2024 Undergraduate Reseach Conference

Time Series Analysis of S&P Index Prices and Volume (2021-2024) with Python

Salvador Vazquez
Advisor: Dr. Yonghong Tong

NIAGARA UNIVERSITY

Abstract

This study utilizes historical data of S&P (Standard and Poor's) prices and trading volume starting from January 1, 2021, to January 1, 2024, to conduct a time series analysis. The objective is to analyze the dataset, uncover trends, patterns, and anomalies, and potentially develop forecasting models. By employing statistical methods and visualization techniques, this analysis aims to provide insights into the behavior of the S&P market over the specified time period, assisting in informed decision-making for investors and financial analysts.

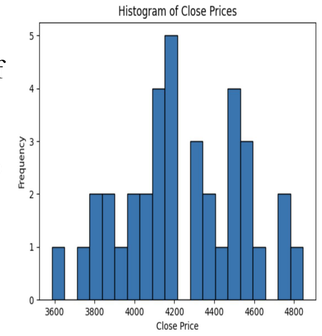
Introduction

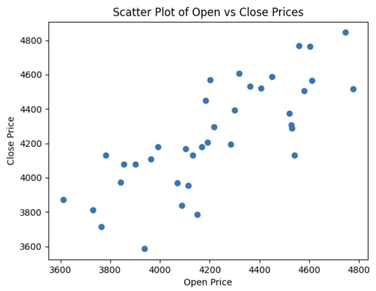
The S&P, also known as the Standard and Poor's 500, is a stock market index that measures the stock performance of 500 large companies listed on the stock exchanges in the United States. In this project, we delve into a comprehensive analysis of the S&P stock market dynamics from 2021 to 2024, utilizing Python as our primary analytical tool. My investigation centers on exploring the trends and patterns exhibited by S&P prices and trading volumes over this period. Utilizing Python's data analysis libraries, including pandas for data manipulation and matplotlib for visualization, I aim to extract actionable insights for investors and financial analysts.Through statistical analysis, I offer stakeholders valuable information to inform their investment strategies and decisions.

Result

I observed fluctuations in stock prices and trading volumes throughout the period under review. Statistical analysis revealed patterns and trends in the data, indicating a correlation between price movements and trading activity. Visualizations, including line plots and scatter plots, helped with illustrating these trends graphically. Overall, the analysis provided valuable insights into the behavior of the S&P stock market during the specified timeframe.

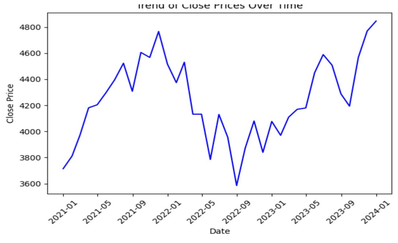
The histogram provides a visual representation of the distribution of closing prices in the S&P stock market dataset. Each bar on the histogram represents a range of closing prices, and the height of each bar indicates the frequency of occurrence of closing prices within that range. By examining the histogram, we can observe the pattern of closing prices and identify any clusters or outliers in the data. This information is valuable for understanding the typical range of closing prices and the frequency of occurrence of different price levels in the dataset. Additionally, the histogram allows us to assess the symmetry and skewness of the distribution, providing insights into the overall shape of the data distribution.





The scatterplot illustrates the relationship between opening and closing prices in the S&P stock market dataset. Each point on the scatterplot represents a specific date, with its coordinates corresponding to the opening price (x-axis) and the closing price (y-axis) for that date. By examining the scatterplot, we can visually assess whether there is a linear relationship between opening and closing prices. Additionally, we can observe the clustering of data points, which may indicate certain patterns or trends in the data. In this context, adding a trendline to the scatterplot allows us to quantify the strength and direction of the relationship between the two variables.

The trendline represents the best-fit line through the data points, providing insight into the overall trend in closing prices relative to opening prices. Analyzing the slope and direction of the trendline can help identify whether there is a positive or negative correlation between opening and closing prices, and to what extent one variable influences the other. It offers a visual and quantitative understanding of the relationship between opening and closing prices in the S&P stock market dataset.



Conclusion

The examination of closing prices through the histogram revealed a distribution pattern that can help investors understand the typical range and frequency of closing prices, aiding in market analysis and the identification of the most common closing prices. Additionally, the scatterplot of open vs. close prices, along with the trendline analysis, provided further understanding of the relationship between these variables, offering insights into trading patterns and market behavior. Furthermore, the trend plot of close prices over time allowed for the visualization of long-term market trends and fluctuations, enabling investors to identify periods of volatility and stability. The analysis has provided actionable insights that can inform investment decisions and strategies.

Future Work

While the analysis has provided valuable insights into the S&P pricing behaviors, some of the areas to explore in order to have a better predictive analysis are the following:

- Develop predictive models using advanced statistical techniques and machine learning algorithms to forecast future movements in S&P stock prices based on historical data.
- Explore the impact of external factors such as economic indicators, geopolitical events, and regulatory changes on the S&P market performance.
- Explore machine learning applications for anomaly detection, pattern recognition, and trend forecasting in the S&P stock market dataset to uncover hidden insights and opportunities.

Acknowledgement:

This research has been supported by Niagara University.

References

Kenton, Will. "S&P 500 Index: What It's for and Why It's Important in Investing." Investopedia, September 22, 2023. https://www.investopedia.com/terms/s/sp500.asp.

Sonkavde, Gaurang, Deepak Sudhakar Dharrao, Anupkumar M. Bongale, Sarika T. Deokate, Deepak Doreswamy, and Subraya Krishna Bhat. 2023. "Forecasting Stock Market Prices Using Machine Learning and Deep Learning Models: A Systematic Review, Performance Analysis and Discussion of Implications" *International Journal of Financial Studies* 11, no. 3: 94. https://doi.org/10.3390/ijfs11030094