Factor Modelling with Portfolio Optimization

With Fama - French Models

Table of Contents

01. Introduction

Problem Statement

This project investigates the application of the Fama-French 5-factor model with momentum in constructing and optimizing equity portfolios. The study analyzes how major technology stocks (AAPL, MSFT, GOOGL, AMZN, TSLA) are influenced by market, size, value, profitability, investment, and momentum factors through regression analysis.

A key challenge addressed is determining whether factor-based optimization can generate superior risk-adjusted returns compared to market benchmarks. The research employs mean-variance optimization to build minimum-volatility portfolios while incorporating realistic constraints like no short-selling. Performance evaluation focuses on Sharpe and Sortino ratios, maximum drawdowns, and benchmark comparisons.

02. Data Description

Overview

This project analyzes factor-based investing using the Fama-French 5-Factor Model and historical stock price data.

The goal is to construct an optimized portfolio by understanding the impact of risk factors on stock returns.
We use two primary datasets:

1. Stock Price Data (retrieved from Yahoo Finance using yfinance)
2. Fama-French Factor Data (from the Ken French Data Library via pandas_datareader)

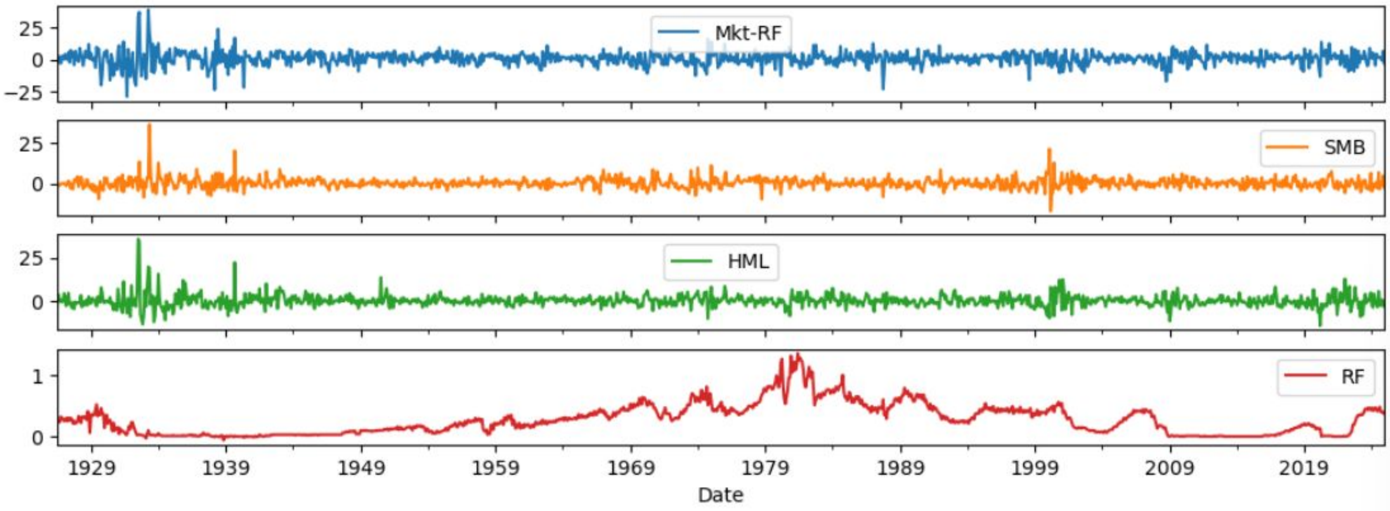
Stock Price Data

- Stocks analyzed: AAPL, MSFT, GOOGL, AMZN, TSLA.
- Source: Retrieved from Yahoo Finance using yfinance.
- Timeframe: Daily adjusted closing prices are resampled into monthly returns.
- Excess returns are computed by subtracting the risk-free rate from each stock's monthly return.
- This data is essential for understanding stock behavior and factor exposure.

- The Fama-French dataset includes monthly factor returns that explain stock performance beyond traditional market risk.
- Factors included:
    1. Market Risk (Mkt-RF): Excess return of the market over the risk-free rate.
    2. Size (SMB - Small Minus Big): Measures the return difference between small-cap and large-cap stocks.
    3. Value (HML - High Minus Low): Represents the return difference between value and growth stocks.
    4. Profitability (RMW - Robust Minus Weak): Captures the return spread between profitable and unprofitable firms.
    5. Investment (CMA - Conservative Minus Aggressive): Measures the return difference between low and high investment firms.
- Momentum Factor (Mom) is included separately to assess trend-following strategies.
- The risk-free rate (RF) is also provided to calculate excess returns.

Fama-French Factor Data

```
array([<Axes: xlabel='Date'>, <Axes: xlabel='Date'>,
       <Axes: xlabel='Date'>, <Axes: xlabel='Date'>], dtype=object)
```



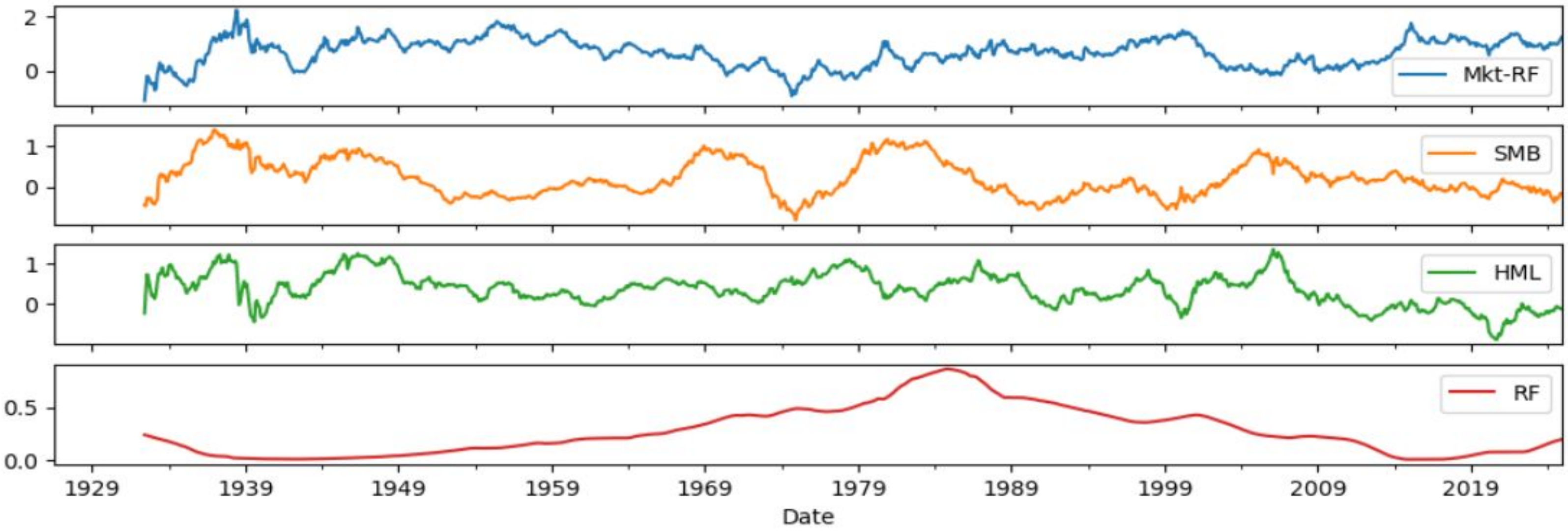
The four subplots show the key factors over time:

- Mkt-RF (Market Risk Premium) – Blue
- SMB (Size Factor) – Orange
- HML (Value Factor) – Green
- RF (Risk-Free Rate) – Red

The fluctuations represent market movements and economic cycles.

Helps in understanding how different factors contribute to stock returns.

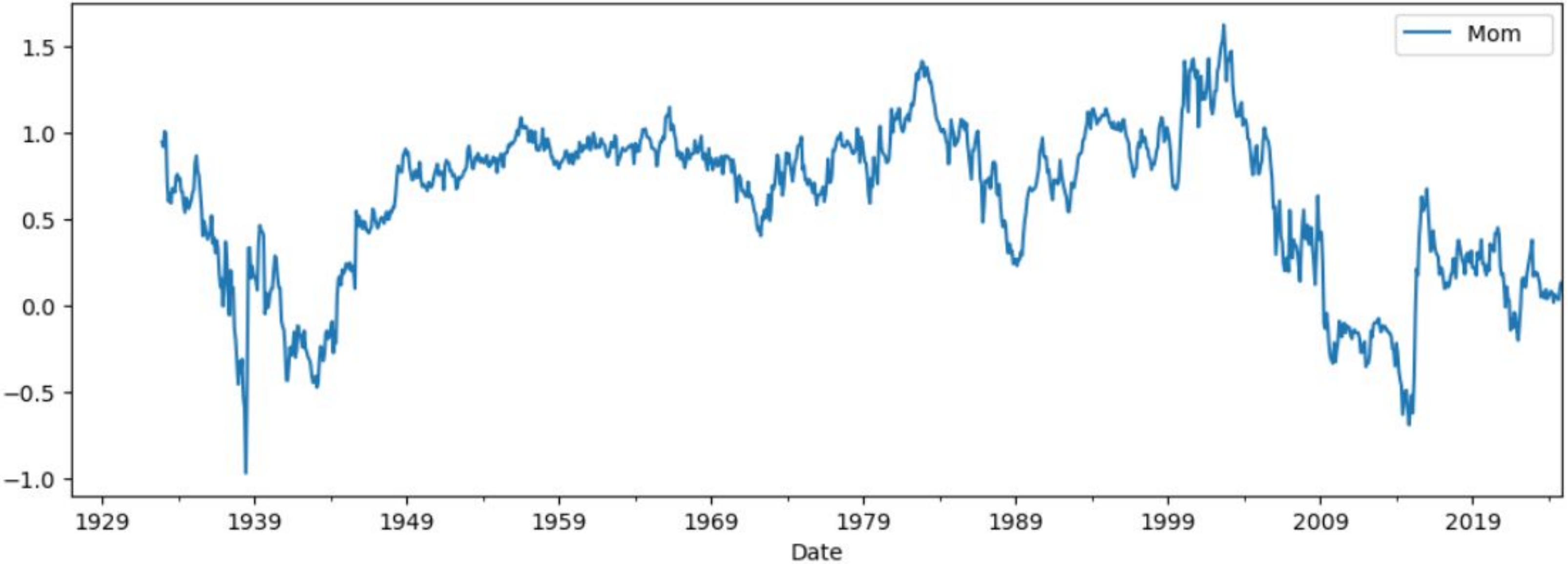
```
array([<Axes: xlabel='Date'>, <Axes: xlabel='Date'>,
       <Axes: xlabel='Date'>, <Axes: xlabel='Date'>], dtype=object)
```



Rolling Mean Analysis of Fama-French Factors

- Market Risk (Mkt-RF): Shows long-term upward trends
- Size (SMB - Small Minus Big): Indicates higher returns for small-cap stocks in certain periods
- Value (HML - High Minus Low): Demonstrates fluctuations in the performance of value vs. growth stocks
- Risk-Free Rate (RF): Long-term declining trend, mirroring historical interest rate trends.
- Key Takeaway: Factor performance varies over time, influencing investment strategies and risk assessment.

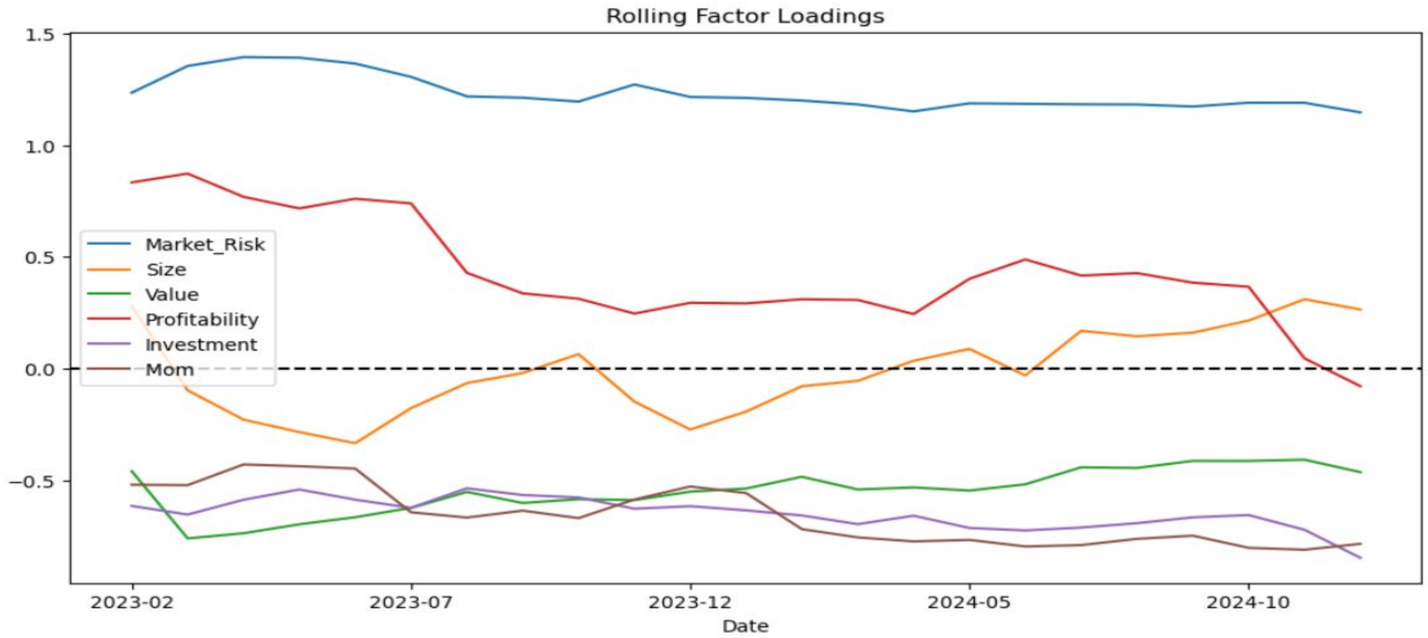
```
array([<Axes: xlabel='Date'>], dtype=object)
```



Represents: Trend-following returns (winners keep winning, losers keep losing).

Key Trends:

- Sharp drop in 1930s-40s (Great Depression, WWII).
- Peaks before 2000 Dot-com Bubble, declines after.
- Falls sharply in 2008 Financial Crisis.
- Post-2010: Volatile recovery.



- Represents: How factor exposures change over time using a 36-month rolling window.
- Key Observations:
  - Market Risk (Blue): Strongest and stable factor.
  - Size (Orange): Fluctuates around zero, showing varying influence.
  - Value (Green) & Investment (Purple): Mostly negative, suggesting weaker returns.
  - Momentum (Red): Declining, indicating weaker trend persistence.
- Importance:
  - Helps analyze portfolio dynamics and adjust strategies based on shifting factor influences.

03. Methodology

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

Portfolio Construction: Optimization

Objective

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

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.

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

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

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.

04. Results

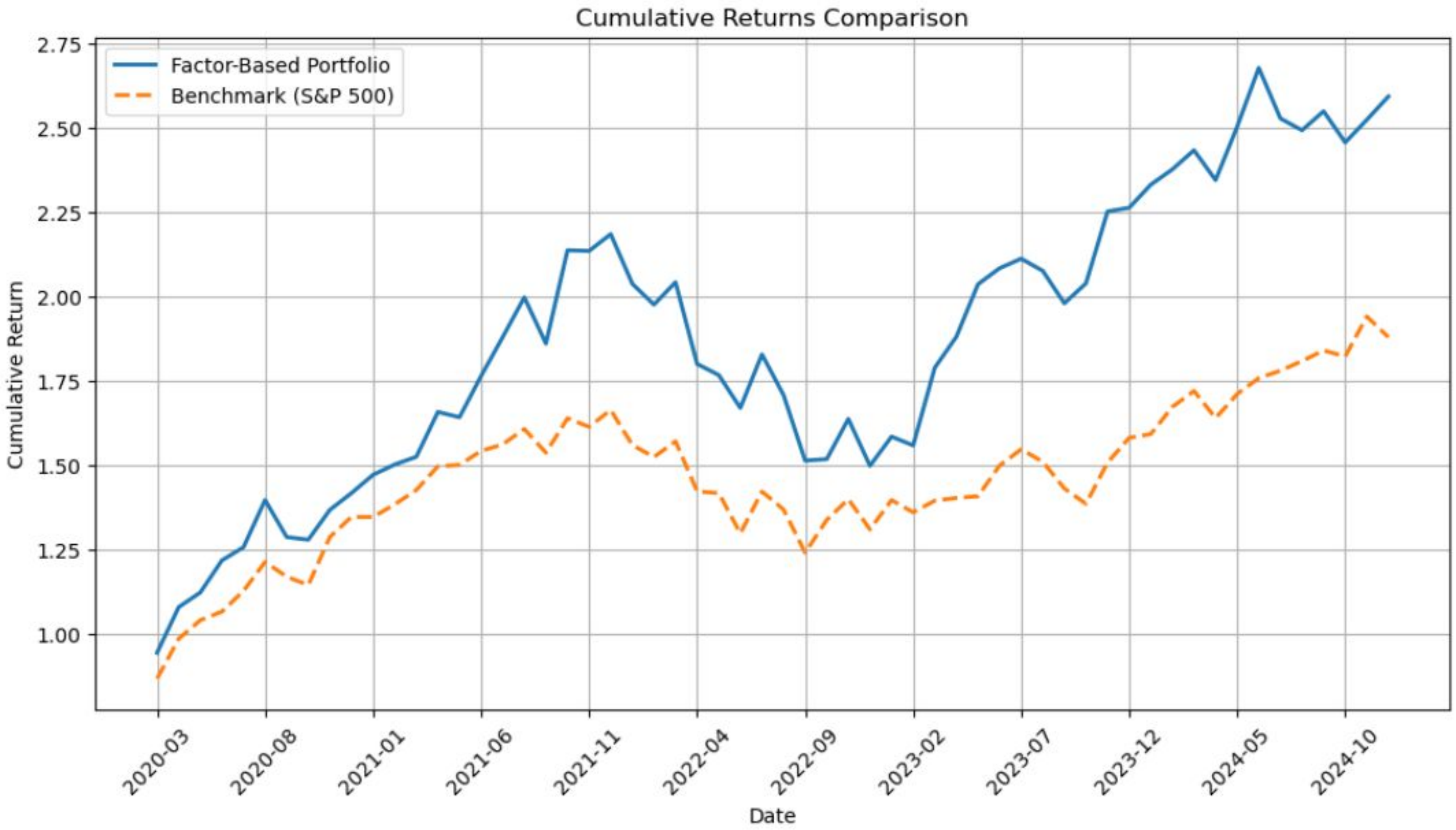
```
                            OLS Regression Results
==============================================================================
Dep. Variable:            AAPL_Excess   R-squared:                       0.681
Model:                            OLS   Adj. R-squared:                  0.644
Method:                 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.
```

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

	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


Cumulative Returns Comparison
Factor-Based Portfolio
Benchmark (S&P 500)
Cumulative Return
2.75
2.50
2.25
2.00
1.75
1.50
1.25
1.00
2020-03
2020-08
2021-01
2021-06
2021-11
2022-04
2022-09
2023-02
2023-07
2023-12
2024-05
2024-10
Date

Sharpe Ratio: 1.03

Sortino Ratio: 1.93

Max Drawdown: -31.46%

05. Conclusion

This project integrates factor modeling through the Fama-French three-factor model, which extends the Capital Asset Pricing Model (CAPM) by incorporating Size (SMB) and Value (HML) factors alongside market risk. By regressing excess stock returns on these factors, we identified how different securities are influenced by broader market dynamics and risk premia. This allowed for a more nuanced understanding of systematic risks beyond traditional beta-based analysis. The factor loadings provided insights into how each stock responds to market conditions, aiding in portfolio construction by selecting assets with favorable risk-return trade-offs. Incorporating factor-based analysis helped in enhancing diversification, mitigating idiosyncratic risks, and optimizing portfolio performance against broader economic conditions.

06. Future Work

- Alternative Factor Models – Integrate macroeconomic indicators, sentiment analysis, and liquidity risks for better risk-return estimations.
- Enhanced Backtesting – Incorporate transaction costs, slippage, and dynamic rebalancing for real-world applicability.
- Risk Management – Apply stop-loss, tail-risk hedging, and trend-following to reduce drawdowns and improve stability.
- Tail-Risk Hedging – Use protective puts, volatility hedging, and dynamic allocation to enhance portfolio resilience.

Thank You

Diya Gupta & Diyanshi Shah