# Optimizing an Investment Portfolio: Forecasting & Minimizing Risk

**Team 6:**

Md Shoaib Ullash

Md Shohidul Haque

Sai Siddhartha Dadireddy

Contents

[**1.** **Introduction** 2](#_Toc184695004)

[**2.** **Dataset Overview** 2](#_Toc184695005)

[**2.1 Data Source** 2](#_Toc184695006)

[**2.2 Key Features** 2](#_Toc184695007)

[**2.3 Data Preparation** 3](#_Toc184695008)

[**3. Descriptive Analytics** 3](#_Toc184695009)

[**3.1 Summary Statistics** 3](#_Toc184695010)

[**3.2 Visualizations** 4](#_Toc184695011)

[3.2.1 Key Takeaways from Yearly Stock Value 4](#_Toc184695012)

[**4. Predictive Analytics** 5](#_Toc184695013)

[**4.1 Methodology** 5](#_Toc184695014)

[**4.2 Results** 5](#_Toc184695015)

[4.2.2 Model Performance Using Error Matrics 5](#_Toc184695016)

[**4.3** **Visualization of Forecasting Results** 6](#_Toc184695017)

[**4.4 Analysis of Model Performance** 7](#_Toc184695018)

[**5. Prescriptive Analytics** 7](#_Toc184695019)

[**5.1 Objective** 7](#_Toc184695020)

[**5.2 Methodology** 7](#_Toc184695021)

[5.3.1 Correlation Matrix 8](#_Toc184695022)

[**5.4 Key takeaways from Optimal portfolio weight** 10](#_Toc184695023)

[**5.5 Project Impact** 10](#_Toc184695024)

[**5.5 Future Scope** 10](#_Toc184695025)

[**6. Conclusion** 11](#_Toc184695026)

## **Introduction**

Investing in financial markets involves balancing risk and return to optimize outcomes. Portfolio optimization focuses on selecting assets to maximize returns while minimizing risk. This project uses predictive and prescriptive analytics to build an efficient portfolio featuring five prominent stocks—Amazon, AMD, Cisco, Netflix, and Apple. These stocks were chosen for their growth potential, volatility, and market relevance, offering a balanced mix of stability and opportunity.

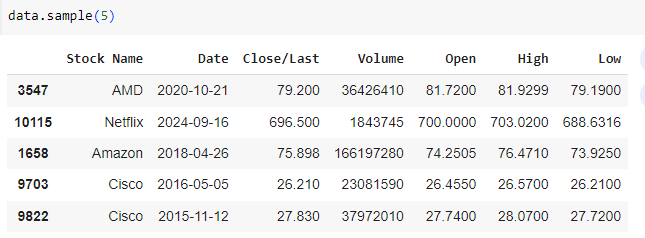
The topic was selected out of a practical interest in managing personal investments in the U.S. stock market. With ownership of multiple stocks, making informed, data-driven decisions became essential. The objectives are to forecast stock prices to identify trends using predictive analytics and optimize investment allocation to achieve a 5% target return with minimal risk. By combining historical stock price analysis, forecasting, and optimization techniques, this project provides a comprehensive strategy for effective portfolio management. It integrates advanced statistical methods to deliver actionable insights for constructing robust investment portfolios, detailed in the methodologies and results.

## **Dataset Overview**

## **2.1 Data Source**

The dataset contains historical stock price data for five major stocks, collected from NASDAQ (<https://www.nasdaq.com/market-activity/quotes/historical>). Each stock has 2,516 rows, adding up to a total of 12,580 rows. The data covers a 10-year period from 2014 to 2024, with daily closing prices and key financial details. This large and detailed data set provides enough information to spot trends, reduce errors, and make accurate predictions. The long-time frame and daily records make the analysis more reliable for forecasting and investment decisions.

## **2.2 Key Features**

The dataset utilized for this project comprises several key attributes that provide a comprehensive view of stock performance. The primary attribute is the Date, which specifies the trading day for each recorded stock price, forming the foundation for time-series analysis. Closing Price, another critical feature, indicates the value of a stock at the close of each trading session. It is a key measure used to analyze historical performance and forecast future trends. Additionally, Daily Returns represent the percentage change in closing prices from one day to the next, offering insights into daily market movements and stock volatility. Complementing this, Weekly Returns are aggregated over a week to align with the project’s mid-term investment analysis focus.

*Fig 1: Sample of the Dataset Showing Key Attributes*

## **2.3 Data Preparation**

To ensure the accuracy and reliability of results, several preprocessing steps were applied to the dataset. First, missing data points, which could have skewed analysis results, were addressed using forward-fill and backward-fill methods. This ensured continuity without introducing biases. Outlier detection was another critical step, where extreme values in stock returns were carefully examined to mitigate their impact on overall analysis. Lastly, feature engineering was employed to calculate weekly returns based on daily price data.

## **3. Descriptive Analytics**

## **3.1 Summary Statistics**

Descriptive statistics provided valuable insights into the historical behavior of each stock in the portfolio. The Mean Returns highlighted the average performance over the analysis period, indicating the overall growth trajectory for each stock. Additionally, the Standard Deviation served as a measure of volatility, with higher values corresponding to increased risk. The maximum and minimum returns highlighted the extremes in performance, showcasing the potential gains and losses investors might face.

A graph of different colored bars

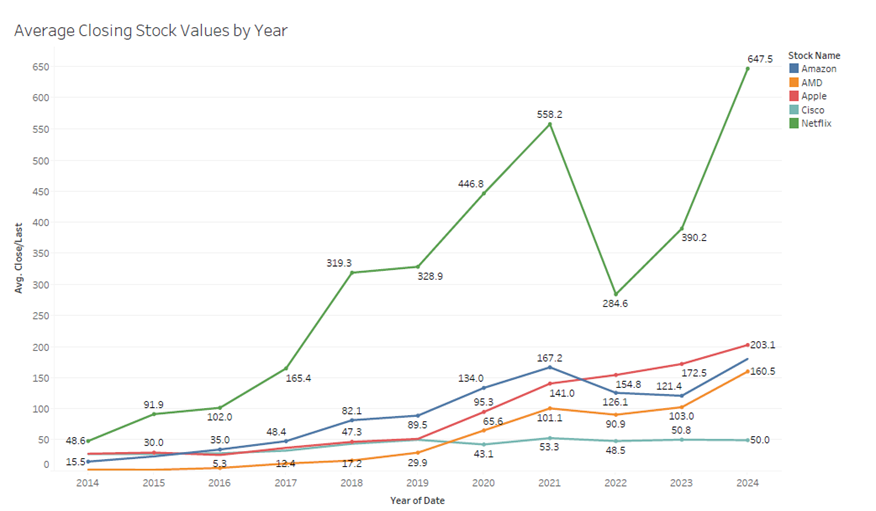
Description automatically generated

*Fig 2: Summary Statistics for Weekly Stock Prices*

The analysis revealed that Amazon and Apple exhibited consistent returns with moderate volatility, making them ideal for risk-averse investors. AMD, on the other hand, displayed the highest variability, reflecting its growth potential but also its associated risks. Cisco maintained stable returns with minimal volatility, appealing to those seeking secure investments, while Netflix demonstrated high returns accompanied by significant volatility, making it suitable for investors with a higher risk appetite.

## **3.2 Visualizations**

To better understand the dataset, the following time series visualizations were generated:

1. **Yearly Stock Value Trend:**

*Fig 3: Yearly Stock Value Trends*

### 3.2.1 Key Takeaways from Yearly Stock Value

The yearly stock value trends provided several actionable insights for portfolio optimization. Stocks like Amazon, Apple, and AMD demonstrated strong upward trends, making them attractive for long-term investment strategies aimed at capitalizing on growth potential. However, the analysis also underscored the importance of managing volatility, particularly for high-risk stocks like AMD and Netflix, which, while offering higher returns, require cautious allocation to mitigate potential losses. Stability emerged as a defining characteristic of Cisco, positioning it as a favorable option for investors prioritizing consistent returns and reduced risk. Furthermore, the weak correlations observed between certain stocks, such as AMD and Netflix, presented opportunities for diversification, allowing for the reduction of overall portfolio risk.

## **4. Predictive Analytics**

### **4.1 Methodology**

The predictive analytics phase employed the **AutoRegressive** (**AutoReg**) model to forecast the prices of Amazon, AMD, Cisco, Netflix, and Apple. This model was selected for its ability to capture time-dependent patterns and lagged observations within stock price data. By analyzing historical data spanning a decade (2014–2024), the AutoReg model predicted future trends for the subsequent 52 weeks (2024–2025). To evaluate the model's accuracy, error metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) were used. These metrics quantified the deviation between predicted and actual values, ensuring the model’s robustness and alignment with historical trends.

## **4.2 Results**

#### 4.2.1 Forecasting Outcomes

The **AutoReg** model generated forecasts that provided actionable insights for portfolio allocation. Stocks like Amazon and Apple displayed consistent upward trends with moderate volatility, reinforcing their suitability for long-term investment strategies. Conversely, AMD exhibited rapid growth potential but with higher price volatility, indicating increased risk for short-term investors. Cisco maintained stable price movements and modest returns, aligning well with risk-averse investment goals. Netflix showcased strong growth potential coupled with significant volatility, requiring a balanced allocation within the portfolio. The model’s ability to accurately predict these outcomes demonstrated its efficacy in capturing the dynamics of diverse stocks, laying a strong foundation for prescriptive analytics.

## 4.2.2 Model Performance Using Error Matrics

The predictive accuracy of the AutoReg model was validated using error metrics. Across all five stocks, the model achieved low MAE, RMSE, and MAPE values, indicating reliable forecasts. For example, the MAPE values ranged between 2–3%, showing that the average prediction error was minimal compared to actual prices.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Stock Name** | **MAE** | **MSE** | **RMSE** | **MAPE** |
| Amazon | 3.93 | 25.38 | 5.04 | 2.2 |
| Apple | 3.3 | 18.12 | 4.26 | 1.65 |
| AMD | 6.08 | 62.27 | 7.89 | 3.8 |
| Cisco | 0.87 | 1.13 | 1.07 | 1.75 |
| Netflix | 14.53 | 335.61 | 18.31 | 2.32 |

## **Visualization of Forecasting Results**

A graph showing the stock market

Description automatically generated with medium confidenceA graph showing a line graph

Description automatically generated with medium confidenceA graph showing the stock market

Description automatically generated with medium confidenceA graph showing the stock market

Description automatically generatedA graph showing a line graph

Description automatically generated with medium confidenceHere, Blue Line indicates historical stock prices, and Orange Dashed Line provides predictions for the last 52 weeks, Red Dashed Line: Forecasted stock prices for the next 52 weeks (2024–2025).

*Fig 4: Stock Trend Prediction for Five Stocks*

These visualizations confirm the model’s ability to align closely with observed data and provide realistic future projections.

## **4.4 Analysis of Model Performance**

The chosen AutoReg model produced low error rates across all stocks, indicating strong forecasting capabilities. The low MAPE values, particularly for Amazon and Apple, reinforce the model’s reliability for predicting future stock prices. The forecasted trends offer actionable insights for portfolio allocation in the subsequent prescriptive analytics phase

## **5. Prescriptive Analytics**

### **5.1 Objective**

The goal of prescriptive analytics is to determine the optimal allocation of investment across five selected stocks—Amazon, AMD, Cisco, Netflix, and Apple—such that the portfolio achieves a specified return of **5%** while minimizing risk. This phase focuses on optimizing portfolio weights using mathematical modeling and leveraging insights from predictive analytics.

### **5.2 Methodology**

#### 5.2.1 Decision Variables

To allocate the total investment across the portfolio, we define the decision variables:

* x1, x2, x3, x4, x5​: Proportion of total investment allocated to Amazon, AMD, Apple, Cisco, and Netflix, respectively.

#### 5.2.2 Optimization Model

The optimization model minimizes the risk of the portfolio while ensuring the return meets or exceeds the specified 5%.

* **Objective Function:** Minimize portfolio risk, measured by the variance of returns:

Where, ​

* *​ : Variance of stock i (square of the standard deviation).*
* : Correlation coefficient between stocks i and j.
* *: Standard deviations of stocks i and j.*
* **Constraints:**
* **Budget Constraint**: x1 + x2 + x3 + x4 +x5 = 1; The sum of all proportions must equal 1 (all capital must be allocated).
* **Return Constraint**: x1μ1+ x2μ2+ x3μ3+ x4μ4+ x5μ5 ≥ A; μi is the expected return of stock i and A is the target return (e.g., 5%).
* **Non-Negativity Constraint**: x1, x2, x3, x4, x5 ≥ 0: No short selling is allowed, so all weights must be non-negative.

#### 5.2.3 Tools and Techniques Used

The optimization problem was solved using Python's scipy.optimize.minimize function, leveraging the SLSQP (Sequential Least Squares Programming) algorithm to determine the optimal weights for minimizing portfolio risk while achieving the target return. The covariance and correlation matrices, derived from historical stock returns using numpy, were used to identify diversification opportunities, while predictive analytics provided expected returns for each stock.

**5.3 Analysis**

While doing analysis in python certain matrix like correlation matrix and covariance heatmap was interpreted for understanding the analysis of this project.

### 5.3.1 Correlation Matrix

The correlation matrix highlights relationships between stocks, guiding diversification:

* Low Correlation: AMD and Netflix (ρ=0.28) present diversification opportunities due to their weak positive relationship. Cisco and Netflix (ρ=0.23) show minimal interaction, making them ideal for risk reduction.
* High Correlation: Amazon and Apple (ρ=0.62) are highly correlated, suggesting they move in the same direction and may increase portfolio risk when combined.

A screenshot of a graph

Description automatically generated*Fig 5: Correlation Matrix*

#### A diagram of a number of stocks Description automatically generated with medium confidence5.3.2 Covariance Matrix

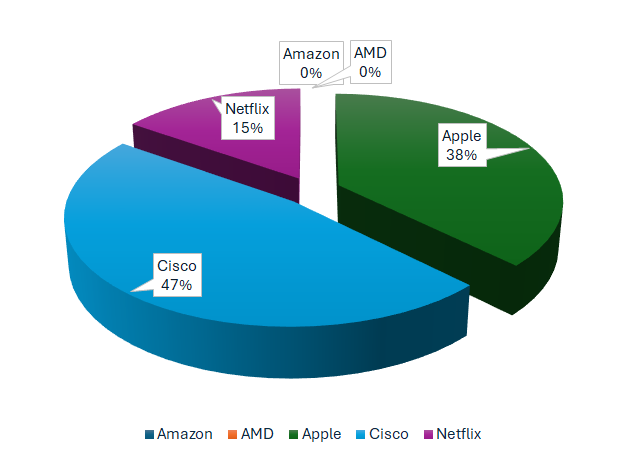
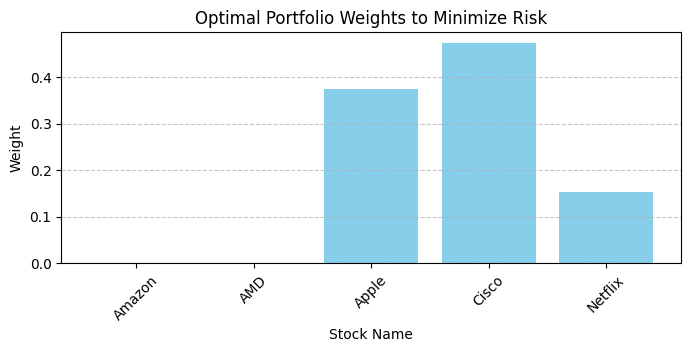
*Fig 6: Covariance heatmap*

The covariance matrix provides additional insights into risk interactions:

* AMD showed the highest self-variance (σ2=0.8782), indicating high volatility.
* Apple and Cisco had the lowest covariance (σ=0.0824), suggesting independence in their price movements.

#### 5.3.3 Optimal Weights

Optimal portfolio weights were calculated to achieve a target return of 5% while minimizing risk. The solution strategically emphasized investments in less volatile stocks, such as Cisco and Apple, while balancing allocations to higher-growth stocks like Netflix. The analysis leveraged the low correlations and covariances identified earlier to distribute investments effectively, reducing the portfolio’s overall risk. By balancing growth-oriented and stable stocks, the optimized portfolio aligned with the project’s objectives of achieving a high return-to-risk ratio.



*Fig 7: Optimal weights for Portfolio*

### **5.4 Key takeaways from Optimal portfolio weight**

The optimal portfolio allocation revealed several critical insights. Diversification emerged as a key strategy, as allocating investments to stocks with low correlation minimized the impact of market fluctuations on the portfolio. The analysis also highlighted the tradeoff between risk and return, with higher-volatility stocks like AMD contributing significantly to potential gains but also to overall risk. Finally, the optimized portfolio demonstrated efficiency by achieving the target return of 5% while minimizing variance, ensuring a balanced and robust investment strategy.

### **5.5 Project Impact**

This project applied advanced analytics techniques to enable more insightful investment decisions. By leveraging portfolio optimization, the study successfully demonstrated how to minimize risk while achieving targeted returns. The findings emphasized the importance of data-driven approaches in constructing efficient portfolios that balance risk and reward effectively. These results contribute to improving investment strategies and highlight the potential of quantitative methods in financial decision-making.

### **5.5 Future Scope**

The project envisions integrating real-time market data to enhance the adaptability of portfolio optimization to dynamic market conditions. Additionally, it aims to explore the use of machine learning for automatic portfolio adjustments, enabling faster responses to market trends. These advancements will ensure better risk management and improved investment performance in evolving financial landscapes.

## **6. Conclusion**

This project successfully showcases a comprehensive approach to investment portfolio optimization through the integration of descriptive, predictive and prescriptive analytics. By forecasting stock prices for Amazon, AMD, Cisco, Netflix, and Apple, the project provides data-driven insights into future price trends. The optimization process strategically allocates investments to achieve a target return of 5% while minimizing risk. Advanced techniques like AutoReg modeling, error metric analysis, and optimization algorithms ensure precise forecasting and effective portfolio allocation. The results underscore the importance of diversification, highlighting how a balanced mix of high-growth and stable stocks reduces overall risk. This approach offers a practical framework for enhancing investment decisions, with potential future developments including real-time market integration and automated portfolio adjustments through machine learning.