**5 Replication**

In this section, we replicate the model built in the reference report.

**5.1 Data**

We build zero-investment long-short based on a list of established factors. The factors we selected can be established by only using the available information such as opening price, closing price, high price, low price, trading volume and market capitalization.

Table 1 shows the constructions of different zero-investment long-short portfolios

The main aspects of these 22 factors are size, momentum, volume and volatility so that we could use them to form portfolios. Among them, compared to the replicated paper, the age factor was removed for we lack the data of the age. Also, the four factors BETA, BETA^2, IDIOVOL and DELAY were also removed for the date coverage in our model is short, so calculating them can cause losing data and reduce reliability.

Next, we will use them to examine the stochastic dominance against 3 asset’s benchmark: stock, bond and risk-free rate.

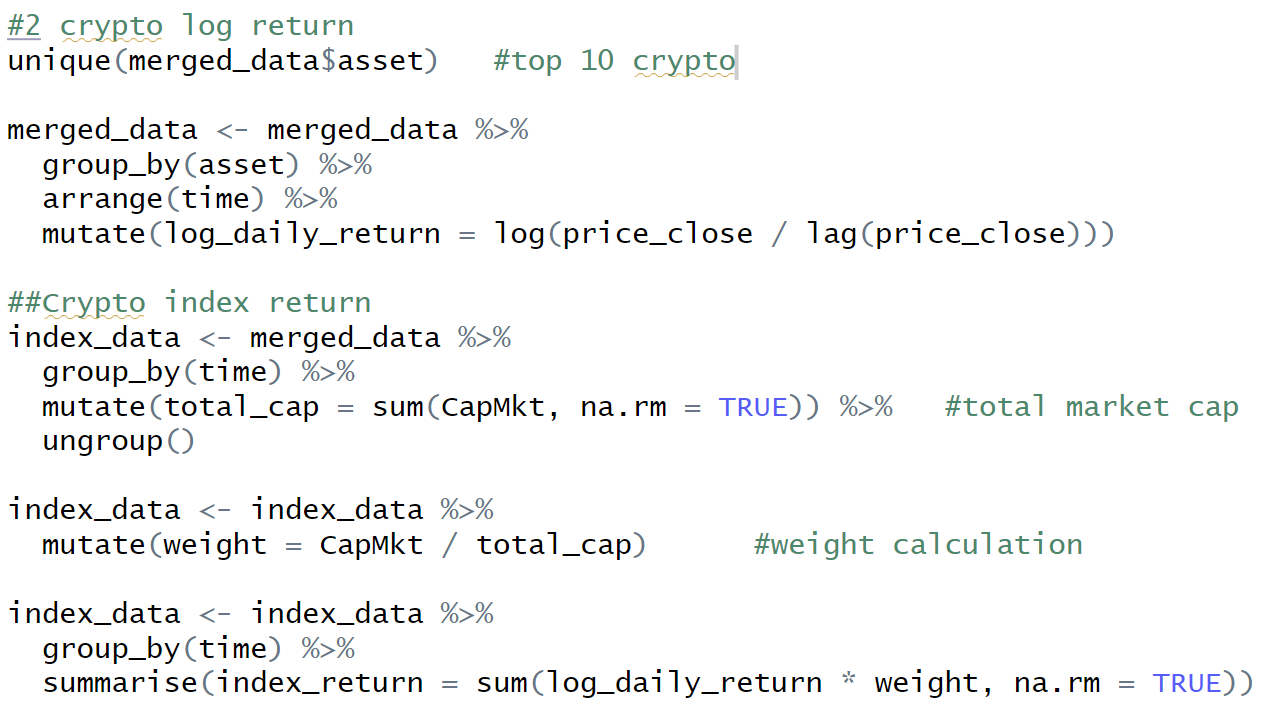
**5.2 Replication of Key Analytical Techniques**

**5.2.1 Data preparing**

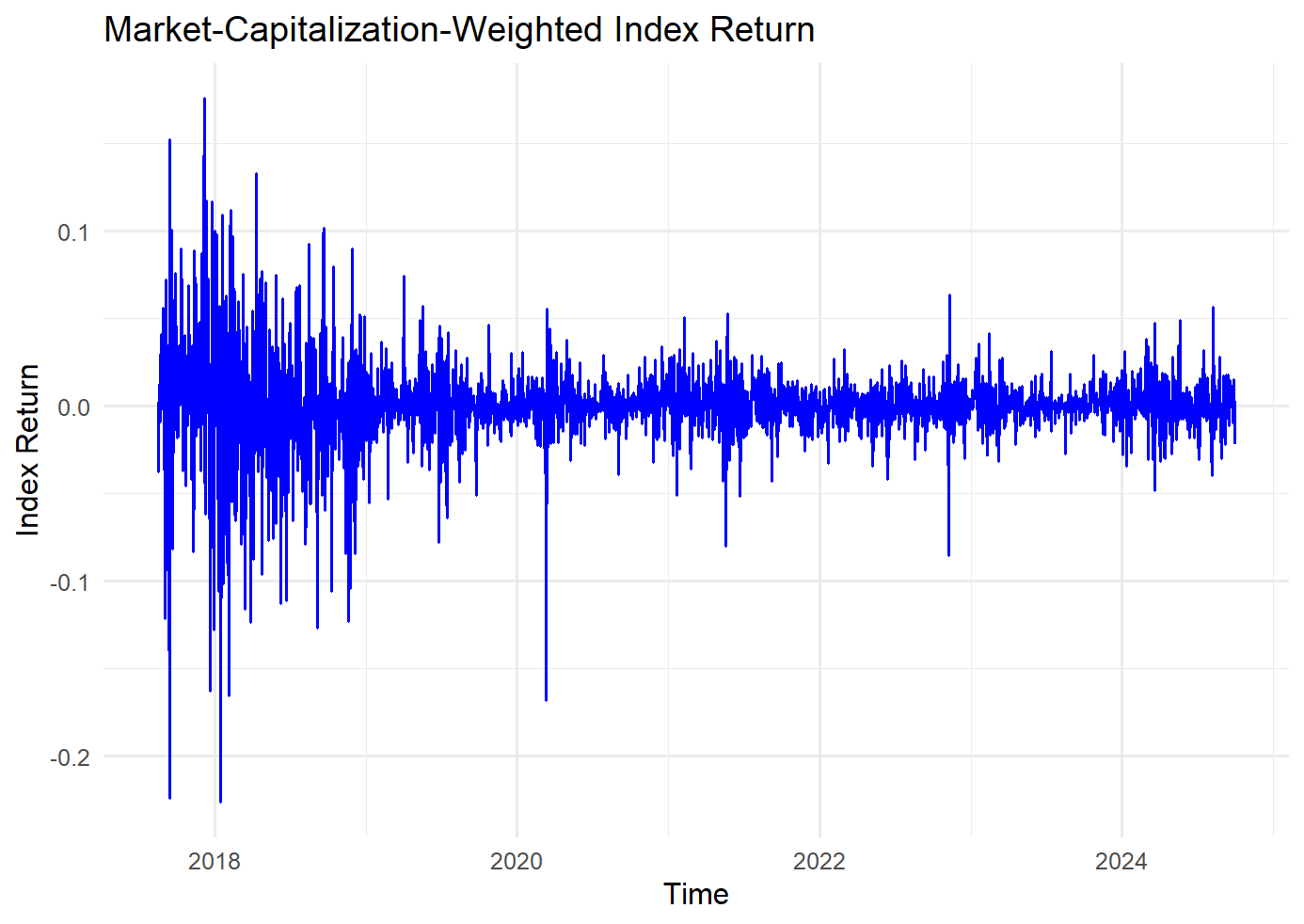
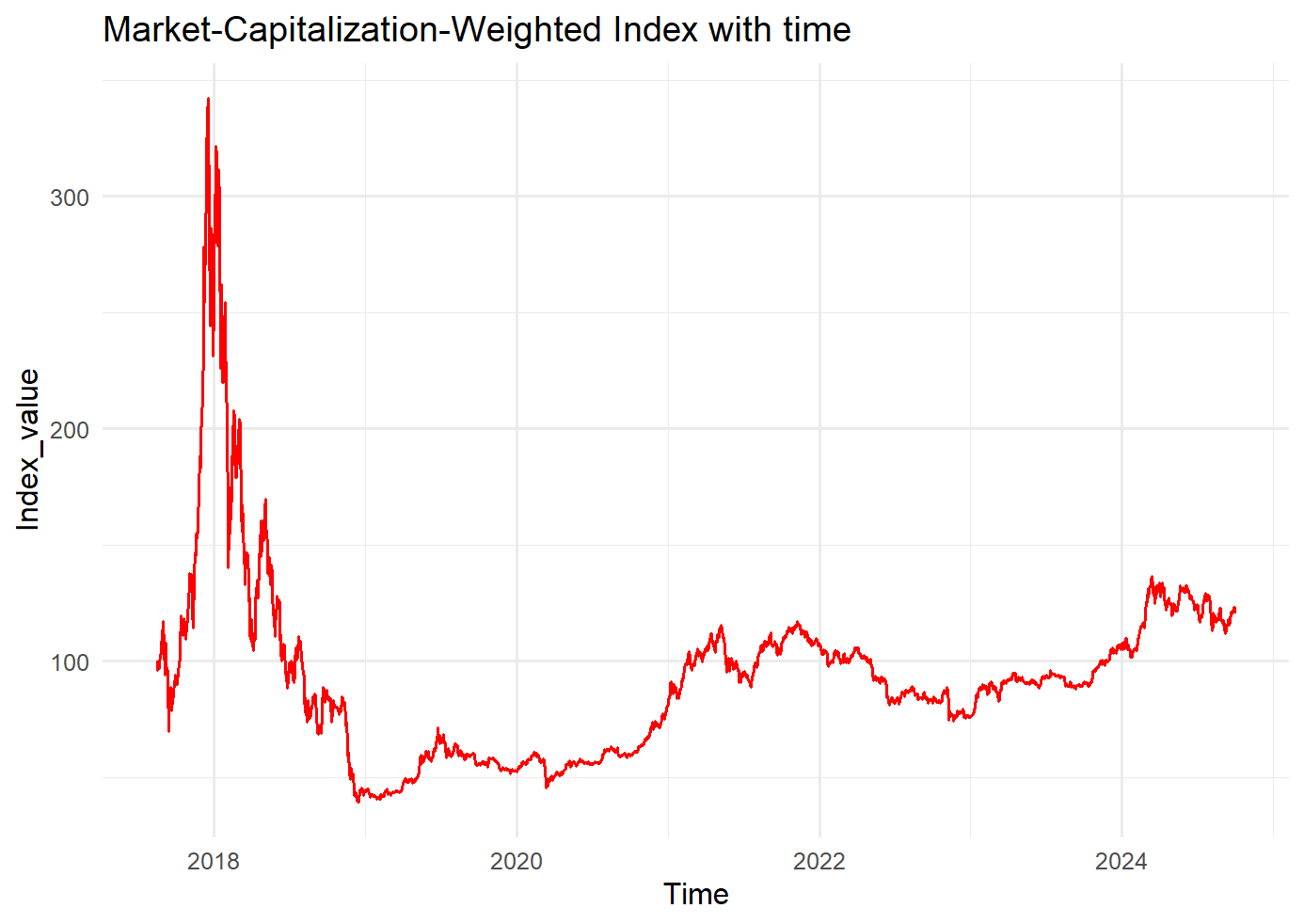
Prepares the data for calculating returns and the weighted index through loading data1 containing market cap data for top cryptocurrencies and data2 containing trading pairs for top stablecoins and cryptocurrencies. Extract asset information from the market field in data2.

Then calculate the Cryptocurrency Log Returns for each cryptocurrency based on its closing price.

Calculate the Market-Capitalization-Weighted Index Return for all the assets on each day (total\_cap), and determines the weight of each asset in the market based on its market cap.

The daily market-cap-weighted index return is computed as the sum of the log daily returns for each cryptocurrency, weighted by their market cap:

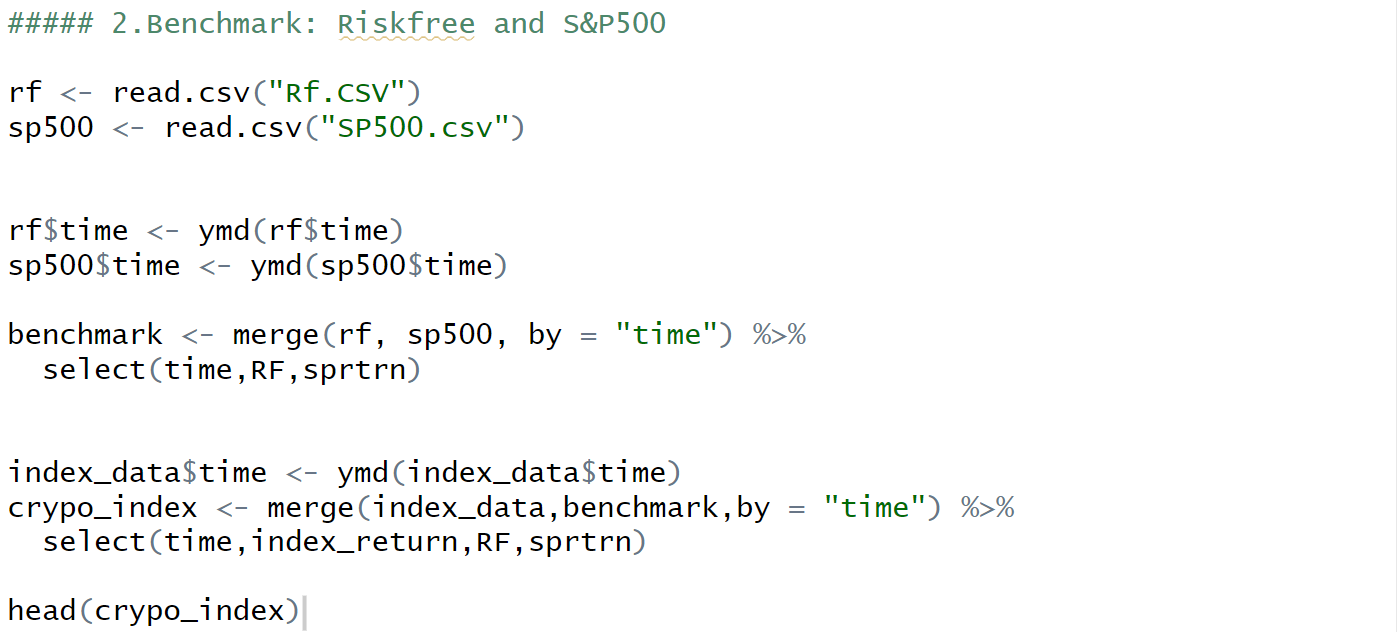
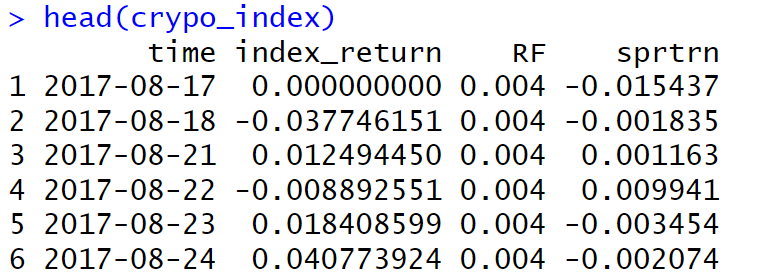
Then we plot the Market-Capitalization-Weighted Index Return and the Market-Capitalization-Weighted Index Value. These plots together provide insights into the daily fluctuations and longer-term trends in the top cryptocurrency market. The index return plot shows the day-to-day variability, while the index value plot demonstrates how the overall market capitalization changes cumulatively over time.

The blue line represents the daily returns of a hypothetical index based on the market capitalization of the top 10 cryptocurrencies.

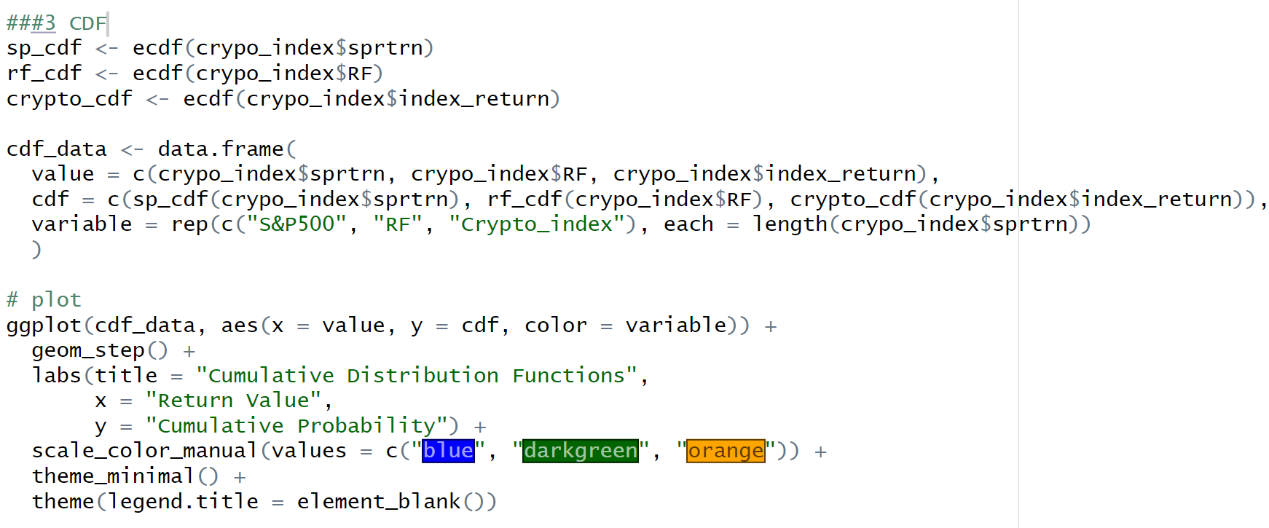
The red line shows how the index value evolves over time. It reflects the cumulative effect of the daily returns.

5.2.2 Set Benchmark: Risk-free and S&P500

We combine the benchmark data with the market-capitalization-weighted crypto index return to compare the performance of cryptocurrencies with traditional financial benchmarks of Risk-free and S&P500.

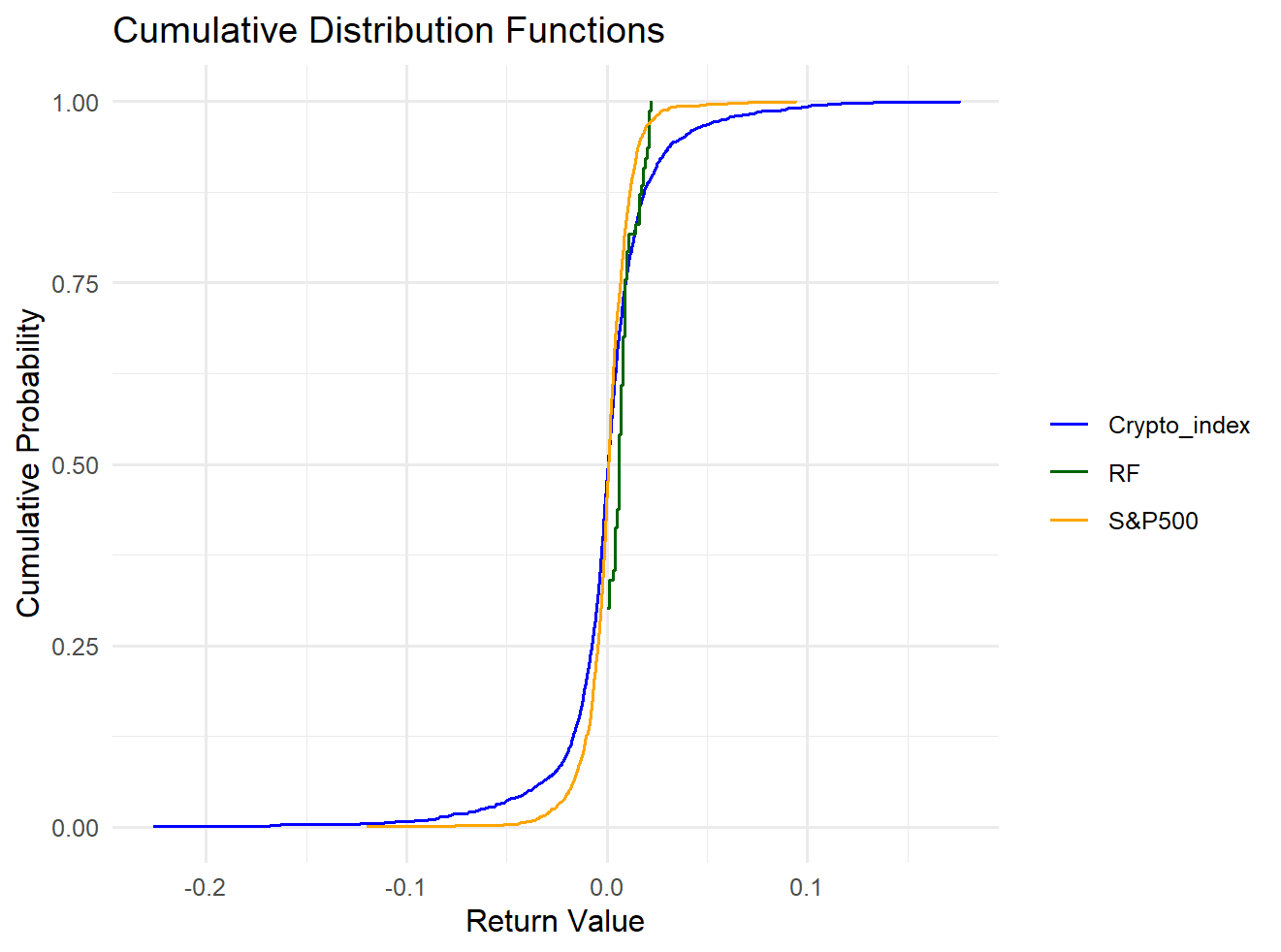


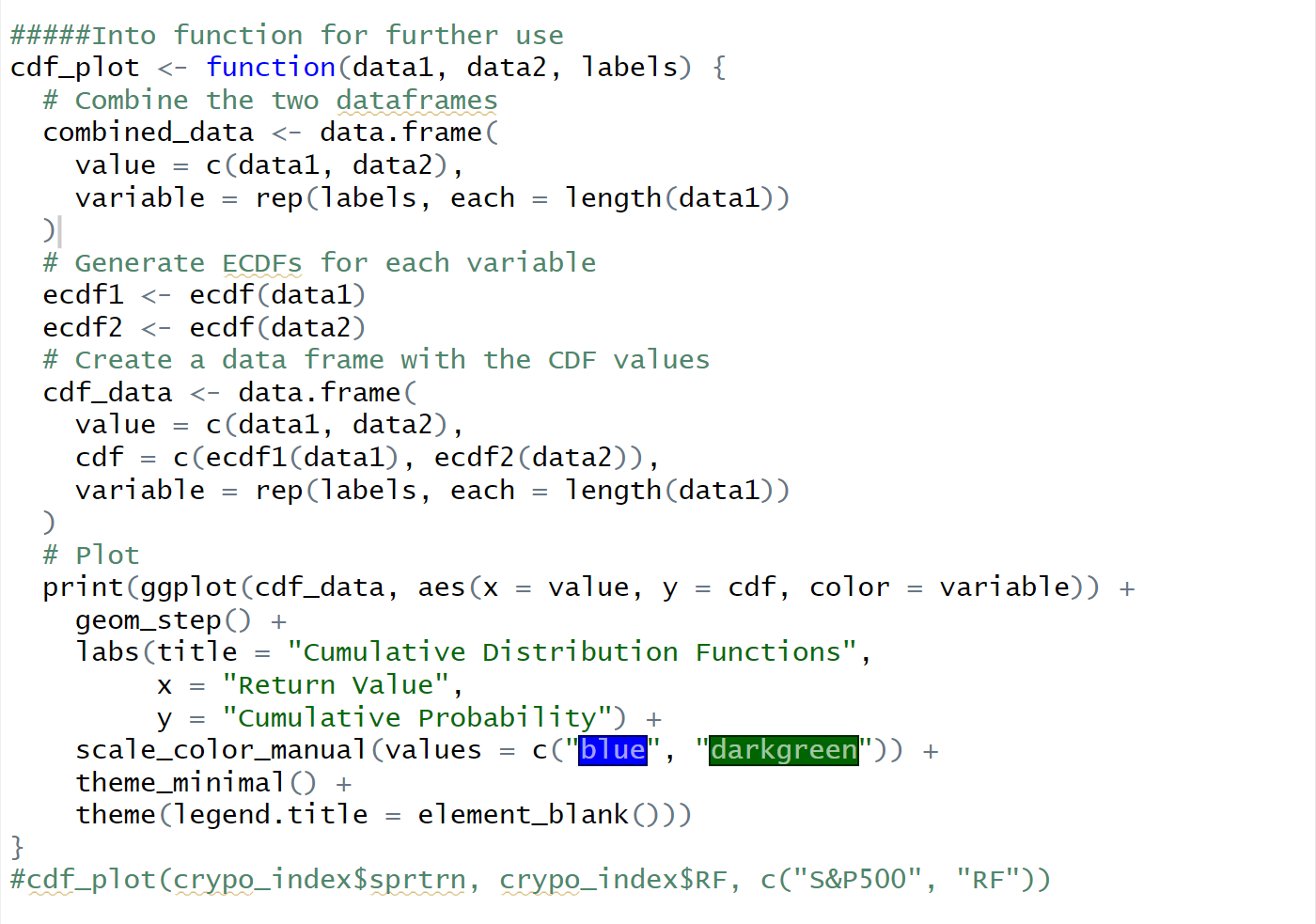
**5.2.2 ECDFs Calculating**

Then calculate the Empirical Cumulative Distribution Functions (ECDFs) for the S&P 500 returns (sp\_cdf), risk-free rate (rf\_cdf), and the crypto index returns (crypto\_cdf). An ECDF function returns the cumulative probability up to a given value in the input data. For example, for a value x, the ECDF function returns the proportion of data points that are less than or equal to x.

The plot compares the distributions of returns for the crypto index, S&P 500, and the risk-free rate.

By observing the shape of the CDFs, you can gauge which asset is riskier (steeper increases indicate more concentrated values), which has more variability, and how they perform relative to each other.

Then introduce the cdf\_plot function as a resuable tool for comparing the distribution of two different sets of data. This helps our later modeling of AFSD and ASSD.



**5.2.3 AFSD & ASSD**

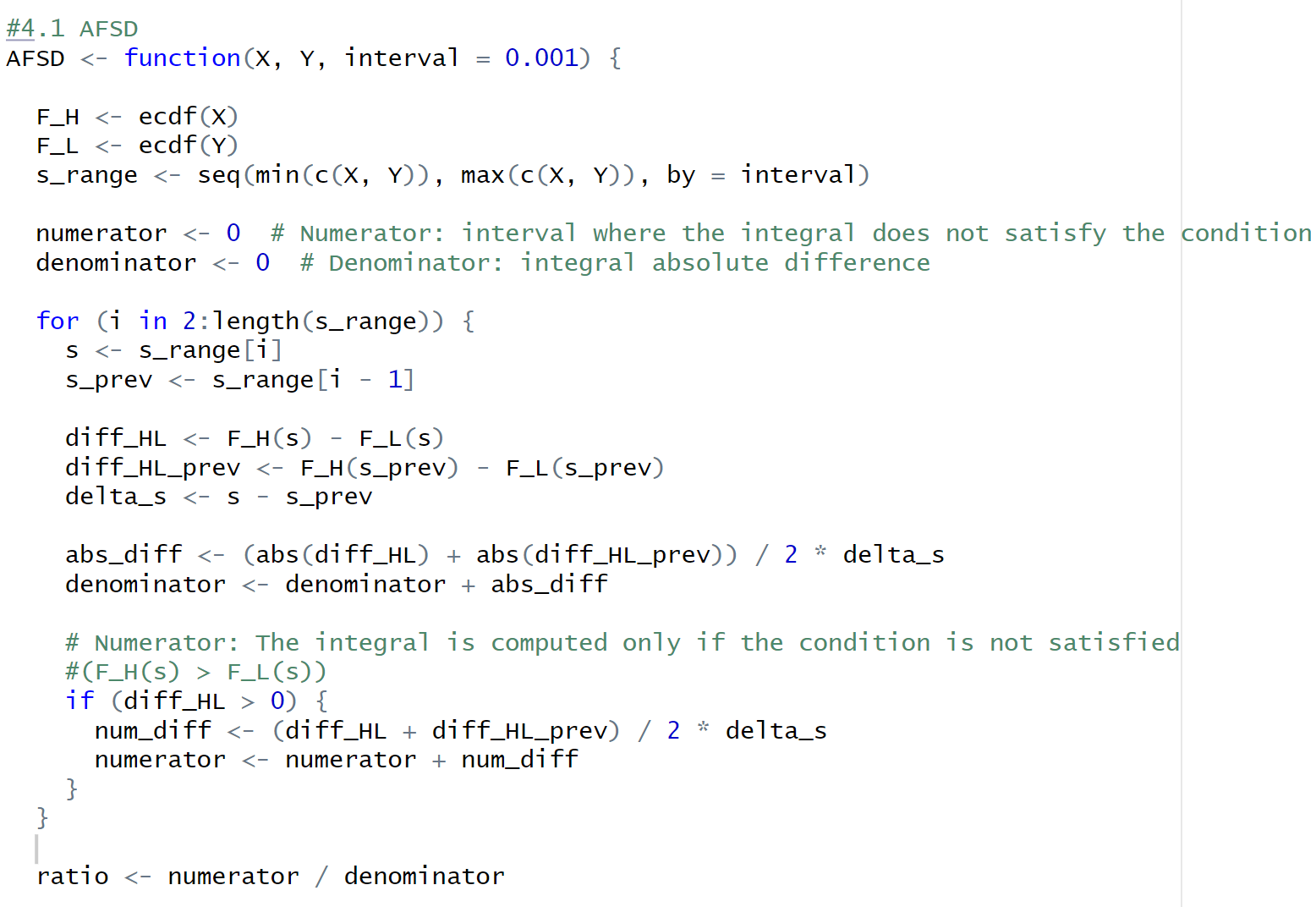
Almost Stochastic Dominance is a concept used in decision theory and finance to compare distributions of returns or outcomes, allowing for some tolerance of violations in the preference ordering. To test whether AFSD and ASSD exist, we need to get the ratio of violation area over total enclosed area for each portfolio and compare if ratios 𝜀1(AFSD) and 𝜀2(ASSD) are lower than AFSD (5.9%) and ASSD (3.2%), showing that they meet the requirement.

For AFSD,

FSD violation range is given by:

The empirical violation area is defined as:

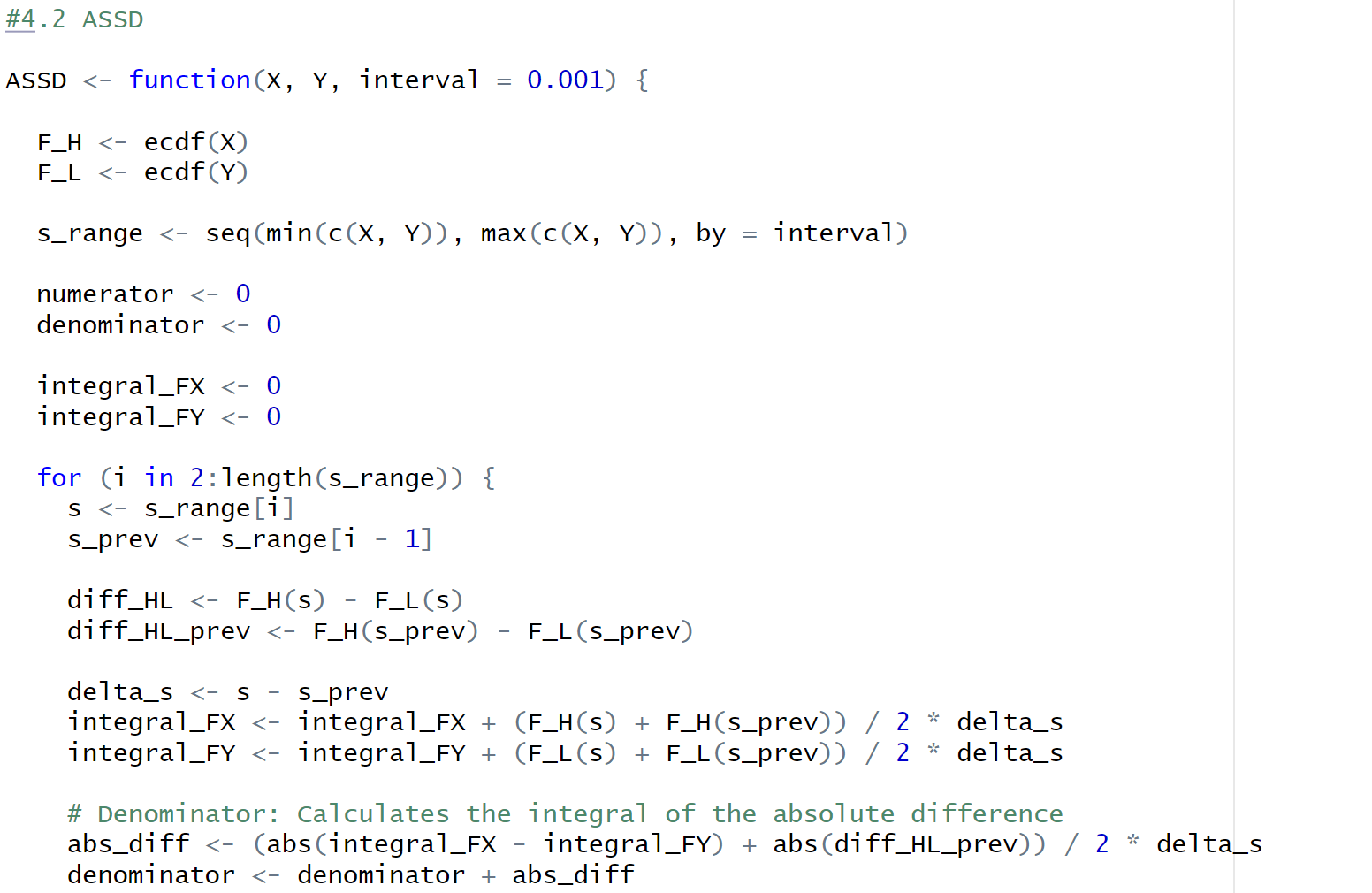
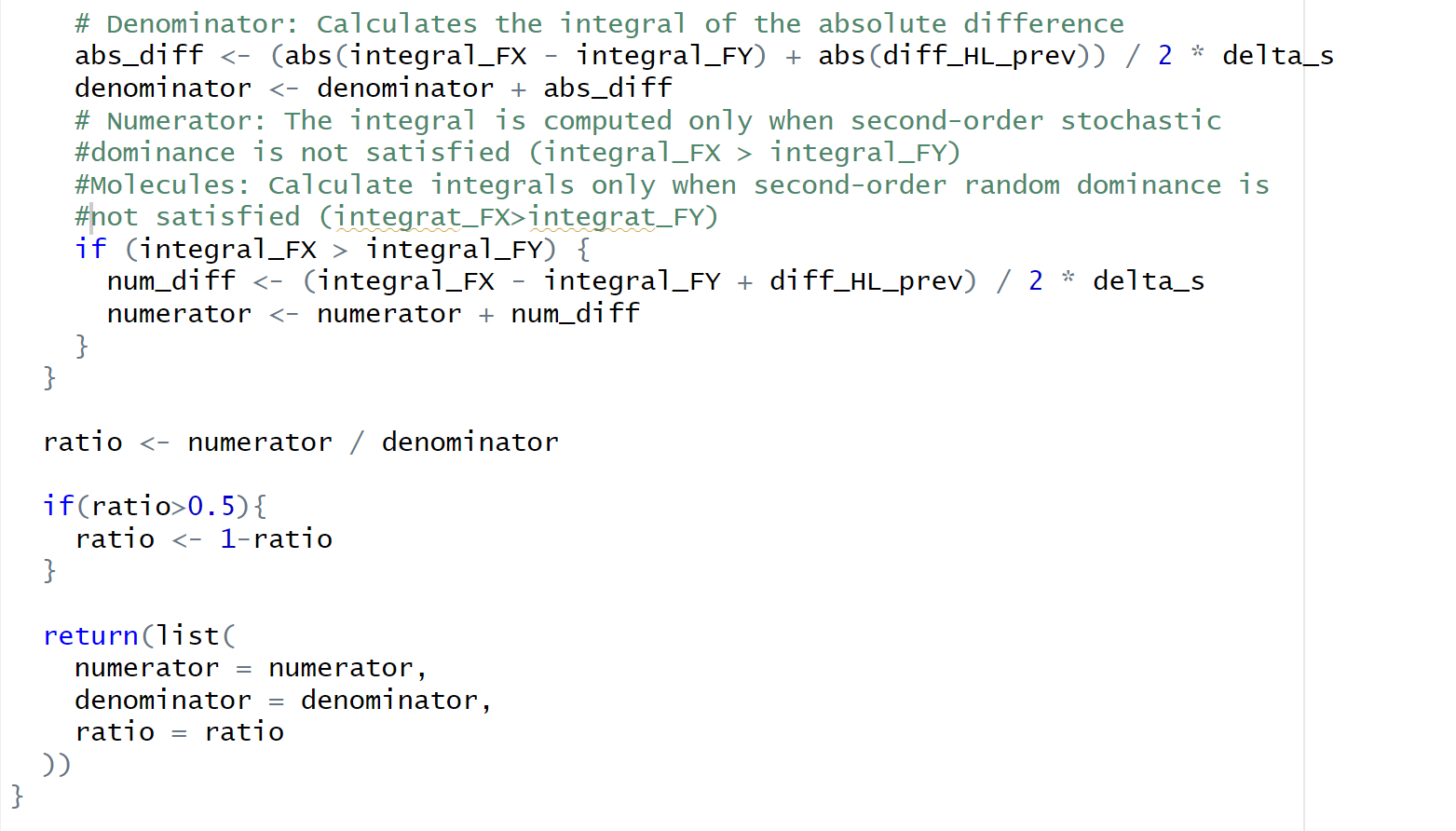
If 𝜀1 <5.9%, AFSD exits.

For ASSD,

SSD violation range is given by:

The empirical violation area is defined as:

If 𝜀2 <3.2%, ASSD exits.



**5.2.4 Factor Calculation**

Depend on the previous 22 factor of the main aspects of size, momentum, volume and volatility, asset characteristics are calculated.

**Momentum Factors:**

Momentum measures the rate of price change over time and is calculated using lagged values of the asset prices.

* **Lagged Values:** Lag1, Lag2, Lag3, Lag4, Lag8 represent the asset's lagged prices over different time periods (7, 14, 21, 28, and 56 days).
* **Momentum (mom):** Calculated as the difference between the current price and the lagged prices over different time horizons (mom1, mom2, mom3, mom4, mom8). This shows how much the asset's price has changed over those time intervals.

**Volatility Factors:**

Volatility factors measure the risk and dispersion of returns over various time periods.

* **Standard Deviation (std):** Calculated over rolling windows of 7, 14, 21, 28, and 56 days (std1, std2, std3, std4, std8). This provides an estimate of how much the price fluctuates within the specified period.
* **Risk-Adjusted Momentum (rmom):** Momentum is adjusted by dividing by the corresponding standard deviation to account for risk. These are also calculated over different time windows (rmom1, rmom2, rmom3, rmom4, rmom8).

**Size Factors:**

The size of an asset is generally measured by its market capitalization (CapMkt), which is the product of price and the number of shares outstanding.

* **Log-Price (lprc):** The natural logarithm of the price to stabilize the variance.
* **Max Price (maxprc):** The maximum price observed over a rolling 7-day window.

**Volume Factors:**

Volume factors reflect the trading activity and liquidity of the asset.

* **Log-Volume (vol):** The natural logarithm of the average volume over a rolling 7-day window.
* **Volume-Price Product (volprc):** The product of volume and price, representing the total traded value.
* **Volume-Scaled Price (volscale):** Volume-price product divided by market capitalization to assess the liquidity relative to the asset size.

The process will compare the portfolios based on ASSD calculations to evaluate if one strategy or factor dominates another.

**Quantile:**

Quantile formation involves sorting assets based on factor values and dividing them into equal-sized groups. This method is used to classify assets according to their factor characteristics.

For each factor, the assets are divided into five quantiles, and this categorization helps analyze the behavior of portfolios constructed from each quantile.

By categorizing the assets into quantiles for each of these factors, the code enables analysis of portfolio returns, risks, and other metrics. Quantile-based sorting allows for examining if certain factor characteristics (e.g., high momentum or low volatility) result in superior performance

**5.2.5 Portfolio Building**

We then analyze financial portfolios, focusing on constructing portfolios based on different factors such as size, momentum, volatility, and volume. Portfolio returns are calculated using equal-weighted and value-weighted methods, with the ultimate goal of assessing the performance of different quantile-based portfolios.

**Step 1: Portfolio Building**

Build various portfolios based on different factors. Combine the size, momentum, volatility, and volume data with the future returns.

**Step 2: Equal-Weighted Portfolio Calculation**

The function calculate\_portfolio\_returns() calculates portfolio returns based on equal weights.

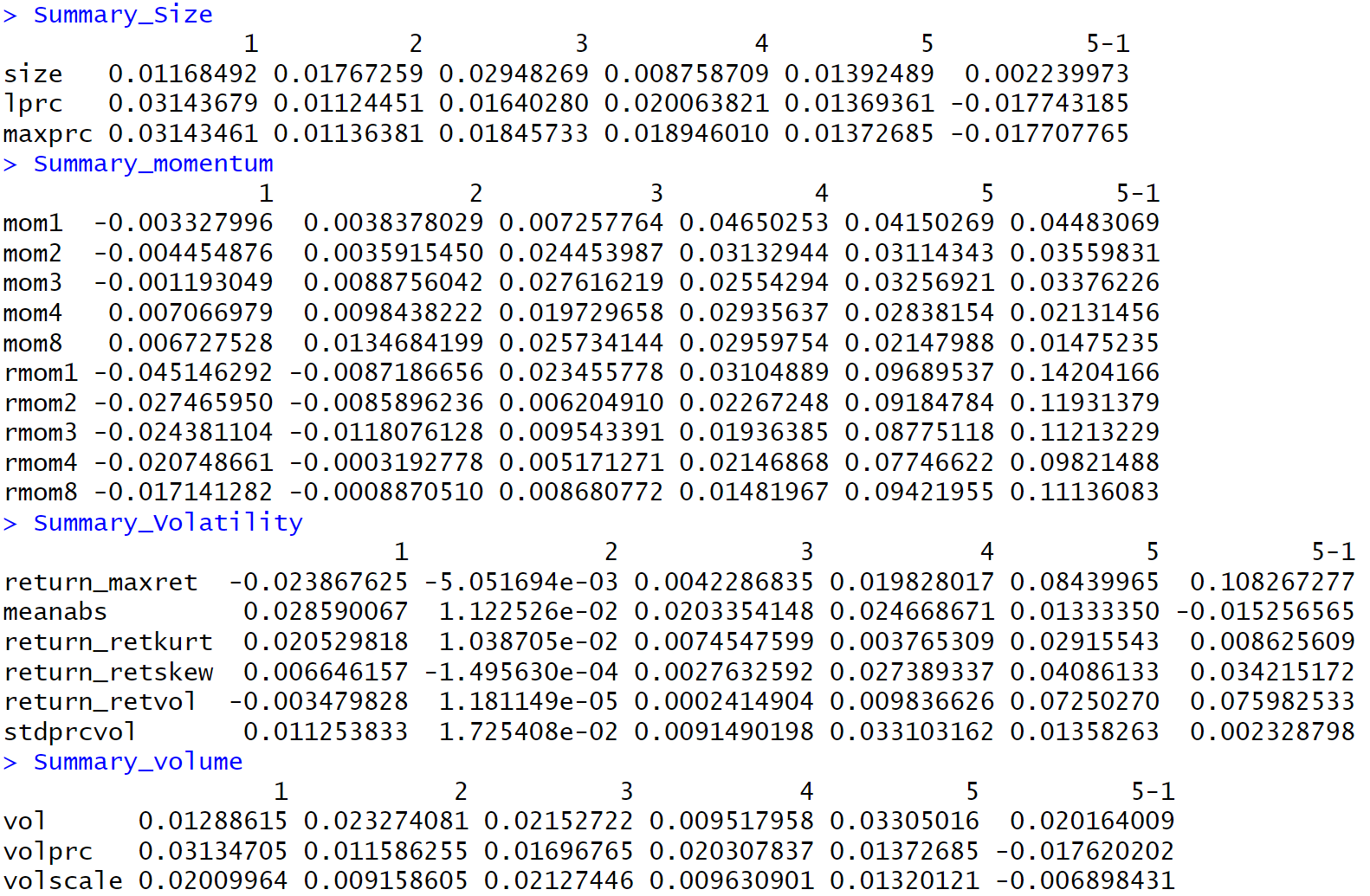
**Step 3: Value-Weighted Portfolio Calculation**

The function value\_weighted\_factor\_portfolio() calculates returns for a value-weighted portfolio as the replicated paper did. The weights of each asset (weights\_port) are calculated by dividing the market value of each asset by the total market value for the period. This normalization ensures that the sum of the weights equals 1. If holding\_period is 1, it uses the default returns (e.g., daily or weekly returns). For longer holding periods, it uses cumulative returns over that period.

This approach ensures that larger assets (by market value) have a more significant impact on the portfolio's returns, following a value-weighted methodology.

**Step 4: Summary Calculations for Different Factors**

Calculates the average returns for different quantiles (1 to 5) across various factors.

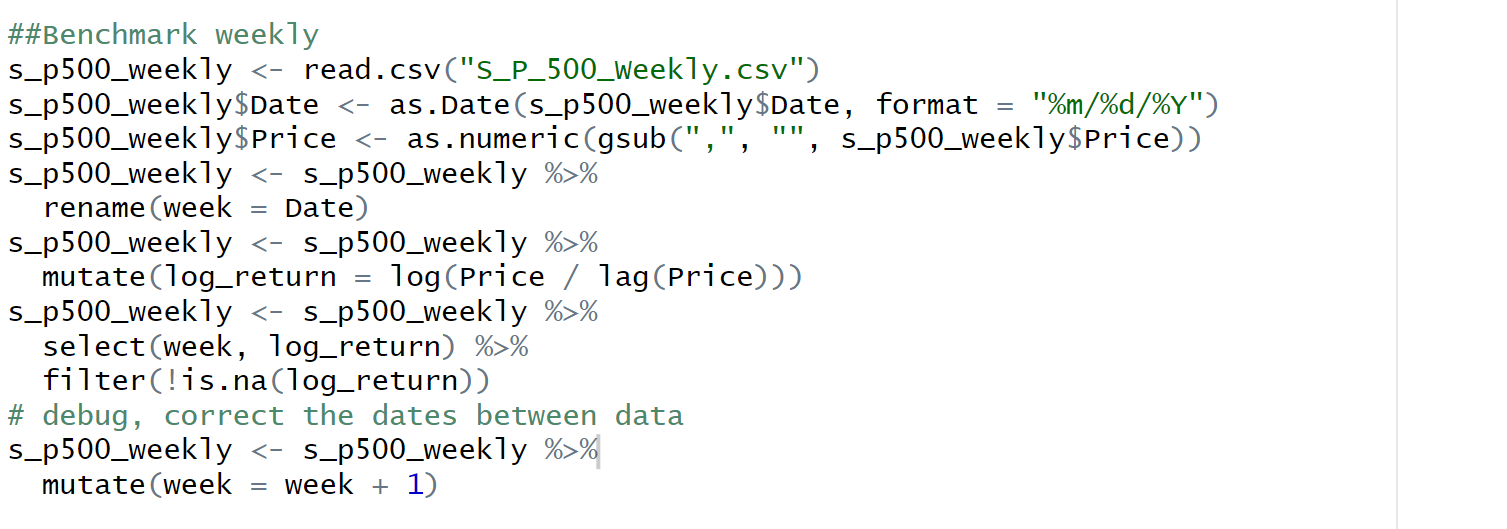


**5.2.6 AFSD and ASSD testing and plotting**

Test and analysis for Almost First-Order and Second-Order Stochastic Dominance (AFSD and ASSD) across different factors and holding periods, using log returns for the S&P 500 index as a benchmark.

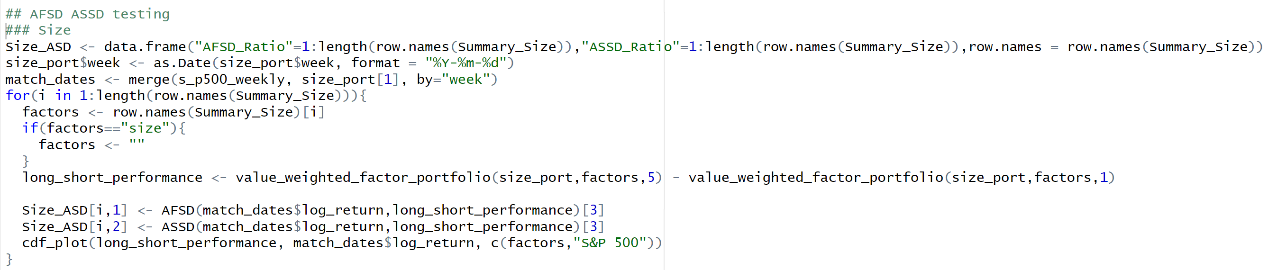
**Step 1: Set the benchmarks of Risk-free and A&P500**

Read and prepare weekly log return data for the S&P 500. It converts the Date column to a Date format, computes the weekly log returns, and adjusts the dates.

**Step 2: AFSD and ASSD Testing**

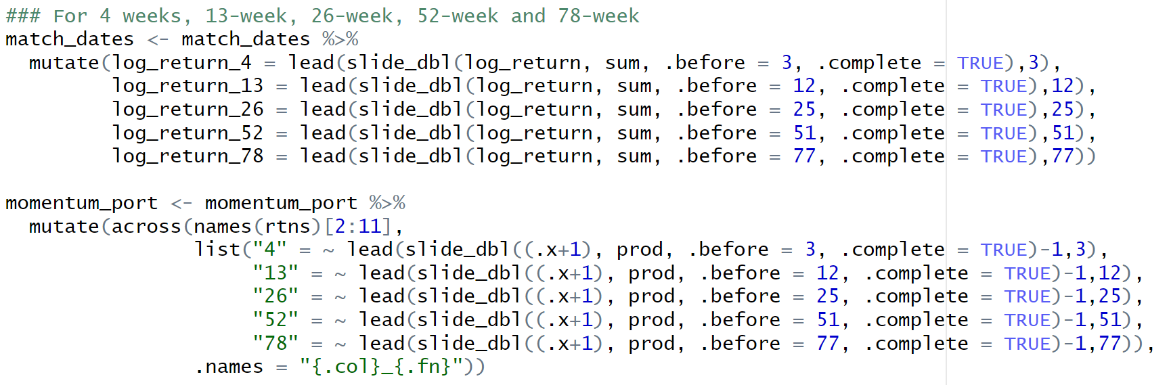
Compare different factor-based long-short portfolio returns with the S&P 500 returns. It uses AFSD and ASSD to measure the dominance relationships between the long-short portfolio performance and the benchmark.

For example, the AFSD and ASSD of size is calculated as follows:

Do the same thing for momentum, volatility and volume.

**Step 3: Factor Analysis by Holding Periods**

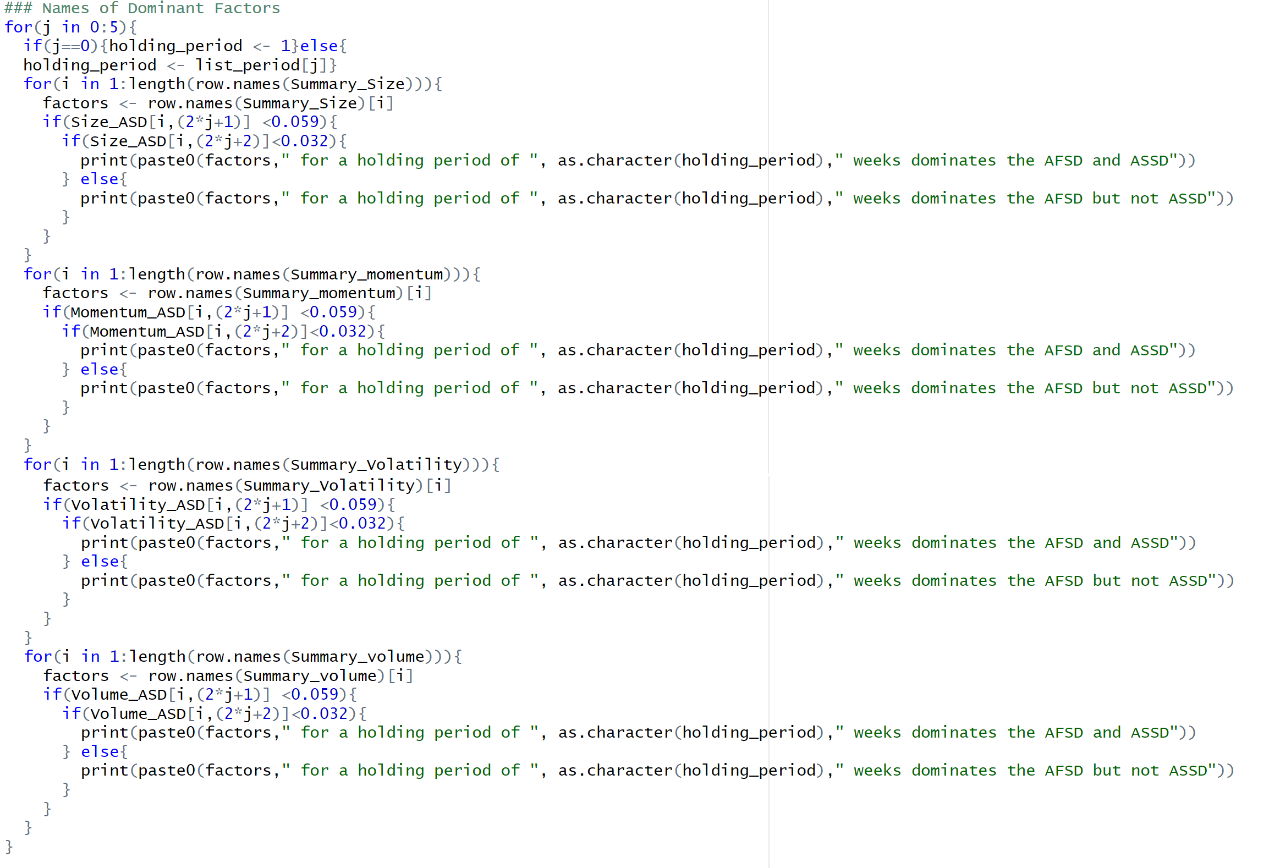
Continue to analyze portfolio performance over different holding periods (4 weeks, 13 weeks, 26 weeks, 52 weeks, and 78 weeks).

For example, apply the different holding periods on momentum.

Do the same thing for size, volatility and volume.

**Step 4: Identifying Dominant Factors**

Evaluate whether factors exhibit dominance according to AFSD and ASSD criteria. Use specific thresholds (5.9% for AFSD and 3.2% for ASSD) to determine dominance.



**5.3 Hypothesis Tests**

After replicating the initial work, it is time to evaluate the hypotheses of the replicated work.

We perform hypothesis testing using Almost First-Order Stochastic Dominance (AFSD) and Almost Second-Order Stochastic Dominance (ASSD) to analyze the performance of different investment strategies.

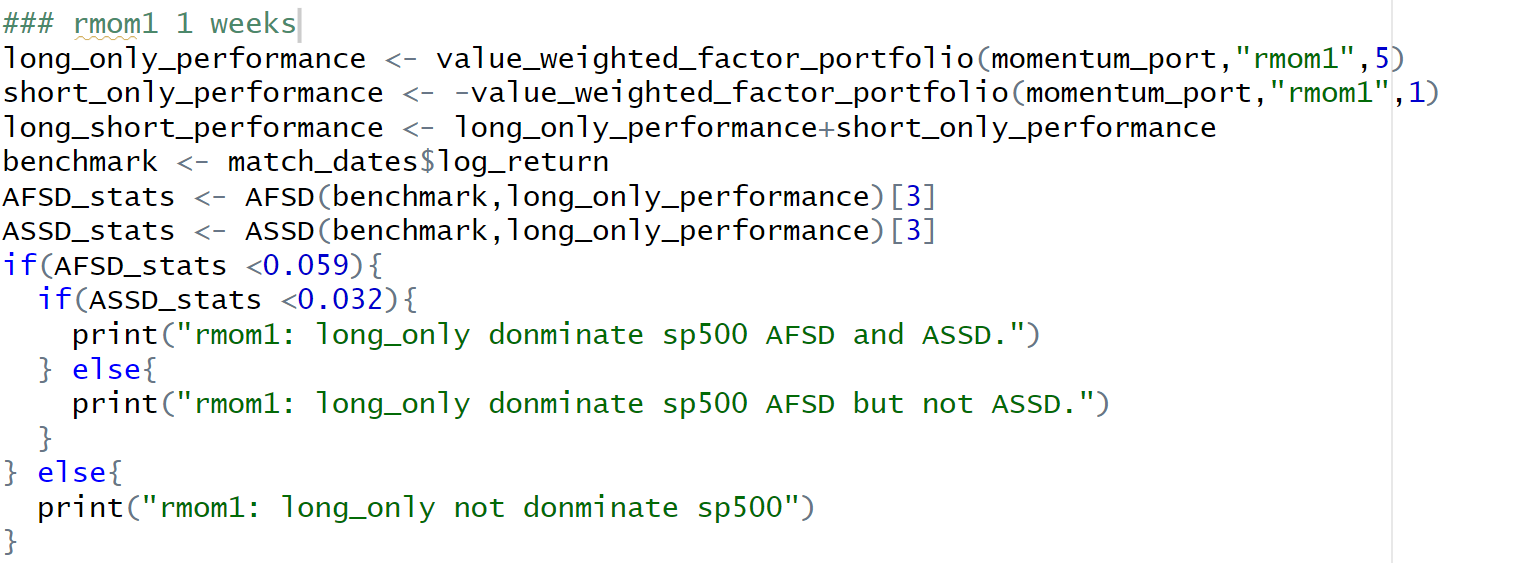
**5.3.1 AFSD and ASSD Testing**

**Step 1: Factor-Based Portfolio Construction**

First, create factor-based long-only and short-only portfolios.

For example, it constructs portfolios based on momentum factors such as rmom1, rmom2, rmom3, rmom4, and rmom8, representing relative momentum over different time periods.

It uses the value\_weighted\_factor\_portfolio() function to create a long-only portfolio by investing in assets in the top quantile (5) and a short-only portfolio by shorting assets in the bottom quantile (1).



**Step 2: Long-Short Portfolio Creation:**

The long-short portfolio is calculated by adding the returns of the long-only portfolio and the negative returns of the short-only portfolio.

**Step 3: Benchmark Data Preparation:**

The benchmark data used for comparison is the log returns of the S&P 500 index, stored in the benchmark variable.

**Step 4: AFSD and ASSD Testing:**

The AFSD() and ASSD() functions calculate the Almost Stochastic Dominance statistics for various combinations of portfolios and the benchmark.

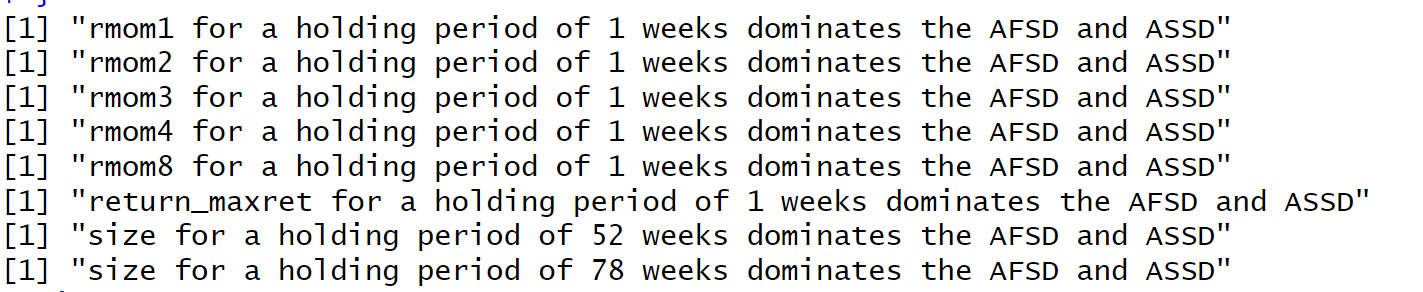
Specifically, the tests are conducted to determine whether:

* The long-only portfolio dominates the benchmark.
* The short-only portfolio dominates the benchmark.
* The short-only portfolio dominates the long-only portfolio.
* The short-only portfolio dominates the long-short portfolio.
* The long-short portfolio dominates the long-only portfolio.

**Step 5: Decision Criteria for Stochastic Dominance:**

If the AFSD statistic is less than 0.059, it indicates that the tested portfolio might dominate the comparison portfolio based on AFSD. If the ASSD statistic is less than 0.032, it indicates dominance based on ASSD.

Based on these thresholds, the code prints statements about the dominance relationships between portfolios.

**Analysis of the Results:**

* **Short-Term Dominance (1 Week):**

The momentum factors rmom1, rmom2, rmom3, rmom4, and rmom8 dominate both AFSD and ASSD for a 1-week holding period. This suggests that momentum strategies are highly effective in generating superior returns over short-term horizons.

The volatility factor return\_maxret also dominates AFSD and ASSD in the 1-week holding period. This indicates that maximum return volatility strategies outperform the S&P 500 in the very short term.

* **Long-Term Dominance (52 and 78 Weeks):**

The size factor dominates both AFSD and ASSD for holding periods of 52 weeks and 78 weeks. This suggests that size-based investment strategies yield better long-term performance compared to the benchmark.

Step 6: Testing Across Different Factors:

The code repeats the process for various momentum factors (rmom1 to rmom8) and size-based portfolios (size\_52\_weeks and size\_78\_weeks), considering different holding periods (52 weeks and 78 weeks).

**Step 6: Hypothesis 2 Testing**

Since the above factors all dominates the AFSD and ASSD, a further testing is used to form strategies to evaluates different trading strategies' performance using AFSD and ASSD criteria. The strategies tested include long-only, short-only, and long-short portfolios, which are compared to the benchmark (S&P 500) and to each other.

* **Calculate Performances**:

long\_only\_performance: The performance of a long-only strategy, which involves buying assets in the highest quantile (top 20%) for the factor rmom1.

short\_only\_performance: The performance of a short-only strategy, which involves shorting assets in the lowest quantile (bottom 20%) for the factor rmom1.

long\_short\_performance: The combination of both strategies, calculated as the sum of long-only and short-only performances.

* **AFSD and ASSD Tests**:

For example, for rmom1

**rmom1: Long-Only vs. S&P 500**:

"rmom1: long\_only not dominate sp500": The long-only strategy does not dominate the benchmark for both AFSD and ASSD, indicating that it does not outperform the S&P 500 based on the stochastic dominance criteria.

**rmom1: Short-Only vs. S&P 500**:

"rmom1: short\_only not dominate sp500": The short-only strategy also fails to dominate the benchmark.

**rmom1: Short-Only vs. Long-Only**:

"rmom1: short\_only dominate long\_only AFSD and ASSD": The short-only strategy dominates the long-only strategy for both AFSD and ASSD criteria. This suggests that shorting assets in the lowest quantile performs better than buying assets in the highest quantile for this factor over the 1-week period.

**rmom1: Short-Only vs. Long-Short**:

"rmom1: short\_only dominate long\_short AFSD and ASSD": The short-only strategy also outperforms the combined long-short strategy based on both stochastic dominance criteria.

**rmom1: Long-Short vs. Long-Only**:

"rmom1: long\_short dominate long\_only AFSD and ASSD": The combined long-short strategy dominates the long-only strategy, suggesting that a mixed approach of both buying high-quantile assets and shorting low-quantile assets yields better performance than a pure long strategy.

Do the same thing for the rest of the factors. The hypothesis is tested across various factors, including different momentum measures (rmom1, rmom3, rmom4, rmom8) and a size factor over longer holding periods (52 and 78 weeks).

**Results:**

**1. Momentum Factors (rmom1, rmom3, rmom4, rmom8):**

* **Long-Only vs. S&P 500**:

For all momentum factors tested (rmom1, rmom3, rmom4, rmom8), the long-only strategy does **not dominate** the S&P 500 for both AFSD and ASSD, indicating that a long-only approach may not outperform the benchmark.

* **Short-Only vs. S&P 500**:

Similarly, the short-only strategy fails to dominate the S&P 500, suggesting it does not provide better returns than the benchmark.

* **Short-Only vs. Long-Only**:

In each case, the short-only strategy **dominates the long-only strategy** for both AFSD and ASSD criteria, suggesting that shorting low-ranked assets yields better performance than buying high-ranked assets in the momentum context.

* **Short-Only vs. Long-Short**:

The short-only strategy also **dominates the long-short strategy** across all momentum factors, indicating that pure shorting is more effective than combining long and short positions.

* **Long-Short vs. Long-Only**:

In some cases (like rmom1), the long-short strategy dominates the long-only strategy, suggesting that including short positions improves performance over a pure long strategy.

**2. Size Factor for Longer Holding Periods:**

* **52 Weeks**:

The long-only strategy **dominates the S&P 500** in both AFSD and ASSD for a 52-week holding period, indicating strong performance.

* **Short-Only vs. S&P 500**:

The results for short-only dominance were not provided, but it can be inferred that it would not outperform the S&P 500 based on previous patterns.

For short-term momentum factors (rmom1, rmom3, rmom4, rmom8), the **short-only strategy consistently outperforms both the long-only and long-short strategies**, suggesting that shorting low-performing assets is more effective than buying high-performing assets.

The **long-short strategy** outperforms the **long-only strategy** in some momentum cases, indicating that including a short component can improve returns.

Over a **52-week holding period**, the **size factor exhibits dominance over the benchmark**, showing that longer-term strategies based on size can yield superior results.

The **momentum-based long-only strategies** do not outperform the benchmark, indicating the limitations of these strategies when used alone for short-term periods.

These results highlight the importance of strategy selection and factor timing in portfolio management, where momentum shorting may be effective in the short term, while size factors offer potential over longer horizons.

**7. Conclusion**

This analysis utilizes non-parametric almost stochastic dominance (AFSD and ASSD) to evaluate the performance of various portfolio strategies based on different factors, including momentum and size, to determine if they can outperform a benchmark of the S&P 500 over different time horizons (1 to 78 weeks). Our findings reveal dominant factors in the short term (1-week) for momentum-based strategies (rmom1, rmom3, rmom4, and rmom8), where short-only portfolios consistently dominate long-only portfolios in terms of both AFSD and ASSD criteria. This suggests that shorting low-ranked assets is more effective than buying high-ranked assets in momentum strategies. Additionally, we find that long-short strategies generally perform better than long-only strategies, particularly for short-term momentum factors.

In the longer term (52 and 78 weeks), the size factor exhibits significant dominance over the benchmark, with long-only portfolios outperforming the S&P 500 under both AFSD and ASSD. This indicates that size-based strategies may be more effective over extended horizons. However, for shorter momentum strategies, long-only approaches do not dominate the benchmark, highlighting limitations in using such strategies in isolation.

Our analysis supports the value of incorporating short positions in portfolio management, particularly for momentum strategies, and suggests that size factors can be effective for long-term investment horizons. However, further research is warranted to explore the performance of additional factors, the impact of different market conditions, and alternative risk metrics to enhance understanding of these strategies.