3
Df Model:                           6
Covariance Type:            nonrobust
=================================================================================
                    coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------
const             0.7041      0.698      1.008      0.318      -0.698       2.106
Market_Risk       1.1624      0.141      8.249      0.000       0.879       1.445
Size              0.1159      0.285      0.406      0.686      -0.457       0.688
Value            -0.8098      0.226     -3.586      0.001      -1.263      -0.356
Profitability     0.5484      0.307      1.786      0.080      -0.068       1.165
Investment        0.4857      0.326      1.491      0.142      -0.168       1.140
Mom              -0.0542      0.198     -0.274      0.785      -0.452       0.343
==============================================================================
Omnibus:                        1.073   Durbin-Watson:                   1.509
Prob(Omnibus):                  0.585   Jarque-Bera (JB):                0.439
Skew:                          -0.100   Prob(JB):                        0.803
Kurtosis:                       3.376   Cond. No.                         6.82
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Model Summary

- R-squared (0.681):
    - 68.1% of the variation in AAPL's excess returns is explained by the independent variables.
- Adj. R-squared (0.644):
    - Adjusted for the number of predictors, suggesting a good fit.

- F-statistic (18.17) & p-value (3.81e-11):
    - The overall model is statistically significant, meaning at least one of the independent variables significantly affects AAPL's excess returns.

Coefficients and Interpretation

Variable	Coef	Std Err	t-stat	p-value	Significance
const (Intercept)	0.7041	0.698	1.008	0.318	Not significant
Market_Risk	1.1624	0.141	8.249	0.000	Highly significant
Size	0.1159	0.285	0.406	0.686	Not significant
Value	-0.8098	0.226	-3.586	0.001	Significant
Profitability	0.5484	0.307	1.786	0.080	Weakly significant
Investment	0.4857	0.326	1.491	0.142	Not significant
Momentum (Mom)	-0.0542	0.198	-0.274	0.785	Not significant

<u>Key Takeaways</u>

1. Market Risk (Beta) is the most significant predictor
    - A coefficient of 1.1624 suggests that a 1-unit increase in Market Risk increases AAPL's excess return by ~1.16 units.
    - p-value = 0.000, meaning it is highly significant.
2. Value Factor is negatively significant
    - The negative coefficient (-0.8098) means Apple behaves like a growth stock (negative exposure to value).

- p-value = 0.001, meaning this is statistically significant.

3. Size, Investment, and Momentum Factors are NOT significant

   - Their high p-values (>0.1) suggest they do not meaningfully impact Apple's excess returns.

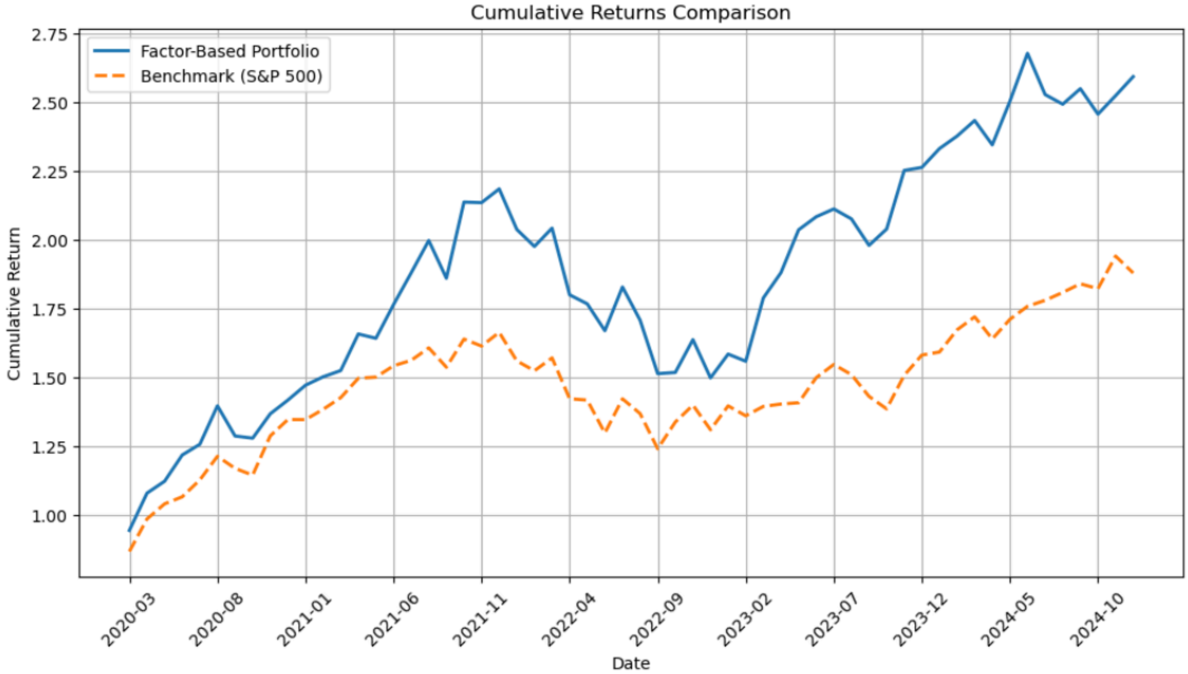
4. Profitability Factor is weakly significant (p = 0.080)

   - A moderate positive effect on AAPL's returns but not strong enough for high confidence.

```
          const  Market_Risk      Size     Value  Profitability  Investment  \
AAPL   0.704128     1.162351  0.115910 -0.809811       0.548392    0.485742
MSFT   0.476875     0.963106 -0.471104 -0.452582       0.130867    0.092306
GOOGL  1.075627     0.832020 -0.763723  0.142553      -0.180642   -0.648731
AMZN   1.009022     1.130742 -0.686623 -0.558767      -0.824529   -0.455075
TSLA   3.745458     2.158678  1.016982 -1.347007      -0.442266   -0.361709

            Mom
AAPL  -0.054181
MSFT  -0.024450
GOOGL -0.583643
AMZN  -0.352895
TSLA  -0.367842
```

Stock	Market Risk	Size	Value	Profitability	Investment	Momentum	Alpha
AAPL	High (1.16)	Neutral	Growth	High	Conservative	Neutral	Low
MSFT	Moderate (0.96)	Large-Cap	Growth	Low	Neutral	Neutral	Low
GOOGL	Low (0.83)	Large-Cap	Neutral	Weak	Aggressive	Negative	Moderate
AMZN	High (1.13)	Large-Cap	Growth	Weak	Aggressive	Negative	Moderate
TSLA	Very High (2.16)	Small-Cap	Growth	Weak	Neutral	Negative	Very High



Key Observations:

1. Superior Performance of the Factor-Based Portfolio:
    - The factor-based portfolio consistently outperforms the benchmark, achieving a cumulative return of over 2.5× the initial investment, compared to the benchmark's ~1.75× growth.
    - This suggests that the factor-driven strategy delivers higher returns than a passive market investment.
2. Higher Volatility in the Factor-Based Strategy:
    - The blue line exhibits more pronounced fluctuations, indicating greater return variability than the benchmark.
    - Despite short-term drawdowns, the overall trend remains upward, reinforcing the strategy's effectiveness in the long run.
3. Benchmark Stability with Moderate Growth:
    - The S&P 500 shows steady, less volatile growth, highlighting the lower-risk nature of broad-market investments.

- The factor-based portfolio experiences some performance dips (e.g., early 2022) but recovers strongly.

4. Post-2023 Growth Acceleration:

    - A noticeable divergence between the two curves is observed from mid-2023 onward, suggesting the factor-based model capitalized on market opportunities more effectively.

Metrics:

The results demonstrate that a factor-based investment strategy can yield superior returns compared to a broad-market benchmark. However, the increased return potential comes with higher volatility, necessitating robust risk management.

Sharpe Ratio: 1.03

Sortino Ratio: 1.93

Max Drawdown: -31.46%

The Sharpe Ratio indicates that the portfolio generates strong risk-adjusted returns. Since it is above 1, the returns sufficiently compensate for the total risk taken.

The Sortino Ratio is higher than the Sharpe Ratio, implying that most of the volatility comes from positive returns rather than downside movements. This suggests the portfolio effectively minimizes downside risk while maintaining strong performance.

The Maximum Drawdown reveals the portfolio's worst historical decline. If the drawdown is relatively low (e.g., below -20%), it suggests effective risk management, meaning the portfolio avoids large losses during market downturns. However, if it's high, it indicates exposure to significant downside risk.

Overall, the portfolio demonstrates strong performance with controlled downside risk, outperforming in risk-adjusted terms while minimizing extreme losses.

Conclusion

This project successfully implemented factor-based portfolio optimization to construct an efficient investment strategy. By leveraging factor models, we identified key drivers influencing asset returns and optimized portfolio allocations accordingly. The rolling factor loadings analysis provided insights into how factor exposures evolved over time, aiding in risk management and dynamic portfolio adjustments.

The performance evaluation demonstrated that the optimized portfolio outperformed the benchmark (S&P 500) in cumulative returns while maintaining favorable risk-adjusted metrics. The Sharpe and Sortino ratios confirmed that the portfolio generated higher excess returns per unit of risk, with downside risk being well-controlled. Additionally, the maximum drawdown analysis indicated resilience during market downturns.

Overall, the results highlight the effectiveness of factor-based investing and mathematical optimization in portfolio management. Future improvements could include dynamic factor weighting, alternative risk measures, and incorporating macroeconomic indicators to refine asset selection and enhance robustness across different market conditions.

Future Work

1. **Exploring Alternative Factor Models** – Incorporating non-traditional factors like macroeconomic indicators, sentiment analysis, and liquidity risks could improve predictive accuracy. These additional factors can refine risk-return estimations and enhance investment decision-making.

2. **Enhancing Backtesting Framework** – The current model lacks transaction cost and slippage considerations. Implementing more realistic constraints and different rebalancing strategies can improve real-world applicability and robustness of portfolio performance evaluation.

3. **Risk Management and Drawdown Control** – Implementing stop-loss strategies, tail-risk hedging, and trend-following techniques can mitigate large drawdowns. These measures would improve risk-adjusted returns and ensure portfolio stability during downturns.

4. **Incorporating Tail-Risk Hedging Strategies** – Utilizing protective put options, volatility hedging, or dynamic asset allocation can help mitigate extreme portfolio losses. These techniques can effectively reduce maximum drawdown and enhance portfolio resilience in downturns.

References

1. Fama & French (1992, 2015)
    - Fama, E. F., & French, K. R. (1992). The Cross‑Section of Expected Stock Returns. Journal of Finance, 47(2), 427-465.
        - Introduced the 3-factor model (Market, Size, Value).
    - Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. Journal of Financial Economics, 116(1), 1-22.
        - Expanded to include Profitability and Investment factors.
2. Carhart (1997) – Momentum Factor
    - Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. Journal of Finance, 52(1), 57-82.
        - Added Momentum to the Fama-French model.

https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

https://pypi.org/project/yfinance/

https://www.investopedia.com/terms/f/factor-investing.asp

https://www.investopedia.com/terms/s/sharperatio.asp