



# Stock Market Analysis

## Aim :

The objective is to develop a community detection algorithm specifically designed for identifying subsets of stocks that exhibit similar behaviours. The effectiveness of this approach will be assessed based on various metrics, including Separability, Clustering Coefficient, Fraction over Median Degree, Conductance, Normalized Cut, and Cut Ratio.

## Abstract :

Community detection, a powerful tool in network analysis, involves partitioning a network into groups of nodes that share common characteristics or exhibit strong connections among themselves. In the context of financial markets, this translates into identifying subsets of stocks that move in tandem, potentially revealing underlying patterns or market sectors.

This project explores the dynamic landscape of financial markets by leveraging community detection algorithms on stock data obtained from Yahoo Finance. The primary objective is to unveil subsets of stocks sharing similar behaviour's, revealing underlying patterns and market dynamics. The proposed approach focuses on the evaluation of key metrics, including Separability, clustering coefficient, Fraction over median degree, conductance, Normalized cut, and Cut Ratio, to assess the effectiveness of the community detection algorithm.

Utilizing network analysis techniques, the project partitions the stock market into cohesive communities, identifying stocks that move in tandem. The Separability metric ensures clear demarcations between these communities, emphasizing the distinctiveness of the identified subsets. Clustering coefficient and Fraction over median degree offer insights into the internal cohesion and degree distribution within communities, enhancing the understanding of network structures.

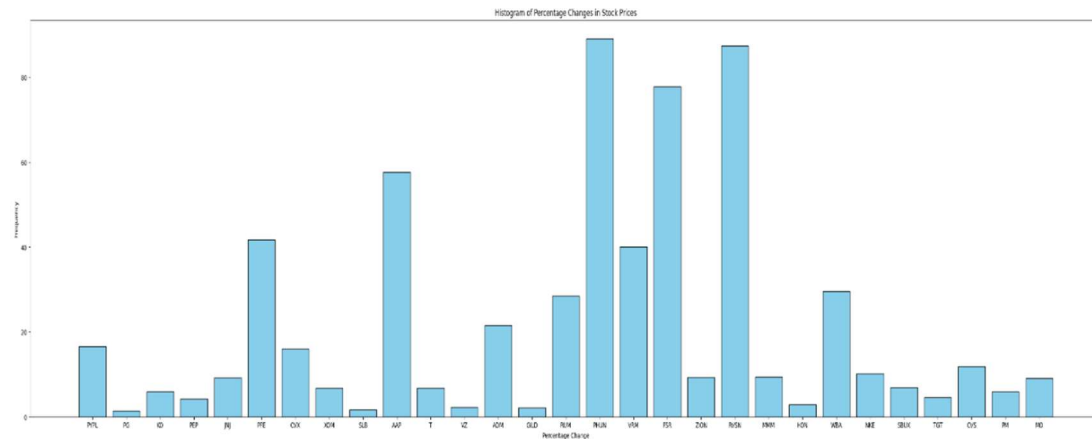
Conductance and Normalized cut evaluate the quality of partitions, considering the balance between internal and external connections. The Cut Ratio metric further contributes to the assessment, offering a comprehensive evaluation of community balance and cohesion. The project utilizes financial data sourced from Yahoo Finance, including historical stock prices and trading volumes, to construct a meaningful network for analysis.

## Introduction :

In the dynamic realm of financial markets, understanding the interconnectedness and shared behaviours among stocks is crucial for making informed investment decisions. The aim of this project is to employ community detection algorithms to unveil subsets of stocks that exhibit similar behaviours. By leveraging network analysis techniques, we delve into the intricate relationships and patterns that may exist within the stock market. The proposed approach will be rigorously evaluated using key metrics, including Separability, clustering coefficient, Fraction over median degree, conductance, Normalized cut, and Cut Ratio.

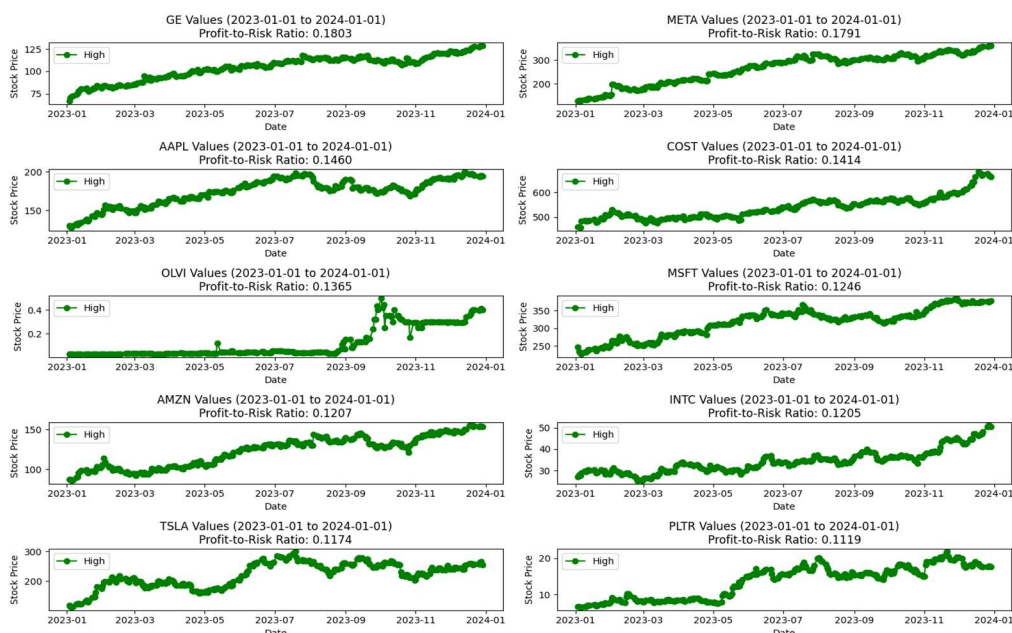
**Separability:** It quantifies the degree of separation between communities by comparing the actual connections within communities to the expected connections in a random network. Higher modularity values suggest stronger separability between communities.

Separability is a fundamental metric in community detection, measuring the distinctiveness of identified communities. In our project, we seek to ensure that the detected subsets of stocks are well-separated, signifying clear demarcations between different groups. This metric will be pivotal in evaluating the efficacy of our community detection algorithm.



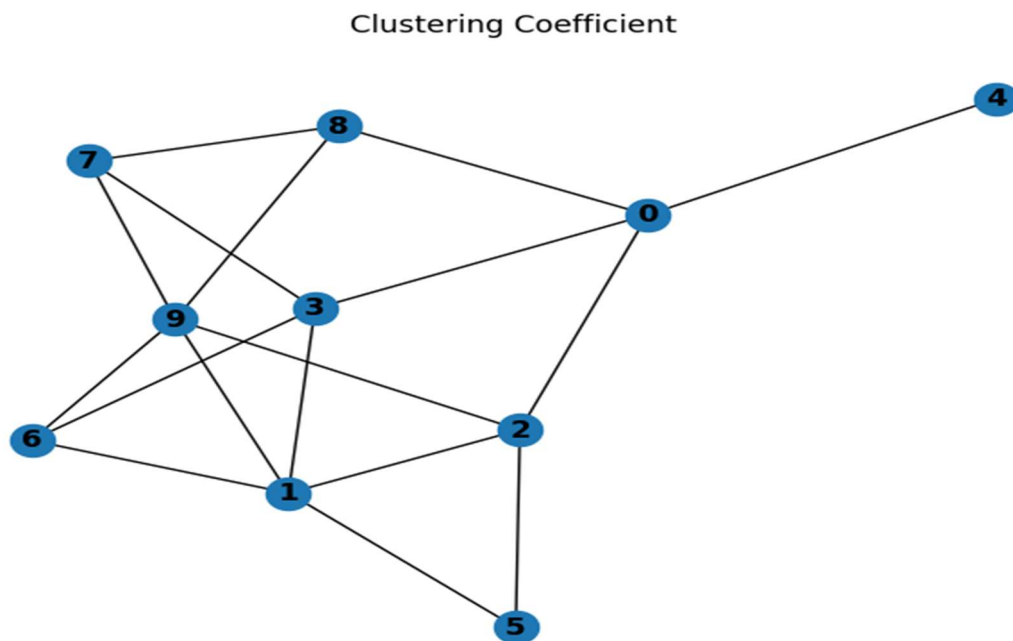
#### Top 10 Most Profitable Stocks (Profit-to-Risk Ratio):

GE: 0.1803  
 META: 0.1791  
 AAPL: 0.1460  
 COST: 0.1414  
 OLVI: 0.1365  
 MSFT: 0.1246  
 AMZN: 0.1207  
 INTC: 0.1205  
 TSLA: 0.1174  
 PLTR: 0.1119



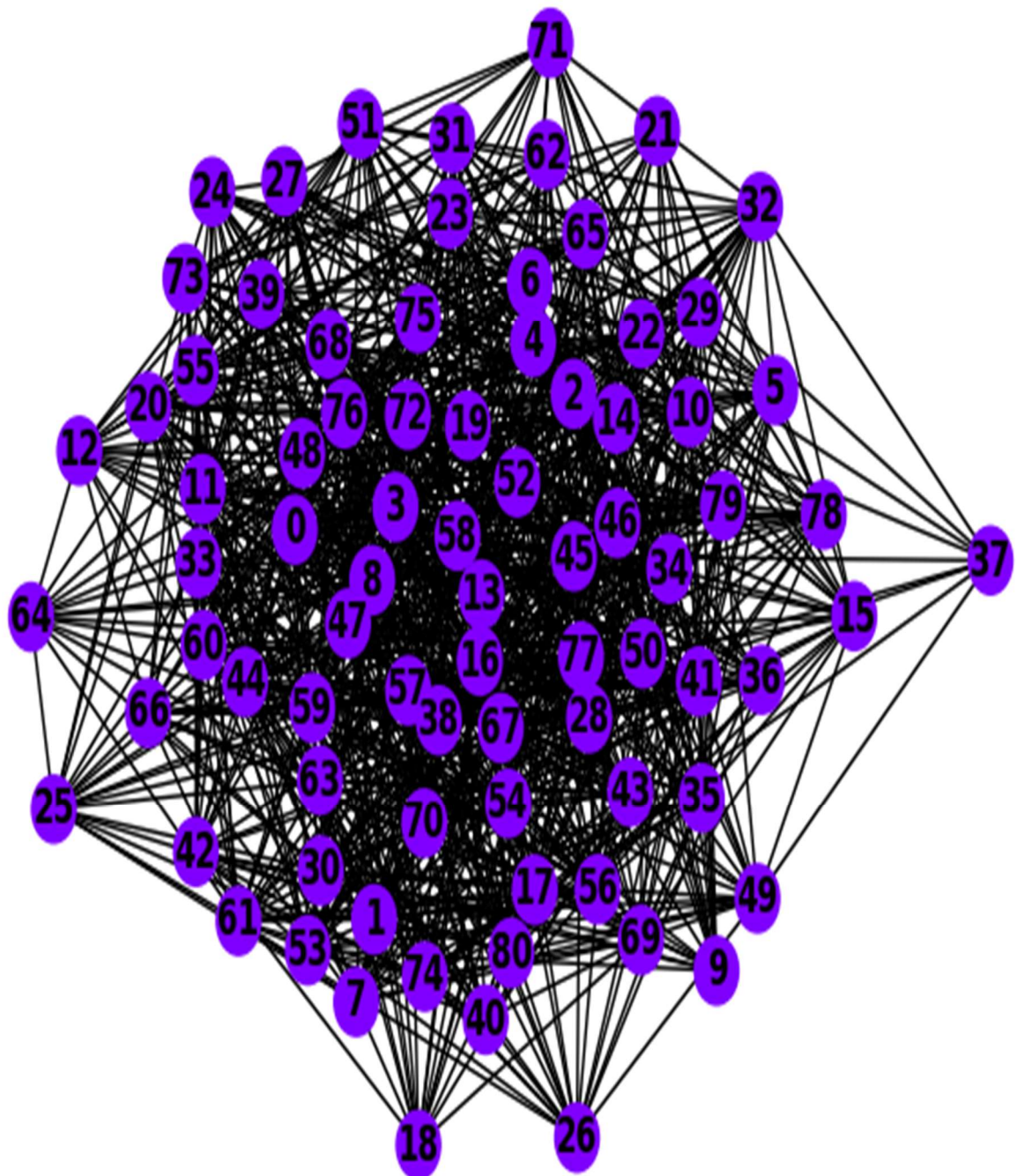
## Clustering Coefficient:

The clustering coefficient provides insights into the density of connections within identified communities. A high clustering coefficient implies strong internal connections, reinforcing the idea of a cohesive group of stocks with similar behaviours.





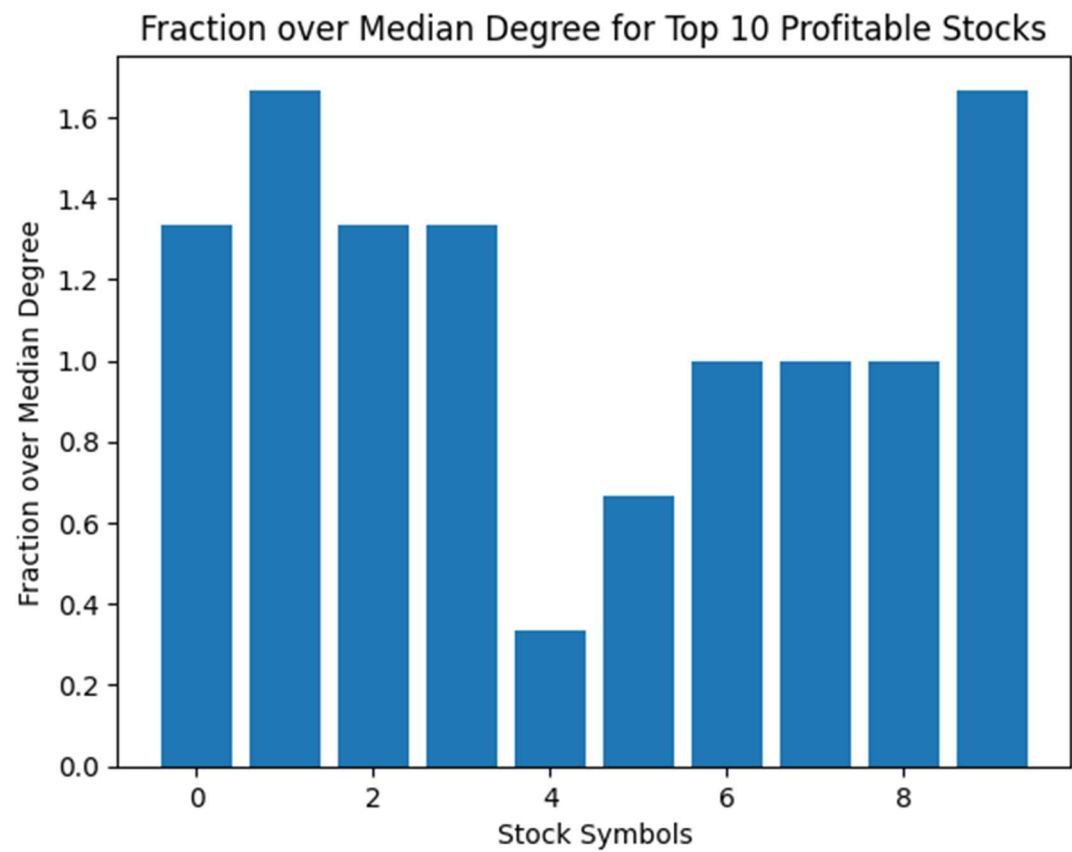
## Louvain Communities





## Fraction over Median Degree:

Fraction over Median Degree offers a comparative measure of the degree distribution within communities, contributing to a nuanced understanding of network structures.



Fraction over Median Degree:

0: 1.3333333333333333  
1: 1.6666666666666667  
2: 1.3333333333333333  
3: 1.3333333333333333  
4: 0.3333333333333333  
5: 0.6666666666666666  
6: 1.0  
7: 1.0  
8: 1.0  
9: 1.6666666666666667



## Conductance:

Conductance measures the quality of partitions by assessing the ratio of external connections to the total degree of nodes in a community. Lower conductance values indicate well-defined communities with minimal connections to the external network.

Conductance

0: 1.0

1: 1.0

2: 1.0

3: 1.0

4: 1.0

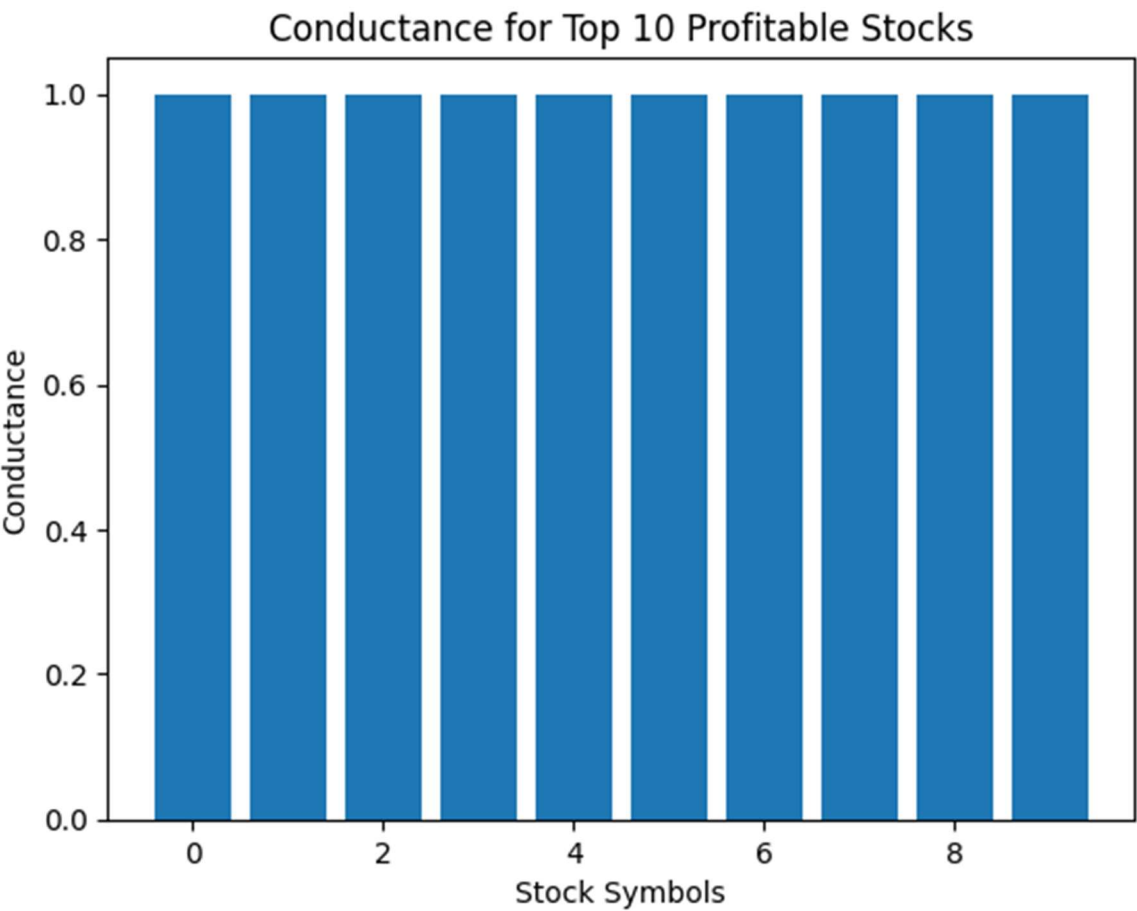
5: 1.0

6: 1.0

7: 1.0

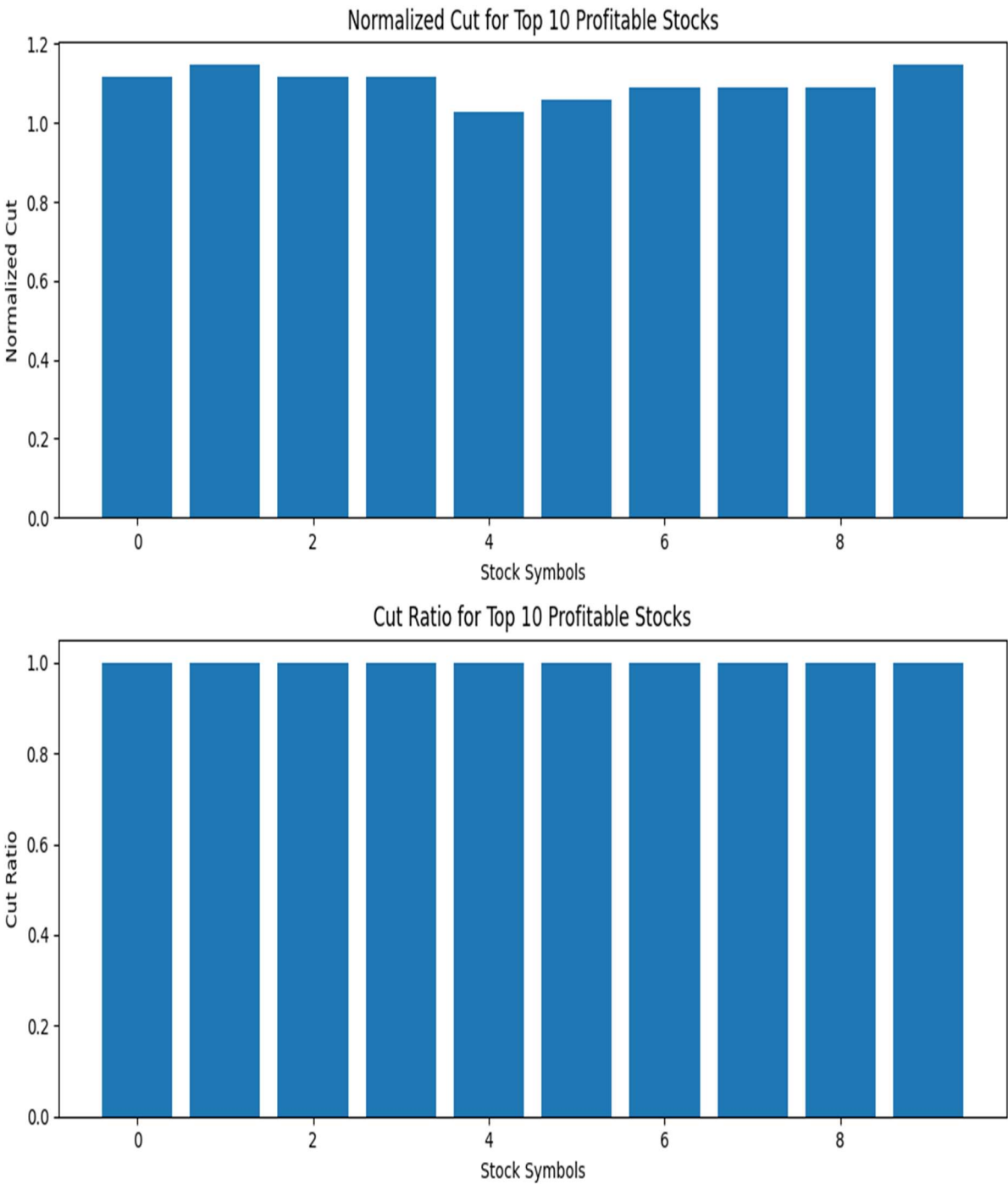
8: 1.0

9: 1.0

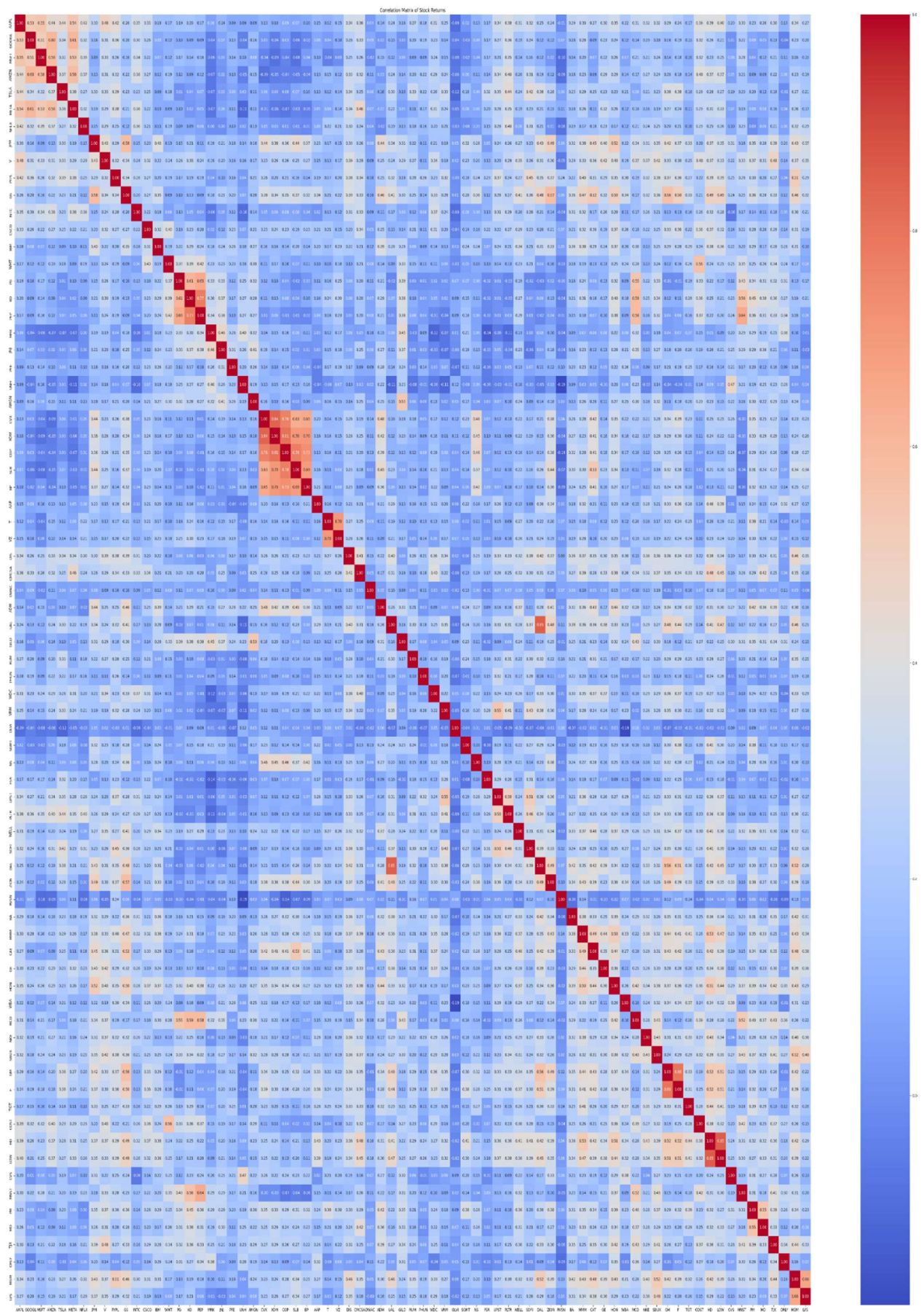


## Normalized Cut and Cut Ratio :

Normalized cut extends this evaluation by considering the size of communities, providing a normalized perspective on the quality of partitions. The Cut Ratio is a metric that evaluates the balance between the internal and external connections of identified communities. A lower cut ratio suggests more cohesive and internally connected communities, adding another layer of evaluation to the effectiveness of the community detection algorithm.



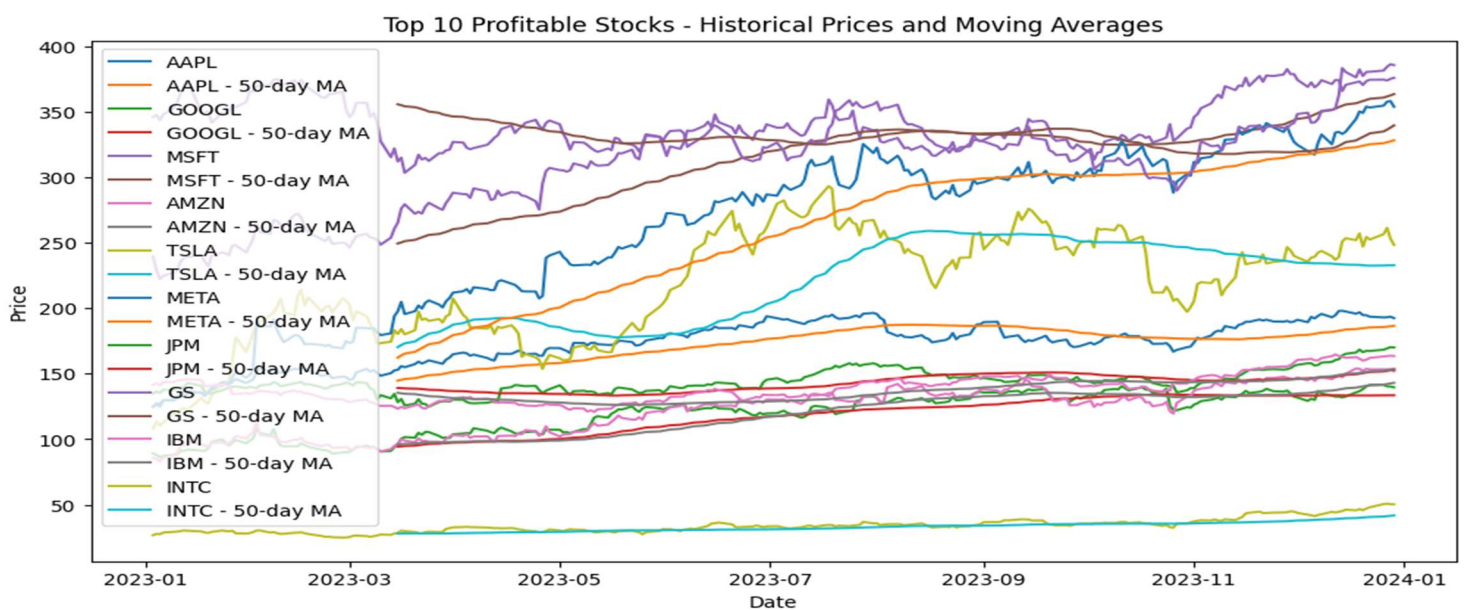
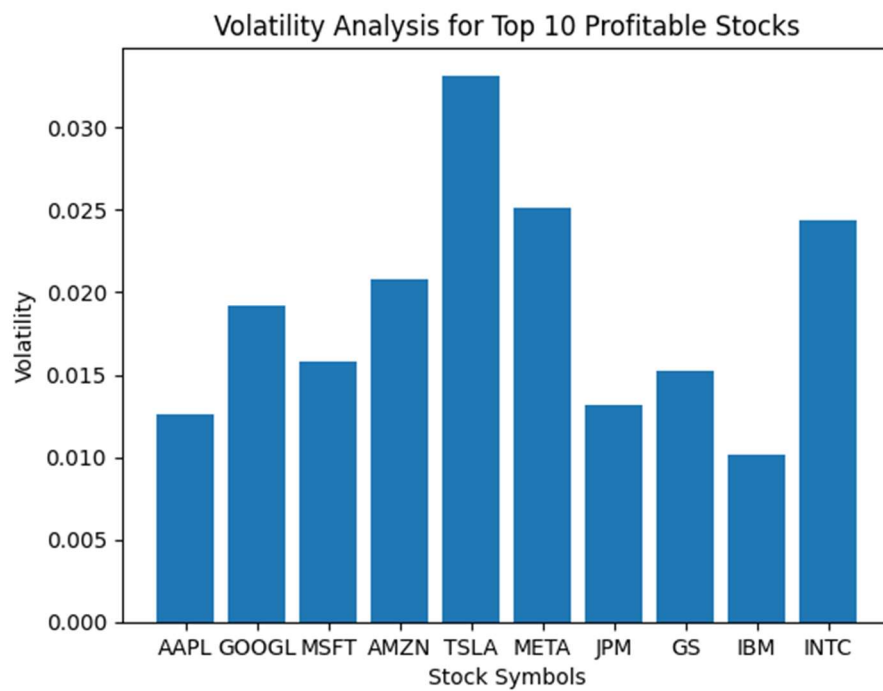






## Volatility:

The standard deviation of returns is computed for each stock, representing the volatility of each stock. A bar chart is created to visualize the volatility of each stock. The x-axis represents stock symbols, and the y-axis represents volatility. The bars are in descending order of volatility, adds a horizontal line indicating the average volatility, and rotates the x-axis labels for better readability.





## Conclusion:

In conclusion, the project aims to contribute to the understanding of stock market dynamics through the application of community detection algorithms. By unveiling subsets of stocks sharing similar behaviours, we aspire to provide investors and analysts with valuable insights for portfolio optimization, risk management, and strategic decision-making. The evaluation framework, encompassing Separability, clustering coefficient, Fraction over Median Degree, conductance, Normalized cut, and Cut Ratio, will serve as a comprehensive assessment of the algorithm's effectiveness in uncovering meaningful patterns within the financial network. As financial markets continue to evolve, this project stands at the intersection of data science and finance, offering a promising avenue for innovative insights and discoveries.

## Reference:

1. Zhou, Y., Chen, Z., & Liu, Z. (2023). Dynamic analysis and community recognition of stock price based on a complex network perspective. *Expert Systems with Applications*, 213, 118944.
2. K. Berahmand, M. Mohammadi, F. Saberi-Movahed, Y. Li and Y. Xu, "Graph Regularized Nonnegative Matrix Factorization for Community Detection in Attributed Networks," in *IEEE Transactions on Network Science and Engineering*, vol. 10, no. 1, pp. 372-385, 1 Jan.- Feb. 2023, doi: 10.1109/TNSE.2022.3210233.
3. Naik, D., Ramesh, D., Gandomi, A. H., & Gorojanam, N. B. (2022). Parallel and distributed paradigms for community detection in social networks: A methodological review. *Expert Systems with Applications*, 187, 115956.
4. Sattar, N. S., Buluc, A., Ibrahim, K. Z., & Arifuzzaman, S. (2023). Exploring temporal community evolution: algorithmic approaches and parallel optimization for dynamic community detection. *Applied Network Science*, 8(1), 64.
5. Yanan Yan, Yuehan Yang. Community detection for New York stock market by SCORE-CCD.