Exploratory Data Analysis Approach to Identify and Analyze a Systemic Financial Risk

Authors: Win Htet; Hnin Oo Wai
Instructor: Professor Joe Byers
WorldQuant University, MScFE Program
winhtet2911@gmail.com; *hninoowai2911@gmail.com*
joe.byers@wqu.edu

Abstract

We attempt to identify and analyze systemic vulnerabilities especially in the global financial markets like the U.S and China using exploratory data analysis approach with the point of view of data analysts from third world countries in mind. This paper is aimed to explore and develop the application of exploratory data analysis (EDA) approach to identify and manage the systemic risk and mitigating the losses that could happen. Our focus shall be more into implementing quantitative approach rather than qualitative approach. We will explore the available information about past literature reviews and try to explore the US market with data visualization.

Keywords: *Exploratory data analysis, data visualization, systemic risk indicators, SWOT analysis, Root Cause analysis, VaR Modeling, ARIMA*

1. Introduction

The financial industry is a very complex network. Systemic risk is known as the possibility that an event at the company or similar level could trigger severe failure to the entire system rather than just some of its constituent parts. It alludes to the potential for a catastrophic financial collapse caused by a chain reaction breakdown of interconnections in the financial industry. This risk had a significant role in the 2008 financial crisis. The recent COVID-19 epidemic also has a major effect on China's banking and real estate sectors' systemic risk. (Huang et al., 2022) Thus, the stability of the financial sector depends on the identification and management of systemic risk.

Data scientists utilize exploratory data analysis (EDA) as a powerful tool to examine, analyze, and summarize the key features of data sets to identify and understand systemic risk. EDA frequently makes use of data visualization techniques. [(Hullman & Gelman, 2021)] stated that a good inference model can help us see the connections between multiple seemingly opposing actions.

The aim of this capstone project is to detect systemic vulnerabilities in the global market like the U.S using EDA techniques. The project will evaluate existing literature on systemic risk and EDA, explain methodology, outline the data collection procedure, and provide the EDA findings. The project will finish with a discussion of the findings' implications and potential future research directions.

2. Literature Review

Several researchers have contributed to the subject of EDA with each providing a distinct perspective and extending our understanding of this complex process.

2.1. Past Literature Reviews and respective Contributions

In this literature review, we will look into the past studies' key contributions made in the field of EDA, with the following summary table.

Table 1. Past Literature Reviews Summary

Sr. No.	Author(s) (Year)	Problem Type	Model Applied	Dataset Used
1	Ma et al., (2022) [1]	Identifying systemic financial risks before and after the COVID-19 outbreak	Copula-GARCH with CES approach	Financial data from the US market
2	Sun et al., (2009) [2]	Identifying systemic events at an early stages and containing steps	Financial soundness indicators, advanced tools like CCA, option iPoD, Multivariate Dependency, etc	Financial data from the US market
3	Craig, (2020) [3]	Developing a systemic risk indicator	Not specified	US financial market data
4	Hullman & Gelman, (2021) [4]	Exploring systemic risks indicator	Exploratory Data Analysis (EDA)	Various financial datasets
5	Adamjee, (2023) [5]	Analyzing stock market data for systemic risks	EDA techniques, statistical analysis	Historical stock market data

We shall look into the contribution of each research mentioned in Table 1. The first paper by [(Ma et al., 2022)] makes an important contribution by estimating systemic risk during the pandemic, identifying vulnerable institutions, and providing macroprudential guidelines. It bridges the gap between exploratory and confirmatory analysis, emphasizing the importance of risk assessment in real time. And the second paper researched by [(Sun et al., 2009)] contributes to the knowledge of systemic risk identification by combining statistical modeling, market data analysis, and Bayesian inference. It emphasizes the importance of taking a comprehensive strategy to improve financial stability. According to [(Craig, 2020)], the Systemic Risk Indicator is a useful tool for tracking systemic risk. It takes a model-driven approach. The fourth paper by [(Hullman & Gelman, 2021)] provides a foundational resource for data scientists, emphasizing the importance of proper EDA. It connects the theoretical and practical divides, allowing researchers to make more informed conclusions. Lastly, the fifth paper by [(Adamjee, 2023)] contains useful information for traders, investors, and analysts. It improves stock market investment decision-making by combining EDA with domain-specific expertise.

2.2. Comparative Analysis on these literature reviews

We shall conduct a comparative study of these previous investigations. The COVID-19 pandemic had a significant influence on global financial markets, prompting a thorough review of systemic financial risks in order to maintain long-term stability. In the first paper, [(Ma et al., 2022)] predicts systemic risk in the United States both before and during the pandemic using advanced statistical models, specifically copula-GJR-GARCH paired with component expected shortfall (CES).The use of complex models is a strength, but assumptions of stationarity and normalcy are possibly incongruent with extreme market conditions. Future study possibilities include expanding the analysis to global markets for cross-country comparisons and investigating other risk indicators. Threats, on the other hand, include potential data restrictions affecting model performance and the dynamic nature of market conditions that may call model validity into question, as well as the potential impact of regulatory changes or market shocks altering SIFIs' risk profiles.

Early detection of systemic occurrences is critical for crisis containment, especially given the worldwide interconnection of financial systems. In the second paper, [(Sun et al., 2009)] focuses on detecting early warning signs of approaching systemic crises by using financial soundness indicators (FSIs) to assess individual vulnerabilities. It emphasizes the importance of taking into account off-balance-sheet exposures and wholesale funding. The article recommends using systemic stress tests and expanding FSIs to include liquidity shortfalls. While integrated exploratory and confirmatory studies via Bayesian model checks are strengths, drawbacks include retroactive analysis, the assumption of linear correlations, and potential oversight of non-financial elements influencing systemic risk. Opportunities exist in developing real-time monitoring systems, investigating machine learning for early warnings, and working with regulators to improve systemic risk surveillance, yet threats include legislative changes affecting FSIs and the possibility for model assumptions to diverge during extreme occurrences, which can be influenced by behavioral biases in analysts' interpretations.

In the third study, the problem states that systemic risk threatens financial stability, prompting the development of an indicator aimed at reflecting market perceptions of widespread insolvency risk in the US banking system. (Craig, 2020) This indicator's strength rests in its use of forward-looking market data to identify systemic threats, comparing average insolvency risk (ADD) and portfolio insolvency risk (PDD) for a holistic picture. It provides timely warnings of increased financial system stress and promotes model-driven inference with an emphasis on uncertainty representation. However, drawbacks include a focus on the financial services industry in the United States, reliance on market perceptions, and potential oversight of non-financial systemic concerns. Opportunities for improvement include broadening the methodology to global contexts, investigating other risk measurements, and investigating the impact of policy initiatives. Threats include potential data restrictions that impair indicator accuracy, quick changes in market dynamics that affect risk assessments, and the possibility of missing basic vulnerabilities due to an overreliance on market data.

In the fourth and the fifth studies, the paper concept is more about EDA. [(Hullman & Gelman, 2021)] highlights Exploratory Data Analysis (EDA) as a critical step in understanding data complexities, emphasizing visual exploration, summary statistics, and iterative strategies for refining research topics. It does, however, address flaws, such as assumptions regarding data quality and a lack of domain-specific insights. Extending EDA to high-dimensional data and investigating interactive tools are two areas for advancement. It serves as a core resource, bridging theory and practice and helping data scientists to make educated decisions. Meanwhile, [(Adamjee, 2023)] focuses on applying EDA to stock market datasets, emphasizing the advantages of using financial time series data for trend discovery. However, constraints include a narrow focus, market efficiency assumptions, and the potential oversight of long-term investment strategy. Extending EDA to additional financial instruments and investigating behavioral biases are two possibilities. This study provides practical insights for traders by combining EDA with domain-specific information, hence improving stock market investing decision-making.

2.3. Capstone Project Focus

Our capstone project will center on three key areas identified in the literature review:

1. **Systemic Risk Identification**: Our goal is to identify the risk using Exploratory Data Analysis (EDA) methodologies. The objective is to uncover early warning signs of systemic problems in the US market.
2. **Financial Stability Assessment**: We will examine financial soundness indicators, stress tests, and market data using EDA. The purpose of this evaluation is to provide insight into the stability of financial institutions in the United States.
3. **Analysis of Market Volatility**: Using EDA on stock market data, we aim to analyze volatility patterns, correlations, and prospective investment methods.

We find that these focused efforts are consistent with the literature's emphasis on EDA's usefulness in predictive modeling, financial stability assessment, and market analysis.

3. Methodology

We will use the following approaches to attain our capstone project objectives:

1. **EDA Techniques**: As part of Exploratory Data Analysis (EDA), we will use data visualization, summary statistics, and pattern recognition to obtain insights into our datasets.
2. **Identification of Systemic Risks**: To identify and assess systemic risks, we use financial soundness indicators, forward-looking market data, and systemic stress tests.
3. **Statistical Risk Management**: To assess prospective financial losses, we will use Value-at-Risk (VaR) modeling as a statistical risk management tool.
4. **Tools for Identifying Risks**: Apart from SWOT analysis, to completely identify and analyze risks, additional root cause analysis methods such as Failure Mode and Effect Analysis (FMEA), Cause and Effect Diagram (Ishikawa or fishbone), and Pareto charts can be used.

3.1. EDA Techniques

Data visualization is a pivotal component of Exploratory Data Analysis (EDA), aiding analysts in discerning patterns and relationships within data through the creation of graphs and charts. This visual representation enhances the understanding of complex datasets. Another way to perform EDA is by using summary statistics, encompassing measures like mean, median, mode, and standard deviation. [(Soni, 2023)] stated in a LinkedIn article that those statistics offer a swift overview of the dataset, facilitating the identification of outliers and anomalies. By providing a snapshot of key characteristics, summary statistics contribute to a comprehensive understanding of the data's central tendencies and variability.

We shall use various economic indicators such as the GDP growth rate, unemployment rate, inflation rate, and interest rate to identify systemic risk in the US market. These datasets are the average annual data points taken from 1980 to 2024 (current).

Another important method used in EDA is correlation analysis. It helps identify correlations between variables in a dataset. This study is quite useful for finding hidden patterns and identifying variables that have a substantial impact on the overall analysis. We can explore these relationships between variables into three basic data exploration methods as mentioned in the table below.

Table 2. Three Basic Data Exploration methods

Sr. No.	Analysis	Variable Relationship	Purpose of method	Techniques
1	Univariate	Single variable	To comprehend the distribution of the variable, detect outliers, and discern any inherent patterns or trends	Histogram, box plot, descriptive statistics (like mean, median, standard deviation)
2	Bivariate	Two variables	To determine how these variables are related and whether a correlation exists	Scatter plots, correlation analysis
3	Multivariate	Three or more variables	To expose deeper understanding, patterns and trends within the dataset	Principal Component Analysis (PCA), cluster analysis

3.2. Systemic Risk Identification Methods

Identifying systemic risks requires identifying potential hazards that may cause financial system instability. (Hartwig et al., 2021)

Among the several ways are:

1. **Financial Soundness Indicators (FSIs)**: These measurements, which include capital adequacy, asset quality, earnings, liquidity, and market risk sensitivity, serve as indicators of the health and stability of the financial system.
2. **Forward-Looking Market Data Analysis**: This is scrutinizing market data in order to predict future systemic threats. Time-series analysis, machine learning algorithms, and econometric models are examples of techniques.
3. **Systemic Stress Tests**: These tests use simulations to examine a financial system's resistance to various stress scenarios. They contribute to the identification of potential systemic threats by exposing vulnerabilities. (Courtnell, 2020)

ARIMA is a time-series analysis used to forecast the future based on historical values. (CFI Team, n.d.) ARIMA stands for Auto-Regressive Integrated Moving Average and the common notation with seasonal component is $ARIMA(p, d, q)$, where

- p is the number of autoregression also called lag order, showing the dependent relationship between existing data and its previous values.
- q is the number of errors in forecasting, also stands as the window size of moving average.
- d is the degree of differencing which enables stationarity of data.

We shall use a general approach using grid search to determine the best order parameters for our ARIMA model. We shall separate the dataset into train and test sets 80-20 and build an ARIMA model to analyze whether this model is able to predict the systemic risk or not.

3.3. Value-at-Risk (VaR) Modeling

VaR is a statistical method used to measure and quantify the amount of financial risk present in an investment portfolio or firm over a certain period of time. The $p(VaR)$ can be defined informally for a given portfolio, time horizon, and probability p as the highest probable loss during that time after we reject all worse events whose probability is less than p. This assumes mark-to-market pricing and no portfolio trading. (Jain, n.d.)

For example, if the 98% one-day VaR is $2 million, there is a 2% chance that the portfolio will lose more than $2 million in a single day if no trading occurs. Informally, a loss of $2 million or more is projected on 1 day out of every 50 (since 2% equals 1/50).

Mathematically, as stated by [(Li, 2016)], given a probability level $p \in (0, 1)$ for the time T and $(T + \alpha)$, for a given p, we may define the VaR of a long position over the period α as

$$p = \mathbb{P}[\Delta V(\alpha) \leq VaR] = F_{\alpha}(VaR) \tag{1}$$

where, $\Delta V(\alpha)$ be the change in financial asset over the time period α, $F_{\alpha}(x)$ be the CDF of $\Delta V(\alpha)$ being ≤ 0.

As for a short position, with the financial position $\Delta V(\alpha) \geq 0$, the VaR becomes

$$p = \mathbb{P}[\Delta V(\alpha) \geq VaR] = 1 - \mathbb{P}[\Delta V(\alpha) \leq VaR] = 1 - F_{\alpha}(VaR) \tag{2}$$

Following this equation, we can define the p-quantile of $F_{\alpha}(x)$ for any CDF with given confidence level of $p \in (0, 1)$ as

$$VaR_p = x_p = inf\{x \mid F_{\alpha}(x) \geq p\} \tag{3}$$

where, inf be the smallest real number, x_p is known.

Another statistical risk management technique is to use standard deviation to calculate the historical volatility of an investment in relation to its yearly rate of return.

3.4. Risk Identification Tools

Risk identification is quite important in identifying and assessing issues that could jeopardize the success of a business or project. It allows the decision maker to assess the risks and determine a go or no-go decision. For instance, by using the SWOT analysis, decision makers can generally identify the Strengths, Weaknesses, Opportunities, and Threats related to a project or dataset. According to (Borsalli, 2022), some other risk identification tools are FMEA (Failure Mode and Effect Analysis), Fishbone Diagram, and Pareto Charts.

FMEA is a technique for identifying any and all possible defects in a product or service, manufacturing or assembly process, or design. It's usually organized as a table, with ratings for severity, occurrence, and detection for each failure mode. These ratings, which are typically on a scale of 1 to 10, are multiplied together to yield a Risk Priority Number (RPN) for each failure mode: (Borsalli, 2022)

$$RPN = Severity \times Occurrence \times Detection \tag{4}$$

The failure modes are then prioritized based on their RPN, and efforts are conducted to limit the occurrence of the failures with the highest priority.

The Fishbone diagram is more of a visual tool than a mathematical one, used to find out potential root causes of a problem, while these root causes are categorized and plotted as branches of a “fishbone” alike diagram. Following that, the Pareto principle, commonly known as 80/20 Rule, states that 80% of consequences originate from only 20% of the causes. Mathematically,

$$Cumulative\ \%\ of\ items = \frac{Cumulative\ Frequency\ of\ items}{Total\ frequency\ of\ items} * 100 \tag{5}$$

4. Data Preparation

We will take more than forty years worth of S&P 500 index data points from Yahoo Finance website. After importing the dataset into our Google Colab - Python Notebook, we initiate the checking of null data, missing data, and outliers, etc. Besides, in order to compare different variables, we use a standard scaler method to make our dataset standardized. After making sure, there are no abnormal data points in the dataset, the basic EDA methods are applied to know more about the dataset.

Firstly, we have used the single variable which is the adjusted closing daily prices of S&P 500 to comprehend the trends and distribution of the dataset. We shall use additional economic indicators like the GDP growth rate, interest rate, unemployment rate and inflation rate to perform bivariate and multivariate analysis.

Table 3. Datasets and Descriptions

Sr. No.	Index	Description	Period	Frequency	Source
1	S&P500	Adjusted Closing Prices of S&P 500	From 1980-01-01 to 2024-01-01	Daily	Yahoo Finance
2	GDP_ Growth_ Rate	Real percent change of GDP growth rate	From 1980 to 2024 (current)	Annual	IMF Datamapper [link]
3	Inflation_ Rate	Inflation rate calculated on average of consumer prices	From 1980 to 2024 (current)	Annual	IMF Datamapper [link]
4	Unemployme nt_Rate	Unemploym ent percent of total labor force	From 1980 to 2024 (current)	Annual	IMF Datamapper [link]
5	Interest_rate	Average of Federal Fund Rate	From 1980 to 2024 (current)	Annual	FRED [link]

Table 3 shows the information about the datasets used.

5. Data Visualization

First of all, we try to visualize the S&P 500 dataset as a univariate analysis. We choose to take 'Adjusted Close Price' as our single variable.

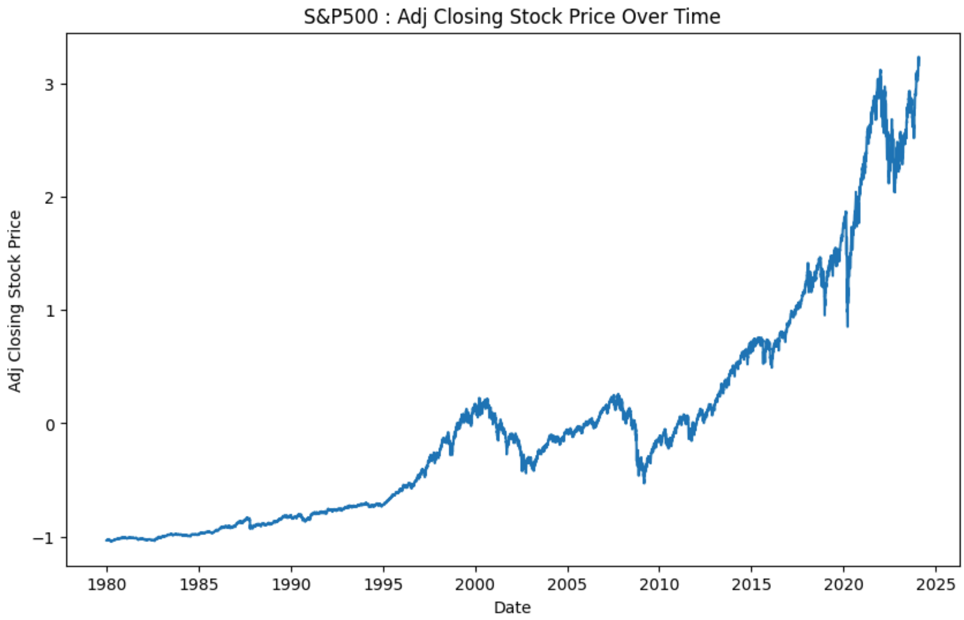

Figure 1. S&P 500 Graph on Adj Closing Price (Daily)

We can find that the Adj Closing price of S&P stock is increasing over time with some fluctuation happening around 2020 (due to pandemic). However, when we compare this trend to the 2008 financial crisis, we can find that the downward trend patterns are different. One noticeable fact is the recovery period. It took more than a year and half for the recession to finalize during the 2008 crisis, and during the 2019 pandemic period, it seems to take around one year for the stock price to bounce back to its original position before the pandemic.

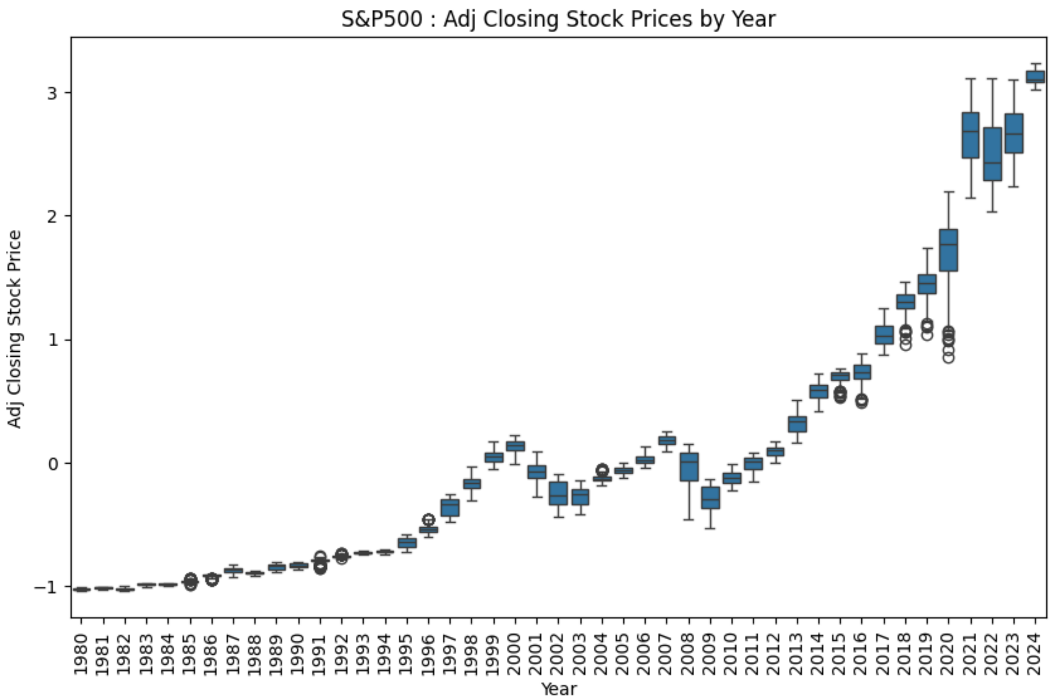

Figure 2. Candle Graph on Adj Closing Price (Daily)

With this plot, we can find that around 2018 to 2020, the price increased over time with some outliers.

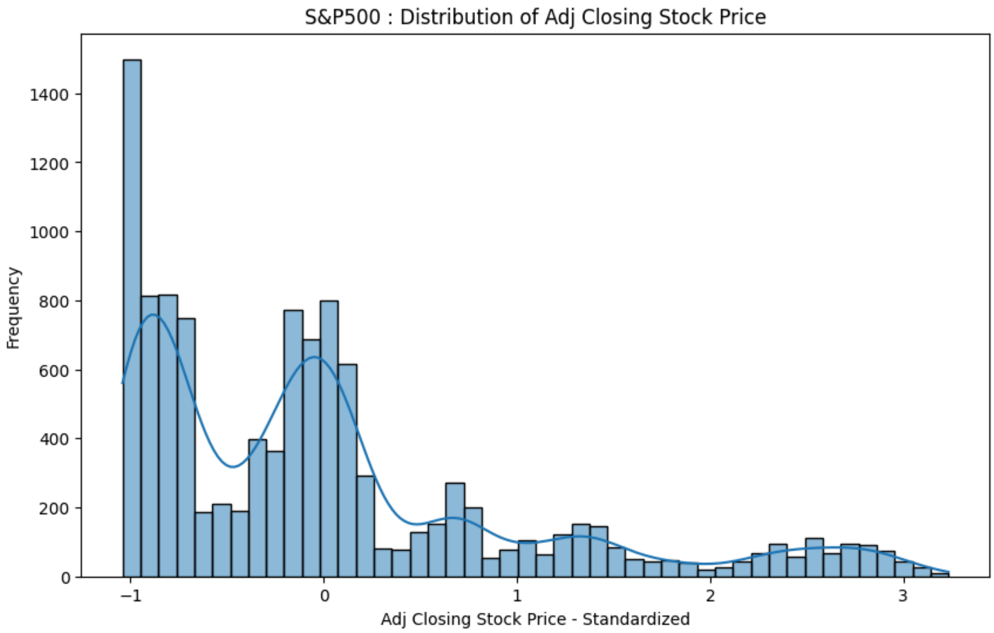

Figure 3. Distribution of Adj Closing Stock Price - Standardized

The x-axis represents the "Standardized Closing Stock Price" ranging from -1 to 3, and the y-axis represents the "Frequency" ranging from 0 to 1400. The blue bars of varying heights represent the frequency of different adjusted closing stock prices. There's also a line graph overlaid on the bars showing the trend in the distribution, which peaks at around an adjusted closing price of 0.

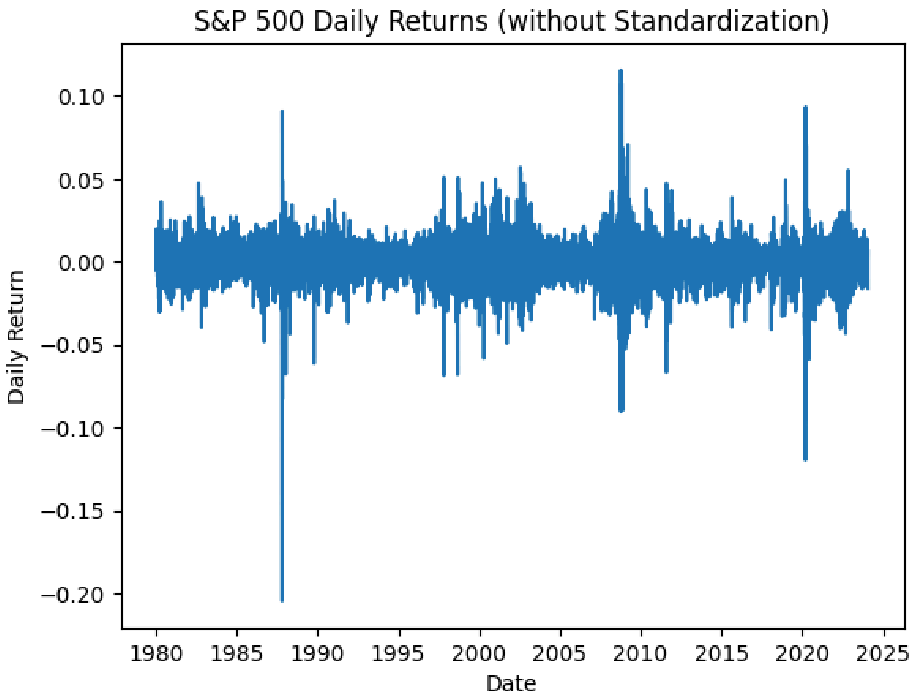

Figure 4. S&P 500 Daily Returns

We can find the huge spikes in the graph during 1987, 2008 and 2020. This could suggest some financial events happening and could be the chance for systemic risk.

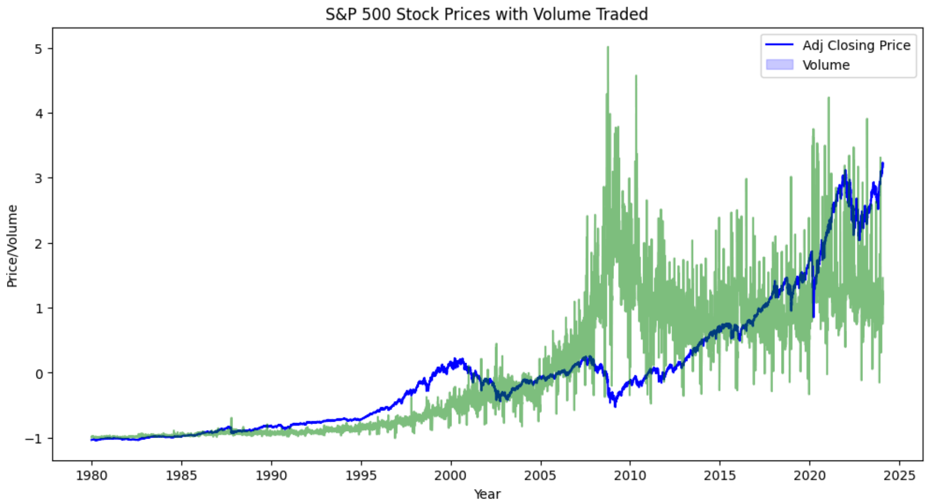

Figure 5. S&P 500 Stock Prices with Volume Traded

The graph shows the combined charts of S&P 500 stock closing prices with volume traded. We can see the similar movement/trend in this graph. We got 70.85% of positive correlation value. Apart from that, we shall need to perform multivariate data exploration to understand more about the circumstances.

6. Finding and Discussion

We can find that the Adj Closing price of S&P stock is increasing over time with some fluctuation happening around 2020 (due to pandemic). We used the external economic indicators mentioned in Table 3 to implement the principal component analysis to analyze the potential risk in the financial market.

6.1. Findings after PCA analysis visualization

PCA reduces the dimensionality of the data by creating new features (principal components) that capture the most variance in the data. We shall try to visualize our results using the scatter plot and the bar plot. We input the interest rate, the inflation rate, the GDP growth rate, S&P 500 prices and the unemployment rate as our initial features.

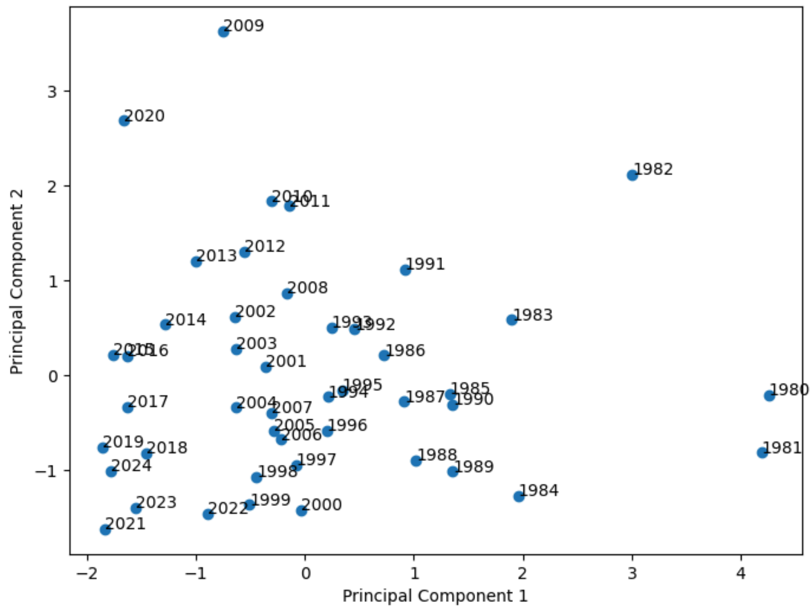

Figure 6. Scatter Plot on the principal components

The scatter plot visualizes the principal components, which are the directions in which the original data varies the most. Each point on the scatter plot represents a year, and the coordinates of the point are the values of the first and second principal components for that year. We can see that the dot com crash period (2000 to 2002) and 2008 financial crisis possess similar systemic risk profiles since they are nearer to each other. Yet, regarding the 2019 pandemic, we can find different positioning on the plot. This can indicate that the 2019 pandemic did not impose as much systemic risk threat as 2008.

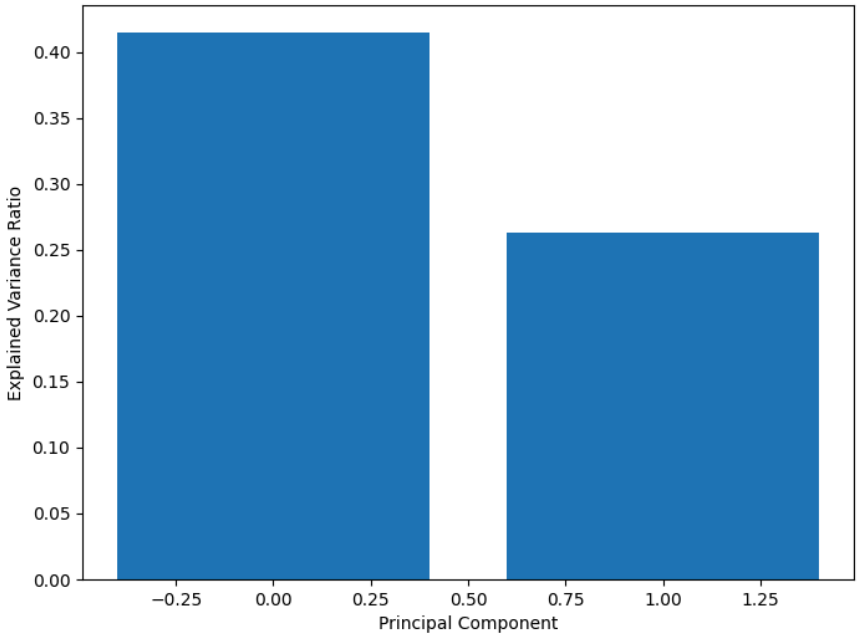

Figure 7. Bar Plot on the principal components

The explained variance ratio [0.41505145, 0.26294005] shows that 41.5% of the variance in the data is explained by the first principal component, and another 26.3% is explained by the second principal component. This means that these two components together explain about 67.8% of the total variance. The remaining variance is explained by other factors not captured in these two components.

More details about code and calculations can be found in the source code file. When we try to explore the PCA components with our initial features, we can interpret the result as follows:

- PC1: The first principal component is strongly influenced by the S&P 500 (-0.515), Interest Rate (0.613), and Inflation Rate (0.502) as they have high values. This suggests that these three variables move together and might represent some form of economic condition. For example, when the S&P 500 goes down, the interest rate and inflation rate might go up, indicating a recovery in the economy.
- PC2: The second principal component is strongly influenced by the GDP Growth Rate (-0.62) and Unemployment Rate (0.667). This suggests that these two variables move together and might represent another form of economic condition. For example, when the GDP growth rate goes down, the unemployment rate might go up, indicating an economic recession.

6.2. Findings after ARIMA model

We will now try to determine whether the ARIMA model can be used to predict or identify the risk.

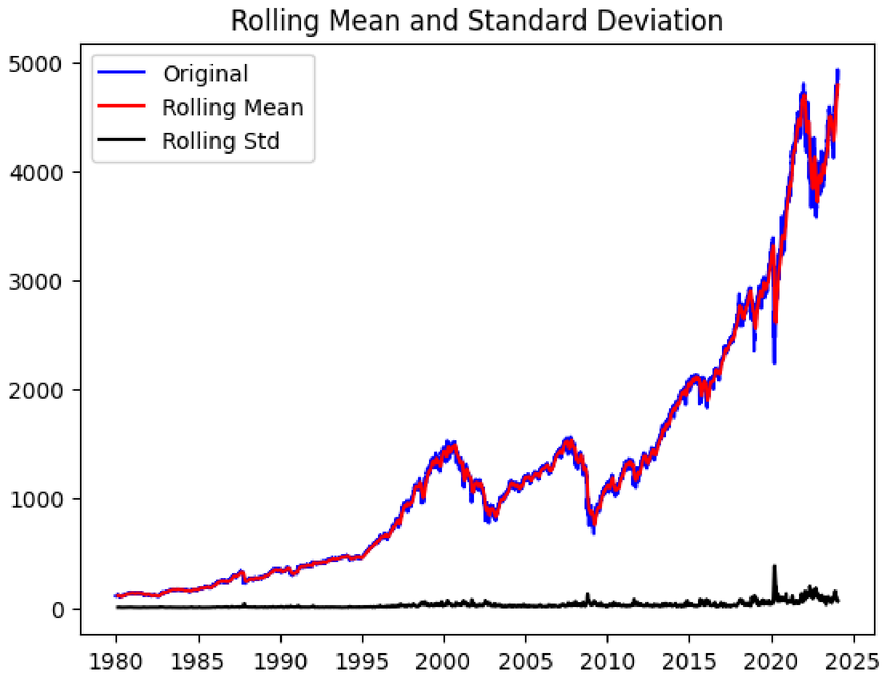

Figure 8. Rolling Mean and Std of S&P 500 closing price

As we try to plot the rolling mean and standard deviation, we can find that the mean and standard deviation seem to increase over time. Now, we use the Dickey-Fuller test to test stationarity in a time-series.

- Test Statistics (2.418): This is the actual value of the test statistic (t-statistics)
- p-value (0.999): This is used to either support or reject the null hypothesis in the test. It reflects the chance that the test findings occurred at random. If p-value is close to 1 (as in this case), it's a strong indicator that the time series is non-stationary.
- No. of lags used (37): This is the number of lagged observations used in the regression when performing the test.
- Number of observations used (11076): This is the number of observations that were used in the regression.
- Critical values: These are the test statistic values at which you reject the null hypothesis if the calculated test statistic is less than the critical value. If the Test Statistics is more than the Critical Value, we fail to reject the null hypothesis and infer that the time series is non-stationary.

We also try to split the dataset into 90% train data and 10% test data to determine the performance of the ARIMA model.

Figure 9. Train and Test Data Split

After that, we will use the auto-ARIMA method to determine the best order parameters for the model. The generated best model is $ARIMA(2, 1, 0)$.

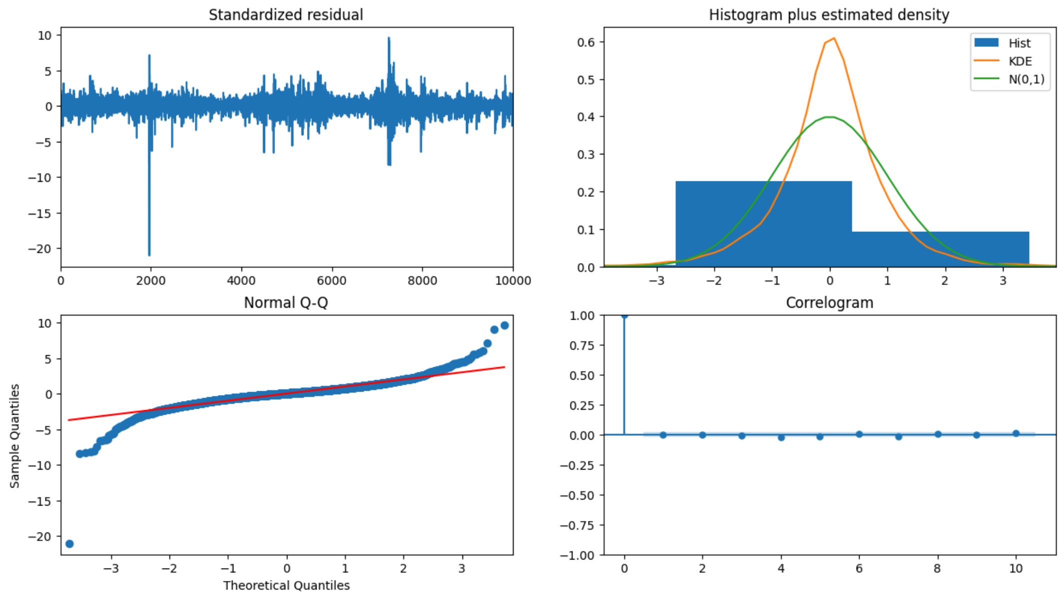

Figure 10. Auto-ARIMA plot diagnostics

The Standardized Residual seems to possess mean 0 and variance 1. The density histogram seems to follow normal distribution with mean 0. The red line seems to align with most of the dots. The ACF plot suggests that no significant point existed.

At the end, we got the forecast report performance as follows: MSE: 0.1026, MAE: 0.288, RMSE: 0.32 and MAPE: 0.0346. The MAPE of 3.46% suggests that the model is 96.54% accurate in predicting the next sequential observations. Besides, we can conclude that this S&P 500 time series does not seem stationary.

7. Conclusion

As the results from the above code and calculation, we can conclude that multivariate data analysis can determine the chance of systemic risk happening. Identifying and analyzing systemic risk is one thing. Predicting the exact time and intensity requires more resources and expertise. General time-series analysis approach like ARIMA is not enough in predicting the trends and unique events of the real-world situation. More complex methods like neural networks and machine learning are necessary.

8. Disclaimer

This paper was created as part of a WorldQuant University degree program towards an MSc in Financial Engineering.

9. Acknowledgement

The authors would like to thank the instructor and WorldQuant University for providing continuous support and guidance with openly accessible resources.

References

[1] Adamjee, U. (2023, March 4). Exploratory Data Analysis on Stock Market Data | by Uzair Adamjee | MLearning.ai. Medium. Retrieved January 15, 2024, from https://medium.com/mlearning-ai/exploratory-data-analysis-on-stock-market-data-5d99fbdf3b04

[2] Borsalli, P. B. (2022, February 28). 7 methods and tools for risk identification. Excellence Blog. Retrieved January 16, 2024, from https://blog.softexpert.com/en/risk-identification/

[3] CFI Team. (n.d.). Autoregressive Integrated Moving Average (ARIMA). Corporate Finance Institute. Retrieved February 13, 2024, from https://corporatefinanceinstitute.com/resources/data-science/autoregressive-integrated-moving-average-arima/

[4] Courtnell, J. (2020, July 3). Systemic Risk Explained, How to Measure and Identify Systemic Risk Items. Process Street. Retrieved January 16, 2024, from https://www.process.st/systemic-risk/

[5] Craig, B. R. (2020, December 3). How Well Does the Cleveland Fed's Systemic Risk Indicator Predict Stress? Federal Reserve Bank of Cleveland. Retrieved January 15, 2024, from https://www.clevelandfed.org/publications/economic-commentary/2020/ec-202028-how-well-does-systemic-risk-indicator-predict-stress

[6] Dhaduk, H. (2021, July 18). Stock market forecasting using Time Series analysis With ARIMA model. Analytics Vidhya. Retrieved February 13, 2024, from https://www.analyticsvidhya.com/blog/2021/07/stock-market-forecasting-using-time-series-analysis-with-arima-model/

[7] Dičpinigaitienė, V., & Novickytė, L. (2018, October 09). Application of systemic risk measurement methods: A systematic review and meta-analysis using a network approach. Quantitative Finance and Economics. Retrieved January 16, 2024, from https://www.aimspress.com/article/10.3934/QFE.2018.4.798

[8] Federal Funds Effective Rate (DFF) | FRED | St. Louis Fed. (n.d.). Federal Reserve Economic Data. Retrieved February 12, 2024, from https://fred.stlouisfed.org/series/DFF

[9] Hartwig, B., Meinerding, C., & Schüler, Y. S. (2021, July 22). Identifying indicators of systemic risk. ScienceDirect. Retrieved January 16, 2024, from https://www.sciencedirect.com/science/article/abs/pii/S0022199621000921

[10] Huang, W., Lan, C., Xu, Y., Zhang, Z., & Zeng, H. (2022, July 16). Does COVID-19 matter for systemic financial risks? Evidence from China's financial and real estate sectors. NCBI. Retrieved January 15, 2024, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9366997/

[11] Hullman, J., & Gelman, A. (2021, July 30). Designing for Interactive Exploratory Data Analysis Requires Theories of Graphical Inference. Harvard Data Science Review. Retrieved January 15, 2024, from https://www.stat.columbia.edu/~gelman/research/published/hullman_gelman_hdsr.pdf

[12] IMF Data Mapper ®. (n.d.). IMF Data Mapper ®. Retrieved February 12, 2024, from https://www.imf.org/external/datamapper/profile/ADVEC

[13] Jain, A. (n.d.). Value at Risk (VaR). WallStreetMojo. Retrieved January 16, 2024, from https://www.wallstreetmojo.com/value-at-risk/

[14] Li, W. (2016, April 25). Value at Risk (VaR) and its calculations: an overview. MST.edu. Retrieved January 16, 2024, from https://web.mst.edu/~huwen/teaching_VaR_Weiqian_Li.pdf

[15] Ma, J., Li, X., Liu, J., Cui, J., Zhang, M., & Sriboonchitta, S. (2022, November 25). Comparison of Systemic Financial Risks in the US before and after the COVID-19 Outbreak—A Copula–GARCH with CES Approach. MDPI. Retrieved January 15, 2024, from https://www.mdpi.com/2075-1680/11/12/669

[16] Soni, S. K. (2023, February 27). The Power of Exploratory Data Analysis (EDA) in Data Science: Understanding the Basics and Best Practices. LinkedIn. Retrieved January 16, 2024, from https://www.linkedin.com/pulse/power-exploratory-data-analysis-eda-science-basics-best-soni

[17] Sun, T., Mills, P., Segoviano, M., Jobst, A., Hesse, H., Gray, D., & Capuano, C. (2009, April 23). Methods to identify systemic financial risks. CEPR. Retrieved January 15, 2024, from https://cepr.org/voxeu/columns/methods-identify-systemic-financial-risks