**Background**

Previously, regarding the concept of S&P 500, people could have heard about it but may not have had the opportunity to get a deeper and intuitive understanding of its constitutions. However, our project aims to solve the problem and create an interactive dashboard or tool that enables users to analyze and visualize portfolio performance, factor exposures, and key risk metrics, thus providing actionable insights for portfolio managers, analysts, and investors by incorporating advanced performance metrics, factor analysis, and historical trends. In the future, we may not be limited to the S&P 500 index alone, but our product can be enriched into a worldwide one-stop platform and become a powerful financial investment assistant for different levels of investors.

**Introduction**

The codebase for our project is extensive and complex, necessitating a modular structure to ensure clarity and maintainability. At its core, combined999\_v3.py serves as the main program file, orchestrating the overall workflow and user interface through Streamlit. To enhance organization, functional components are systematically separated into dedicated files, forming a cohesive package. This modular approach not only streamlines development and debugging but also allows developers to focus on specific functionalities without navigating monolithic code. Each file is designed to handle a distinct responsibility - such as user input processing, performance metric calculations, or AI integration - while communicating seamlessly through well-defined interfaces. By structuring the code as a package of specialized modules, the project remains scalable, with each component easily extendable or replaceable as needed, ensuring long-term maintainability and collaboration efficiency.

**Dealing with users’ inputs**

We allow both file upload and manual typing to let users create a portfolio using S&P 500 constituents. Regarding the file uploaded, users need to follow certain instructions regarding the format of the data organization, and we accept both CSV and XLSX versions. For manual input, we have enabled adding, deleting, and refreshing input rows. However, currently we don’t consider the short stock scenario, so the weight of each stock should be within a range of 0 to 100%. And if the user doesn’t input any weight into some stocks, then the system will automatically allocate the remaining part equally. However, the sum of existing weights should not exceed 100%, otherwise there will be a warning. Moreover, to locate the target stock, both the stock ticker and company’s name is acceptable, and we have enabled fuzzy search and autocomplete functionality to tolerate some extent of mistakes and raise the convenience of searching.

**Archive for investment portfolios**

We also have a memory function to enable the archive of investment portfolios utilizing JSON (JavaScript Object Notation) files, so that users may select the portfolio they want and don’t need to re-enter each time they open the website. Currently, the maximum number of the archive is 10, but it can be modified by changing the underlying code if PremiaLab wants. We list all the archives on the main dashboard page, and users can view the detail constructions of the portfolio without processing the following analysis. Similarly, users can create, edit, rename or delete the archive following the same construction rules, and they can even see the construction period as well as the last modified time to awaken their memories. We also create a one-click removal function for refreshing and updating.

**Portfolio Composition Analysis**

This tab provides a comprehensive breakdown of portfolio constituents and their characteristics through multiple analytical components:

1. ​Portfolio Constituents Table

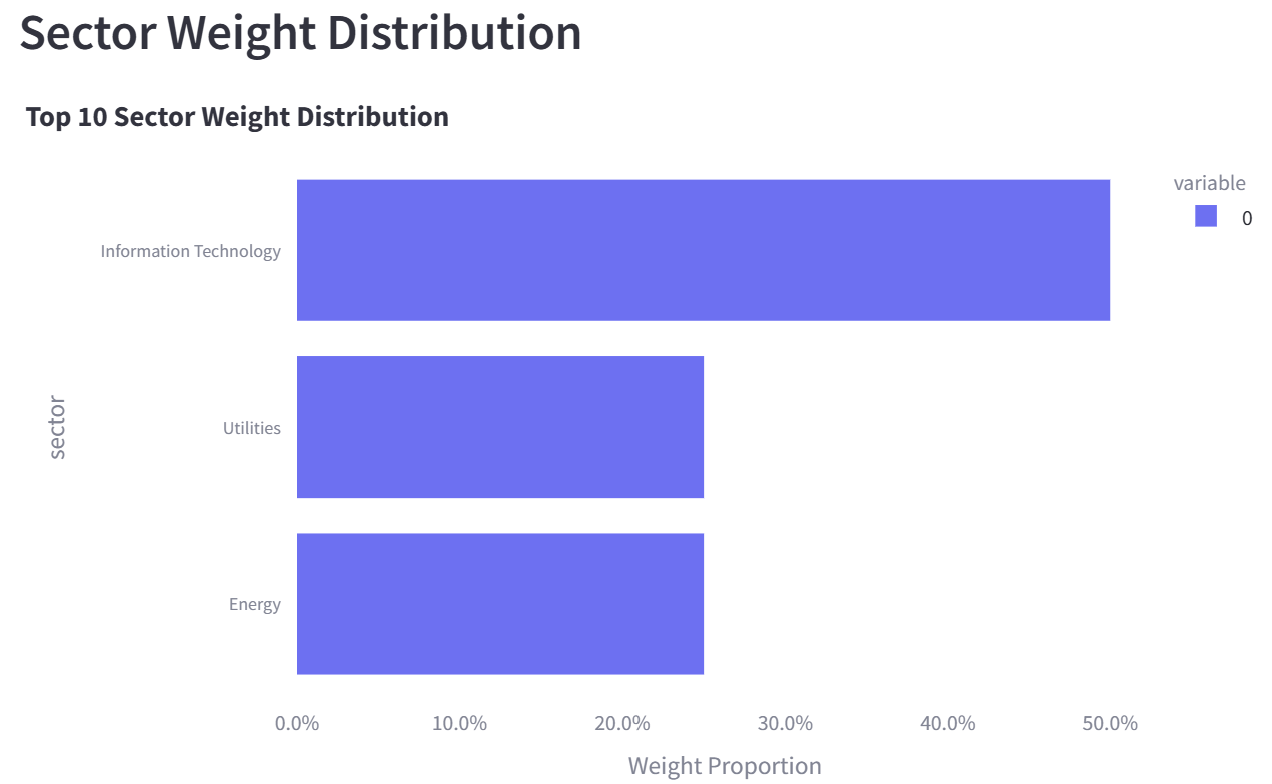
The system generates a structured DataFrame displaying all tickers with their corresponding weight allocations (formatted as percentages). The implementation (*portfolio\_df*) demonstrates precise weight formatting to two decimal places, ensuring transparency in position sizing.



Utilizing static reference data (*static\_data*), the module constructs a detailed information panel for each constituent. The implementation features robust error handling (try-except blocks) to gracefully manage missing data scenarios, maintaining system stability when ticker information is unavailable.

2. ​Sector Weight Distribution Visualization

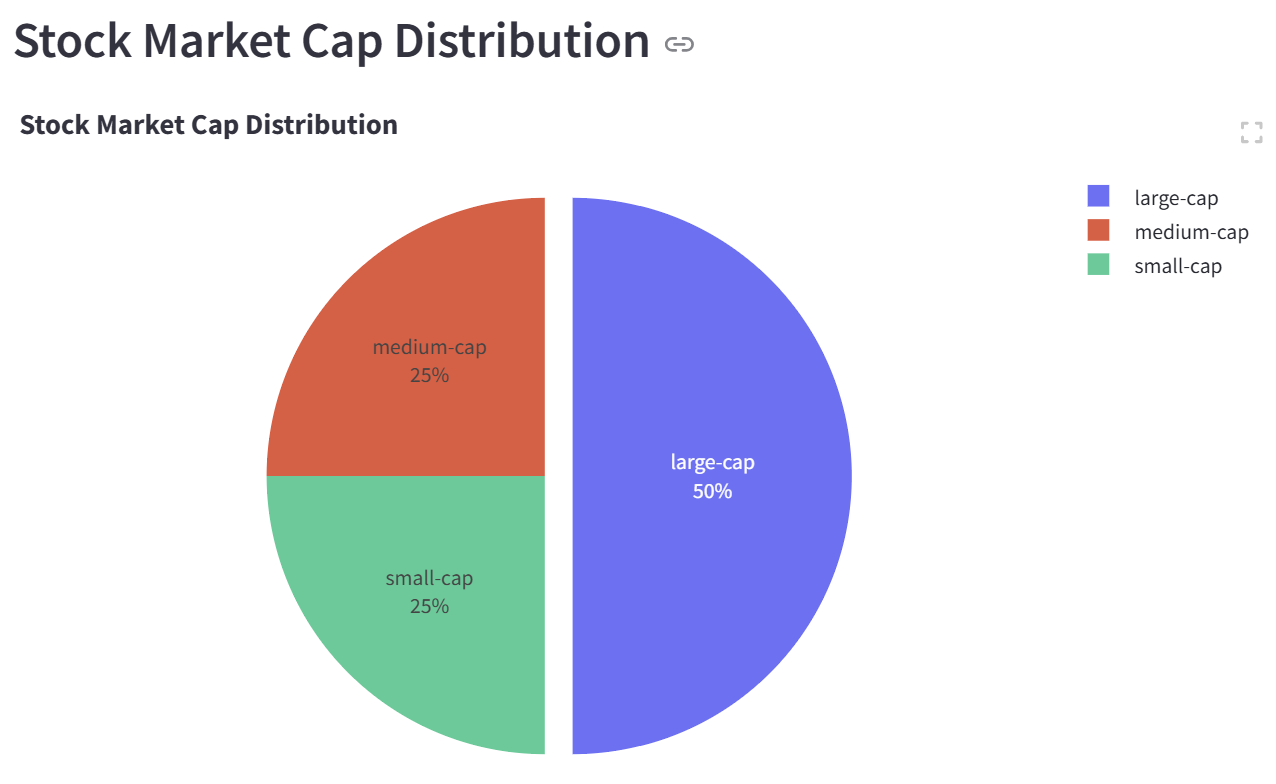
The sector analysis employs a horizontal bar chart (*Plotly* Express) to display the top 10 sector exposures. Here are the key features: Dynamic height adjustment based on sector count ; Percentage-formatted x-axis for intuitive weight interpretation; Customized margin parameters ensuring label visibility.



The visualization highlights concentration risks through a monochromatic blue color scheme for consistent visual interpretation. (Plot shows an example of visualization)

3. ​Market Capitalization Profile

This part includes a pie chart representation of market cap categories (large/mid/small-cap), which provides this information:1) Weight-based aggregation of capitalization buckets; 2) Automatic index alignment between weights and cap categories; 3) Clear visual segmentation of portfolio capitalization structure.

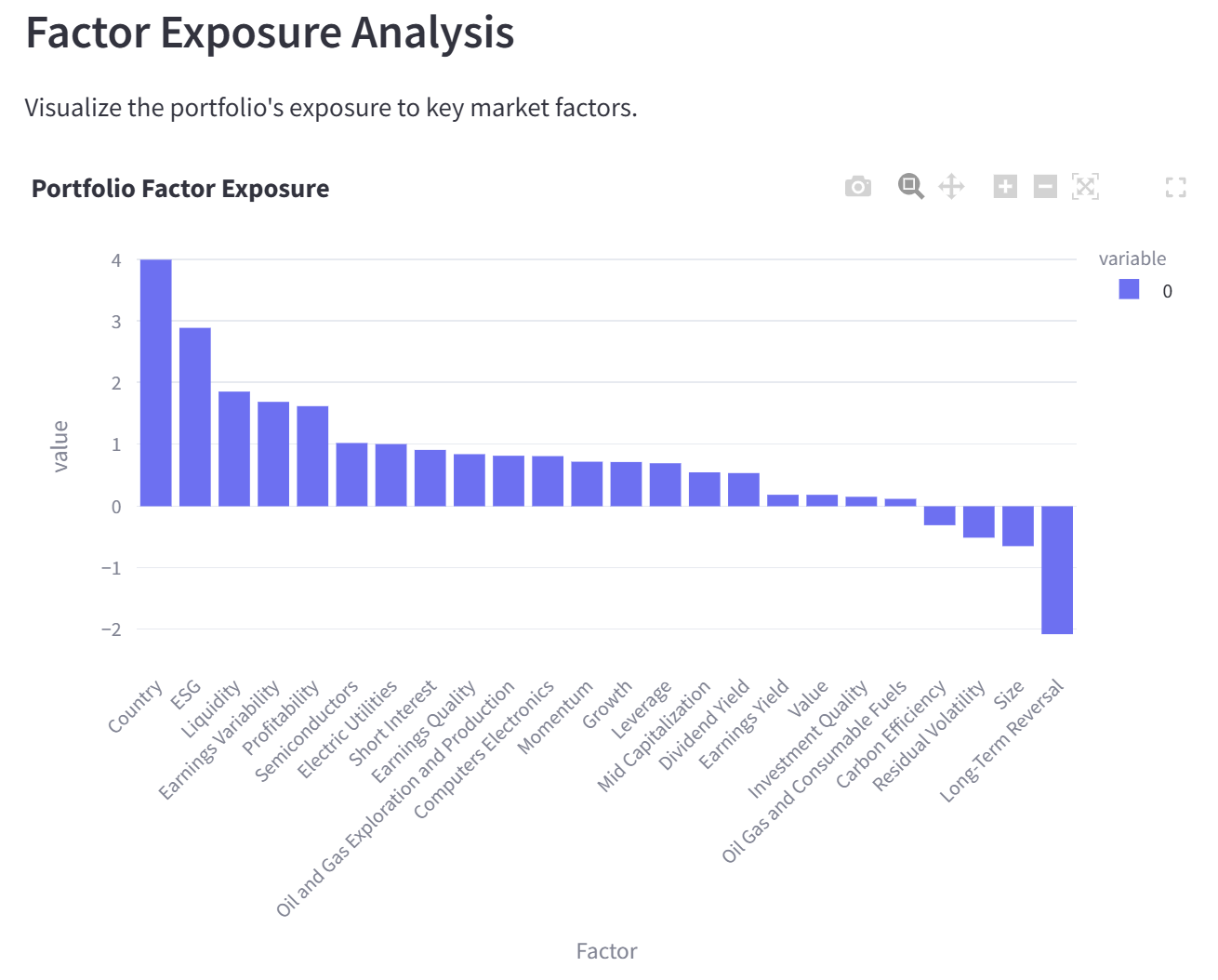


**Factor Risk Analysis**

This tab delivers advanced factor risk diagnostics through multivariate statistical visualization:

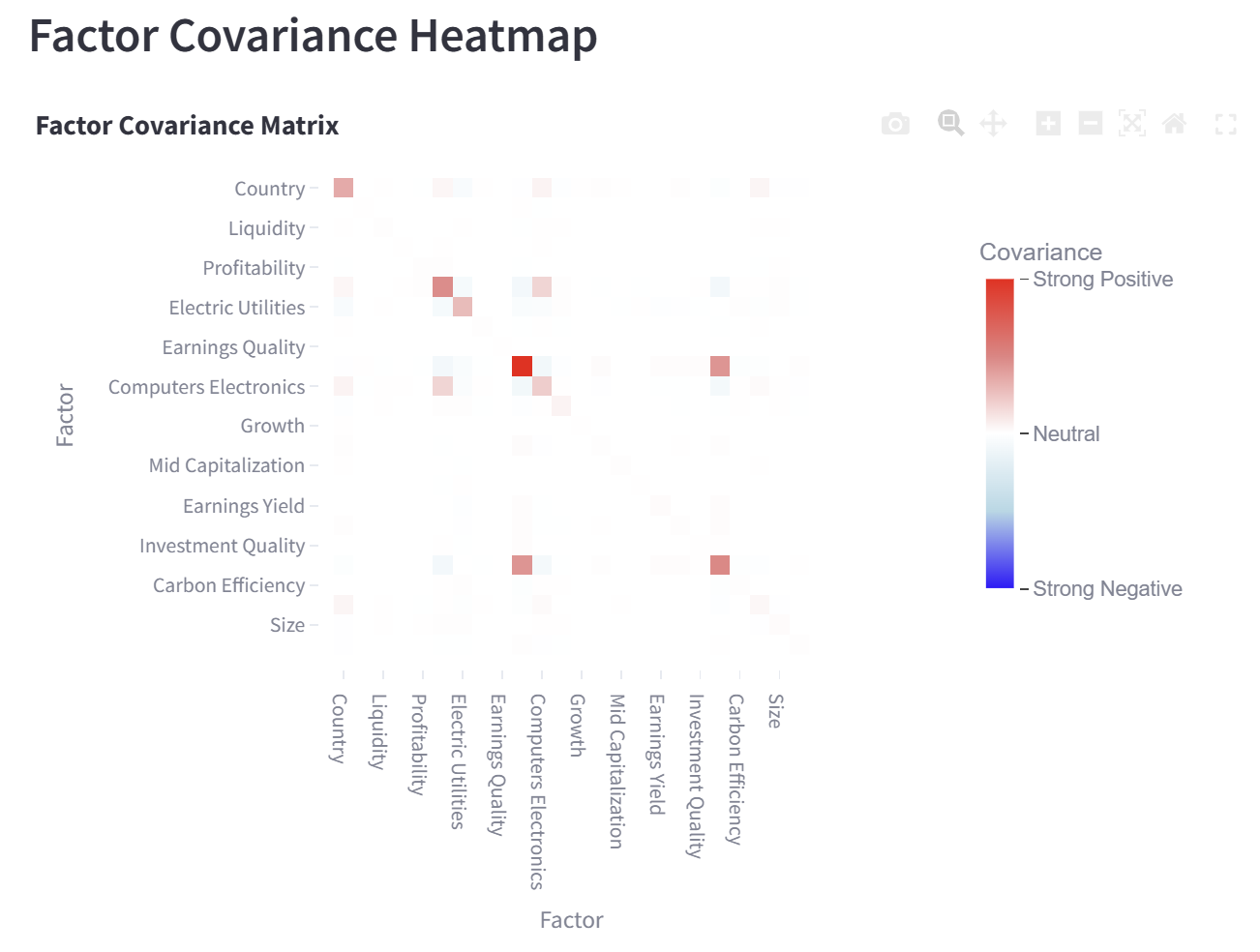
​1. Factor Exposure Bar Chart

The analytical engine (*factor\_exposure\_analysis function*) performs ticker-specific factor exposure aggregation when constituents are specified. After automatically filtering zero-exposure factors, this part of code draws a plot of factor exposure bar chart, aiming to help with identifying the where risks from in the chosen portfolio.



​2. Factor Covariance Heatmap

This sophisticated visualization tool begins by dynamically selecting only the most relevant factors based on your portfolio's actual exposures, filtering out insignificant ones to focus your analysis. It then applies a carefully designed blue-white-red color scale with five distinct shades, where blues indicate negative relationships, whites show neutrality, and reds reveal positive connections - all automatically balanced through intelligent symmetric scaling that ensures color intensities accurately reflect relationship strengths. The professional-grade color bar completes the picture with clear semantic labeling from "Strong Negative" to "Strong Positive," precise tick marks for accurate interpretation, and responsive sizing that adapts to any display, working together to transform complex factor relationships into an intuitive, instantly understandable visual map.

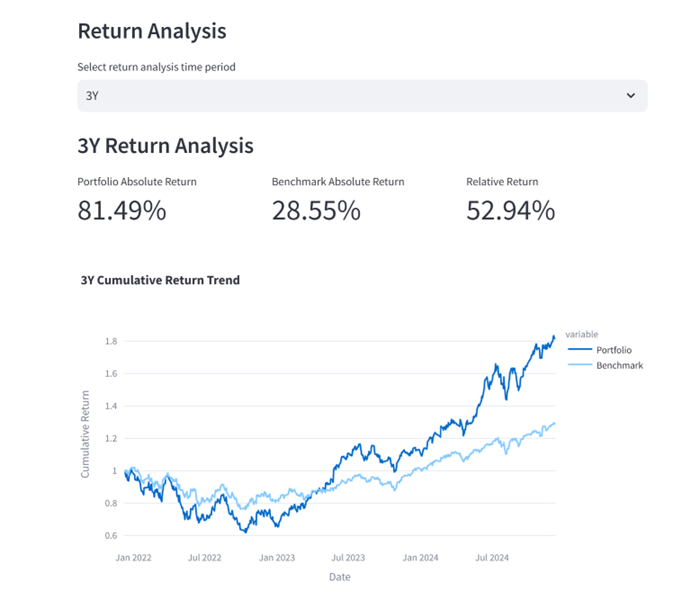


The matrix enables rapid identification of factor interaction patterns, particularly valuable for detecting unintended factor correlations or crowding effects.

**Rolling Volatility Analysis**

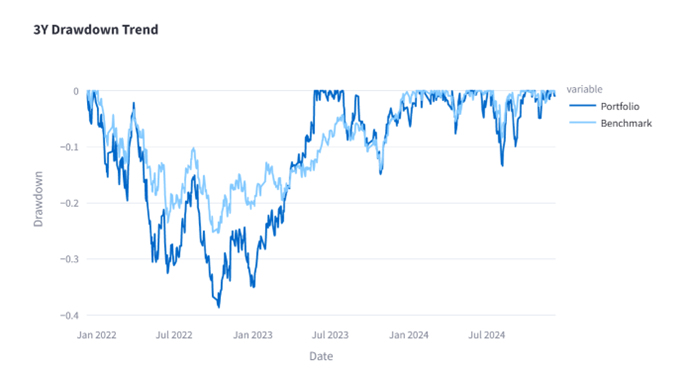
In the complex and ever-changing stock market, investors face many uncertainties. Accurately assessing the risk and return characteristics of the investment portfolio has become an urgent need for investors.

This module allows users to select the window size and analysis period. The volatility trend is displayed in a visual chart by calculating the annualized rolling volatility of the portfolio and the benchmark in the selected time window. This function helps investors intuitively understand the volatility of the investment portfolio in different periods, compare it with the benchmark, and judge the risk volatility characteristics of the investment portfolio.



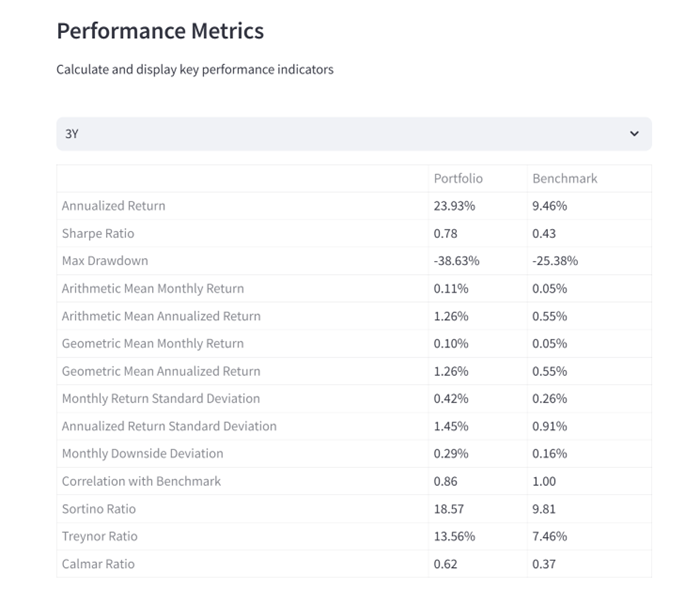
### **Cumulative Return Trend**

The cumulative return trend function will dynamically display, based on the selected time range, the changing trajectory of the cumulative returns of the investment portfolio and the benchmark over time in a visual chart, presenting the process of return accumulation by linking the return data at each time point. This enables investors to observe the return performance at different time points, capture the growth or fluctuation patterns of returns, and then determine the stability of the investment portfolio's return performance during the selected period, as well as the dynamic changes in the return differences compared to the benchmark, providing strong support for investment planning.



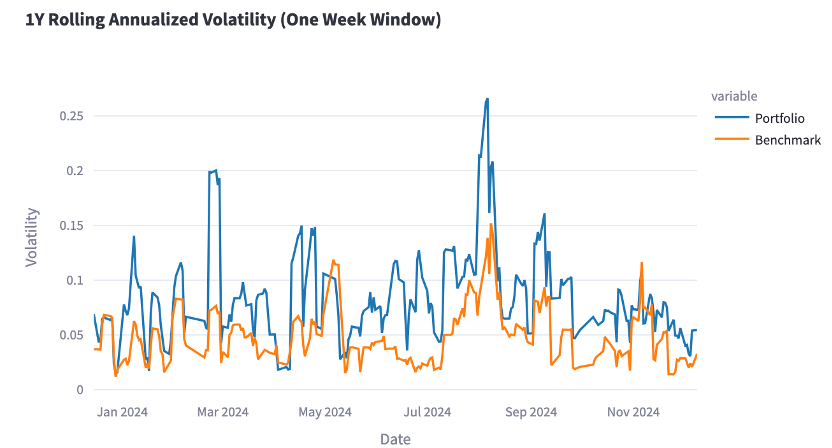
### **Drawdown Trend**

The drawdown trend function, according to the selected specific period, uses a chart to display in real - time the changes in the decline of asset prices from the phased peak over time for the investment portfolio and the benchmark during that period, clearly presenting the dynamic evolution of the degree of value shrinkage during market downturns. Investors can use this to understand the risk exposure level of the investment portfolio during the market decline phase of the corresponding period and know the maximum possible loss of asset value. By comparing the drawdown trends of the investment portfolio and the benchmark, the ability of the investment strategy to control risks during that period can be evaluated, and the risk - resistance performance of the investment portfolio can be measured, providing a reference for risk management and investment decision - making.



### **Performance Metrics**

The performance metrics function systematically calculates and displays a series of key indicators in the investment field, including annualized return, Sharpe ratio, maximum drawdown, etc., based on the selected time period, and quantitatively evaluates the performance of the investment portfolio during that period from multiple dimensions. The annualized return standardizes and converts the returns of the selected period on an annual basis, making it easy to compare the returns of different investment durations and products and measure the long - term value - increasing ability; the Sharpe ratio comprehensively considers returns and risks, measures the additional return per unit of risk, and reflects the return advantage after risk adjustment; the maximum drawdown represents the extreme risk that the investment may face during that period and measures the risk - resistance ability of the investment strategy, etc. These indicators evaluate the investment portfolio from aspects such as return level, risk degree, risk - adjusted return, and return stability. By comprehensively analyzing them, investors can comprehensively and in - depth understand the risk - return characteristics of the investment portfolio, make scientific and rational investment decisions, and optimize the investment portfolio allocation.



Based on this graph, specific volatility data for the portfolio and benchmark are given. These indicators quantify the portfolio's risk level and benchmark, making it easier for investors to assess their risk exposure. In addition, after the user inputs the risk-free rate based on actual conditions, the webpage can directly calculate the Sharpe ratio to help measure risk-adjusted returns.

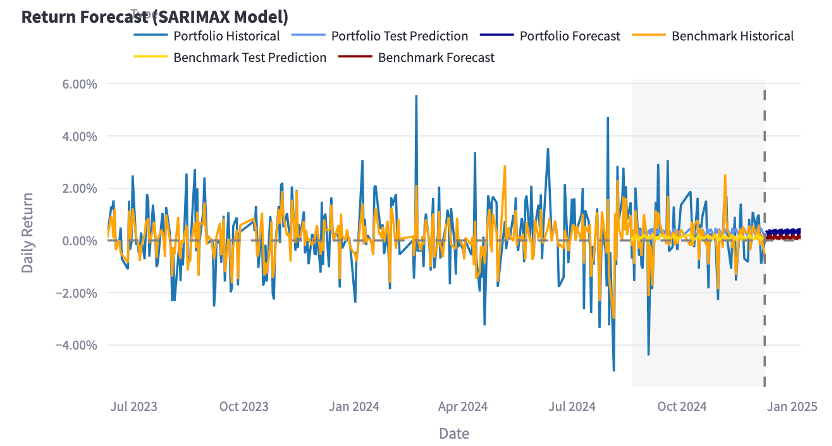
Users can also set the market impact percentage and the number of days the impact lasts to simulate the impact of extreme market conditions on the portfolio. The system calculates and displays the average volatility and Sharpe ratio of the portfolio after the impact, as well as the comparison of portfolio returns before and after the impact, to help investors test the portfolio's resilience under extreme conditions and prepare for risks in advance.

The advantage of this function is that it displays the volatility of the portfolio through rolling analysis and multiple indicators and compares it with the benchmark. It can also flexibly customize the analysis window and period, which is comprehensive, intuitive, and targeted to meet the needs of different investors.

The outlook is to incorporate more factors, such as macroeconomics, to improve the accuracy of analysis, achieve real-time monitoring, and set risk warnings to help investors control risks promptly.

**Return Prediction Analysis**

The webpage provides a variety of forecasting models, including Linear Regression, ARIMA, SARIMAX, Exponential Smoothing, and Prophet. Users can learn the basic information of each model on the webpage and then choose the model and forecast period to get results like the following figure.



The system will give the root mean square error (RMSE) and mean absolute error (MAE). These indicators measure the degree of deviation between the model's predicted value and the actual value, helping investors judge the model's reliability. At the same time, the webpage will also give the predicted average return and predicted volatility of the portfolio and benchmark, providing investors with information on the expected performance of the portfolio and benchmark shortly. Users can also enable the automatic model selection function, and the system will directly select the optimal model and give various related indicators.

The advantage of this part is that it provides a variety of models, flexible forecast periods, and uses professional indicators to evaluate model performance, which is convenient for investors to choose on demand and ensure reliable forecasts.

In the future, hybrid models can be developed, artificial intelligence applications can be deepened, and market sentiment factors can be incorporated to make return forecasts more accurate and in line with the market.

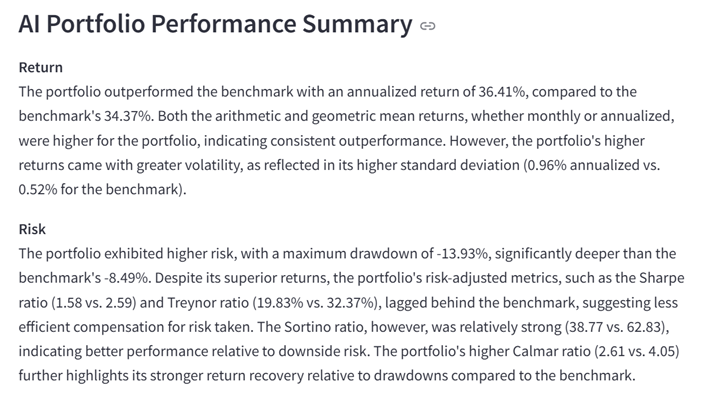
**The integration of AI**

We actively leverage artificial intelligence (AI) to enhance the user experience on our website, focusing on two key areas. First, our AI-driven summarization models deliver concise and insightful portfolio performance information, transforming complex metrics into user-friendly narratives for better decision-making. Second, we have integrated an investment chatbot, enabling users to interact seamlessly and receive real-time, personalized responses to their investment queries. These AI-powered features collectively elevate engagement, accessibility, and the overall value provided to our users.

1. **AI summary**

We engineered a solution that harnesses generative AI large language models (LLMs) to distill complex financial reporting data into concise, engaging summaries tailored for dashboard users. The goal was to convert intricate portfolio performance metrics, including returns and risk indicators, into clear, narrative-driven summaries that are both professional and accessible, seamlessly integrating with a dashboard’s visual interface.

As demonstrated in the provided example, the AI effectively produced succinct, professional summaries structured in two paragraphs under the headings "Return" and "Risk." These outputs captured essential metrics, such as annualized returns and maximum drawdown, while benchmarking them against relevant standards. The engaging yet polished narrative style ensured the summaries were user-friendly and visually compatible with dashboard displays, enhancing decision-making for the audience.



1. **Investment chatbot**

We also include an investment chatbot in our website, The Investment Chatbot on our website is an AI-powered feature designed to provide users with personalized investment guidance. Users can ask a wide range of investment-related questions through a natural language interface, such as inquiries about stock trading, portfolio management, or market trends. The chatbot leverages advanced AI technology to process user queries, analyze relevant data, and deliver accurate, real-time responses.

