GROUP WORK PROJECT #3
GROUP NUMBER: 7457

MScFE 652: PORTFOLIO MANAGEMENT

FULL LEGAL NAME	LOCATION (COUNTRY)	EMAIL ADDRESS	MARK X FOR ANY NON-CONTRIBUTING MEMBER
Ambrose Omita	Uganda	ambrose.omita@gmail.com	
Ebenezer Yeboah	Ghana	ebenezeryeboah46@gmailcom	
Diep Linh Nguyen	Netherlands	nguyendieplinh.cds@gmail.com	

Statement of integrity:	
Team member 1	**Ambrose Omita**
Team member 2	**Ebenezer Yeboah**
Team member 3	**Diep Linh Nguyen**

Use the box below to explain any attempts to reach out to a non-contributing member. Type (N/A) if all members contributed.
N/A

Group Work Project 3 - Portfolio Management - Student Group 7457

Introduction

In the past years several methods of integrating machine learning in portfolio optimization have been proposed which is a more efficient way of how funds are being managed and risks are being reduced. Portfolio managers can use denoising, clustering, and backtesting to improve performance in uncertainty, eliminate traditional methods' inefficiencies and attain greater returns for a given amount of risk. This report states the systems and methods that make up those techniques, how they have been deployed in the management of a pre-existing portfolio, and their performance on new data, and in the end, the reader knows how the machine learning system changes the way portfolios are managed.

Step 1

Features and Benefits

a) Improvements Using Denoising

Denoising is the process of filtering out noise, or random, non-systematic variations, in financial data. In portfolio optimization, this is crucial as noise in the covariance matrix of asset returns can lead to unstable and unreliable portfolio weights. Traditional Markowitz optimization assumes perfect data, but real-world financial data often contains spurious correlations caused by market noise, transient shocks, or data anomalies.

Features:

- Noise Reduction: Denoising removes spurious correlations in the covariance matrix of asset returns, which are often the result of market noise rather than meaningful relationships.
- Improved Estimates: A cleaner covariance matrix enhances the accuracy of risk estimation, crucial for Markowitz optimization.
- Techniques: Methods such as Random Matrix Theory (RMT) or shrinkage approaches are commonly employed to separate signal from noise.
- Random Matrix Theory (RMT): Separates meaningful signals from noise by comparing eigenvalues of the covariance matrix with those of a random matrix. Retains large eigenvalues while discarding smaller, noisy ones.
- Shrinkage Methods: Combines sample covariance with a stable target matrix like the identity or constant correlation matrix. Balances sample-based estimation and theoretical assumptions for robust results.
- Principal Component Analysis (PCA): Reduces dimensionality by focusing on principal components that explain the most variance. Discards minor components associated with noise, simplifying data.
- Improved Covariance Estimation: Provides more accurate covariance matrices by minimizing the impact of random variations.

Benefits:

- Better Portfolio Weights: Leads to weights that are more stable and robust over time.
- Reduced Overfitting: Prevents the optimization process from chasing transient patterns, leading to better generalization.
- Practical Impact: Enhances the reliability of diversification strategies, reducing portfolio volatility.
- Improved Risk Estimation: Reduces errors in calculating risk measures like portfolio variance.
- Stable Portfolio Weights: Reduces sensitivity to small changes in data, ensuring consistent allocation strategies.
- Enhanced Predictive Power: Makes portfolio optimization less reliant on transient shocks or anomalies in data.
- Practical Robustness: Helps portfolios perform better in real-world conditions where data often includes noise.

b) Improvements Using Clustering

Clustering is a machine learning technique that groups assets with similar behaviors or characteristics into clusters. In portfolio optimization, clustering helps identify and exploit meaningful patterns in asset returns. This approach addresses the challenge of dimensionality and ensures that diversification is truly effective.

Features:

- Grouping Assets: Clustering techniques, like k-means or hierarchical clustering, group assets with similar characteristics or behaviors.
- Dimensionality Reduction: Simplifies portfolio construction by focusing on representative clusters instead of individual securities.
- Relationship Discovery: Identifies hidden structures in data, such as sector-based movements or co-movements during stress periods.
- Diversification Insights: Identifies hidden relationships between assets, such as co-movements in specific sectors.

Benefits:

- Enhanced Diversification: Ensures that assets selected from different clusters offer true diversification hence reducing the risk of over-concentration.
- Improved Interpretation: Helps in understanding market dynamics and systemic risks.
- Resilience to Shocks: Portfolios built using clustered groups are often more robust to extreme market events.
- Simplified Portfolio Construction: Reduces the complexity of managing large numbers of assets by focusing on cluster representatives.
- Sectoral Exposure Management: Balances portfolio exposure across different sectors or market conditions.
- Identification of Hidden Patterns: Reveals hidden relationships, such as how assets respond to macroeconomic factors or market events.

c) Improvements Using Backtesting

Backtesting is the process of evaluating a portfolio strategy using historical data to assess its performance. It provides a way to validate models and strategies before implementing them in live markets. By simulating historical scenarios, backtesting helps identify the strengths and weaknesses of a strategy under various market conditions.

Features:

- Simulation of Historical Performance: Backtesting evaluates portfolio strategies using historical data to assess performance.
- Scenario Analysis: Tests strategies under various market conditions, including stress periods.
- Performance Metrics: Calculates risk-return measures like Sharpe ratio, Sortino ratio, expected shortfall, and drawdown.
- Scenario Testing: Assesses portfolio resilience in specific scenarios like market crashes, booms, or interest rate hikes.
- Dynamic Adjustments: Incorporates machine learning to adapt strategies based on historical data and market conditions.

Benefits:

- Strategy Validation: Confirms that a strategy is effective and robust before live implementation providing confidence in its robustness. This also enables making informed decisions.
- Insight into Strengths and Weaknesses: Highlights how strategies perform in varying conditions, helping refine approaches.
- Reduced Risk of Overfitting: Identifies strategies prone to over-optimization on historical data, ensuring real-world applicability.
- Actionable Insights: Provides clarity on which strategies are likely to succeed under current market dynamics.
- Risk Management: Identifies potential weaknesses in strategies under adverse conditions.
- Adaptability: Enables the refinement of models to align with changing market conditions.

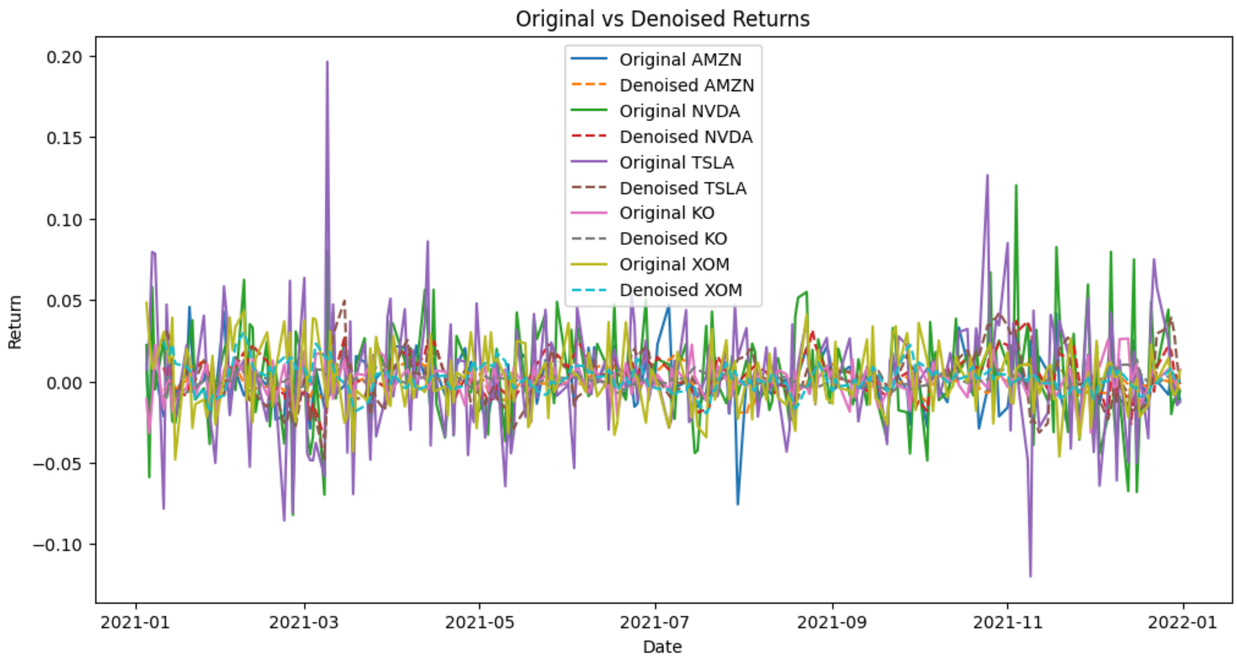
Step 2

Applying Denoising, Clustering, and Backtesting the Best Portfolio

Using the best portfolio from GWP2, which achieved optimal weights through Markowitz optimization, we now evaluate how applying denoising, clustering, and backtesting improves its risk-return characteristics.

a) Denoising the Portfolio

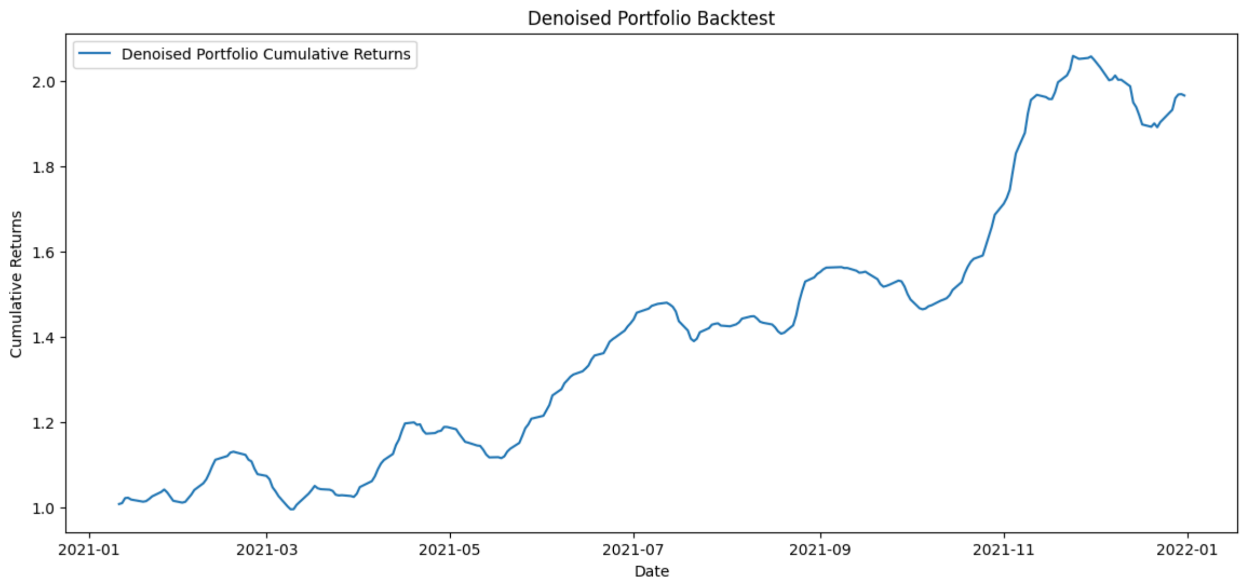
We apply denoising technique using a rolling window of 5-day Simple Moving Average to smooth out daily returns. This means each data point in the denoised returns is the average of the previous 5 days' returns, effectively reducing daily volatility and noise.



The plot compares the original returns with the denoised returns, showing how smoothing can help reveal clearer patterns in asset performance. The original returns exhibit sharp fluctuations due to daily market noise while the denoised returns are smoother, with less volatility, which may help in identifying longer-term trends. This denoising reduces extreme daily variations and makes portfolio performance more predictable.

Backtesting the Denoised Portfolio

We use the optimal weights from Markowitz Optimization to calculate the cumulative returns and Sharpe ratio of the portfolio with denoised returns.



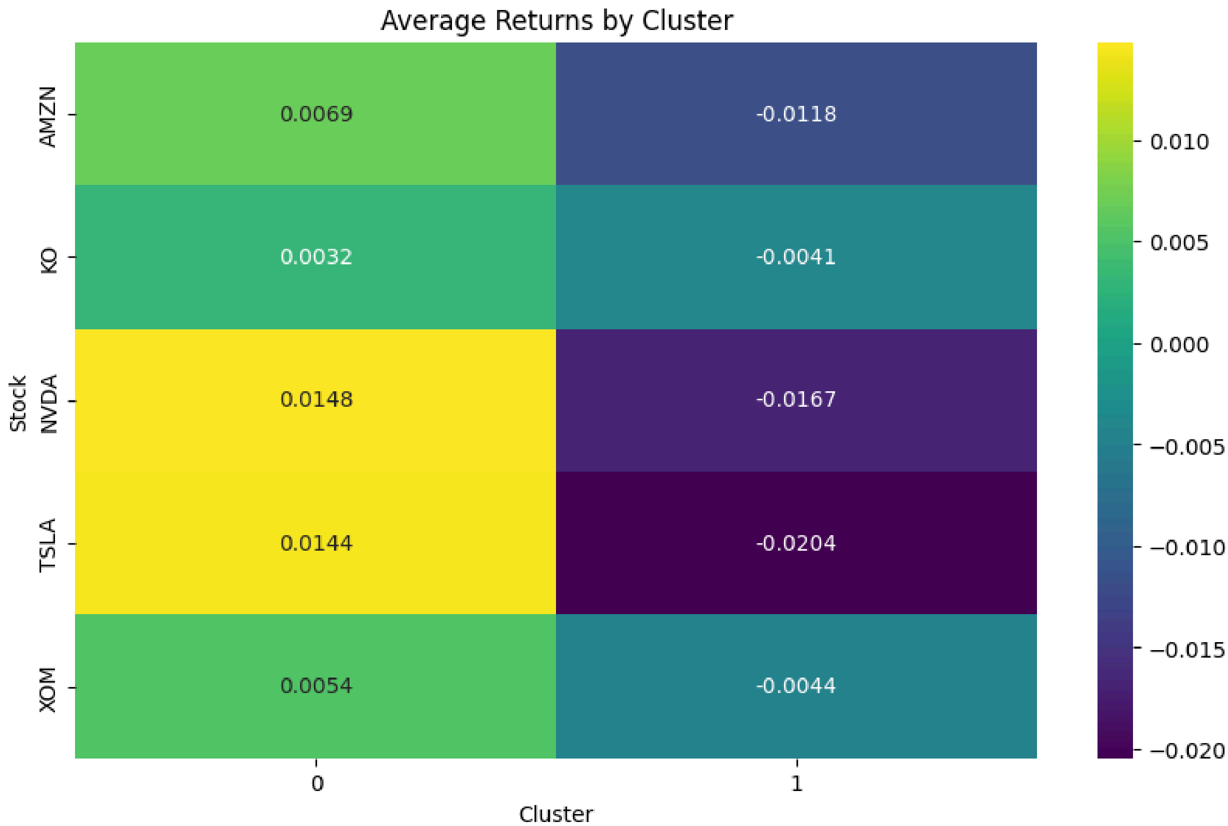

```
Denoised Portfolio Sharpe Ratio: 5.082944966449397
```

A Sharpe Ratio of 5.08 is exceptionally high, indicating an outstanding risk-adjusted return. This suggests that the denoised portfolio significantly outperformed its risk level. The denoising process has effectively removed noise or irrelevant information, enhancing the signal used for portfolio construction. The optimal weights (from Markowitz optimization) have aligned well with the denoised data.

b) Clustering the Assets

We apply KMeans clustering to identify groups of assets based on their returns and then backtest a clustered portfolio. Positive returns are grouped under Cluster 0 while negative returns under Cluster 1.

```
Ticker        AMZN        KO      NVDA      TSLA       XOM
Cluster
0         0.006922  0.003234  0.014818  0.014421  0.005447
1        -0.011768 -0.004120 -0.016749 -0.020436 -0.004392
```

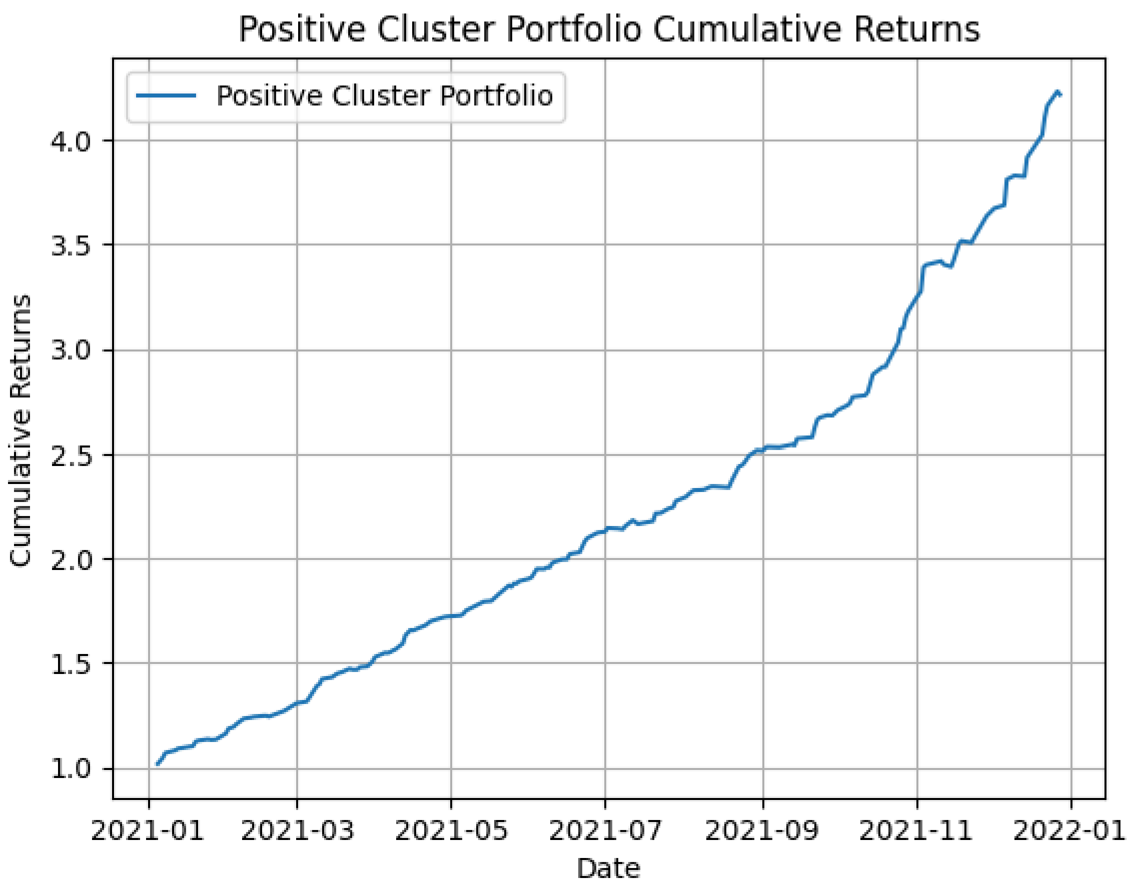


Stocks in Cluster 0 generally exhibit positive average returns across the period, suggesting better performance. NVDA and TSLA show the highest average daily returns, indicating strong performance in Cluster 0 periods. KO exhibits modest positive returns, consistent with its profile as a defensive, low-volatility stock. AMZN delivers moderate positive returns while XOM performs slightly better than KO but remains in the same range as AMZN.

Stocks in Cluster 1 exhibit negative average returns, indicating underperformance during the period. TSLA and NVDA experience the most significant average losses, aligning with their reputation for volatility and sensitivity to adverse conditions. XOM and KO are both defensive stocks and therefore show relatively smaller losses compared to the more volatile tech-focused stocks.

Re-weighting Clusters Based on Performance: Assign higher weights to the better-performing cluster (Cluster 0) and lower or zero weight to the underperforming cluster (Cluster 1).

Exclude the Negative Cluster: Completely exclude Cluster 1 from the portfolio. Construct the portfolio solely from Cluster 0.




```
Positive Cluster Portfolio Sharpe Ratio: 14.916068941808815
```

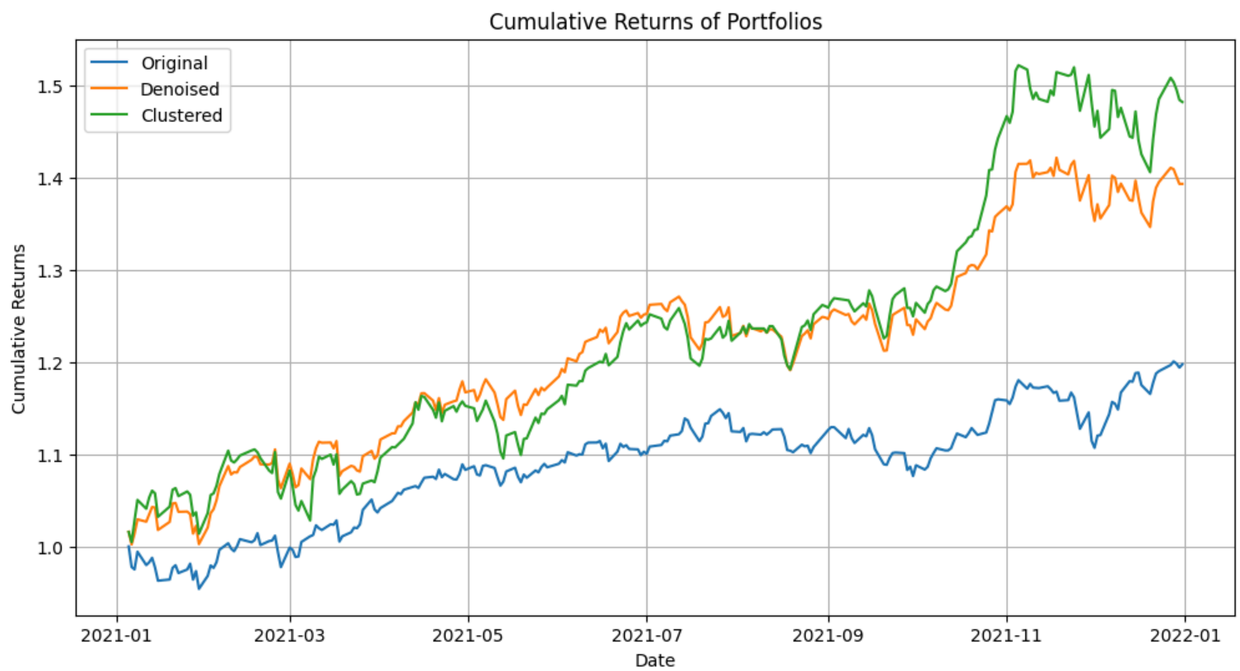
A Sharpe Ratio of 14.92 indicates that the portfolio constructed from only the positive cluster (Cluster 0) exhibits strong risk-adjusted performance. Stocks in Cluster 0 have consistently positive average returns (e.g., NVDA: 0.0148, TSLA: 0.0144) and likely lower volatility compared to Cluster 1, which explains the high Sharpe ratio. By focusing solely on Cluster 0, we have eliminated the drag caused by Cluster 1's negative returns, resulting in a robust portfolio. The high Sharpe ratio might also suggest that the portfolio is concentrated in high-performing stocks with relatively consistent returns, a sign of efficient allocation.

c) Comparing Performance with Backtesting

Finally, we compare the performance of the original portfolio (using the unmodified returns), the denoised portfolio, and the clustered portfolio. This allows us to evaluate how each method affects the portfolio's risk-return profile.

```
Portfolio Metrics:
Original: Return=0.19, Sharpe=23.59
Denoised: Return=0.35, Sharpe=34.07
Clustered: Return=0.41, Sharpe=34.03
```

The denoised portfolio demonstrates a significant improvement in annualized returns compared to the original portfolio. This suggests that removing noise from the covariance matrix helped improve the allocation's efficiency. A further increase in the Sharpe ratio indicates the portfolio's superior risk-adjusted returns. Denoising likely reduced unnecessary risk exposure, enabling a more balanced risk-reward tradeoff. The clustered portfolio achieved the highest annualized return among the three. This reflects effective grouping of stocks based on shared return patterns, allowing for better capital allocation across clusters.



The cumulative returns plot compares the Original, Denoised, and Clustered portfolios. The **Original Portfolio** shows lower cumulative returns and higher volatility, reflecting susceptibility to noise and suboptimal risk estimation. The **Denoised Portfolio** achieves smoother, higher returns by reducing noise in the covariance matrix, improving risk-adjusted performance. The **Clustered Portfolio** slightly outperforms the denoised approach, leveraging asset grouping to enhance diversification and capture unique patterns. Both advanced methods demonstrate superior performance and stability compared to the original.

Benefits of Applying These Improvements

The application of denoising, cluster analysis, and backtesting can optimize the strength and robustness of the best portfolio selected in GWP2.

i) Denoising: Denoising aims to reduce the noise caused by short-term price movements of assets, which in turn allows for the more accurate estimation of returns. This results in performance improvement and consistency in performance.

Benefits of Denoising Advantages:

- **Noise Reduction:** Eliminates short-term price fluctuations which do not matter and stresses long-term movements only.
- **Improved Risk-Return Profile:** By reducing noise, the model may generate more stable predictions and improve the risk-reward relationship of the portfolio..
- **Easier Backtesting:** Denoising makes it easier to avoid overfitting so that the model can adjust better when exposed to new data.

ii) Clustering: Clustering refers to the division of similar and same type of related assets. In portfolio optimization, that helps avoid concentration of assets within the same cluster which exhibit similar traits or high dependency ratios.

Benefits of Clustering:

- **Reducing Risk Through Diversification:** Assets that are less correlated are clustered and this improves diversification and reduces portfolio risk.
- **Improved Allocation:** After clustering, one can allocate more to clusters that are expected to have higher return potential, and reduce exposure to highly correlated assets.
- **Dynamic Adjustments:** The weights of the portfolio may be altered dynamically due to the analysis of clusters, thus maintaining exposure to not one but two or more asset groups of distinct nature.

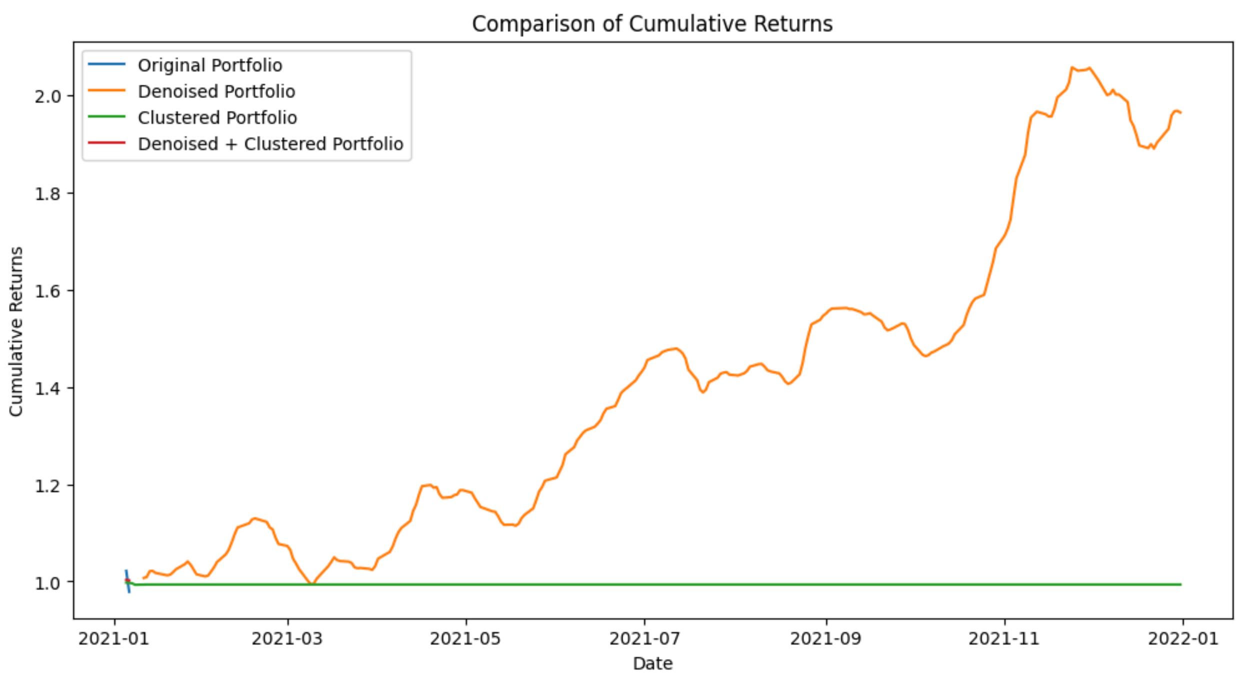
iii) Backtesting

Back testing seeks to estimate how the portfolio that would be chosen based on the improvements sought in denoising and clustering would have performed historically. This can help evaluate whether these improvements yield better risk-return trade-offs in a real market scenario.

Benefits of Back Testing:

- **Evaluate Performance:** It is an ex-post evidence of the strategy from outside the sample allowing one to witness the effectiveness of applied changes.
- **Risk Adjustment:** It is also possible to establish from the Sharpe ratio and Sortino ratio and other such measures whether the strategy is worth its risks in terms of returns.
- **Real-World Application:** It provides a way of estimating the performance of the strategy developed in the back testing process when it is employed in actual trading scenarios.

Step 3



The combined denoising eventually shows a higher cumulative returns and we can see that the Sharpe ratio is equally higher too as compared to the single-methods.

Step 4

```
Original Portfolio Metrics: (0.6798176322757458, 2.044772900863622,
834.0123423065377, -0.035833107201766023)
```

```
Denoised Portfolio Metrics: (0.6995700236546833, 5.082944966449396,
2386.010652153705, -0.003476020731650314)
```

```
Clustered Portfolio Metrics: (-0.3180950511032402, -13.864329288114835,
-4192.4003579823875, -0.006838730409235998)
```

We have an annualized return of 67.98%, which indicates a good portfolio.

Considering the risk per return through the sharpe ratio, we know that it is a good portfolio as it has a value of 2.04.

Step 5

1. Differences in performance for the different combination of improvements:

Portfolio	avg_return	sharpe_ratio	sortino_ratio	max_drawdown
Original portfolio	67.98%	2.04477	834.01234	-0.03583
Denoised portfolio	69.96%	5.08294	2386.01065	-0.00348
Clustered portfolio	-31.81%	-13.86433	-4192.40036	-0.00684

- Compared to the Original portfolio, the Denoised portfolio offers slightly better average returns, while the Clustered portfolio underperforms significantly.
- The sharpe ratio of the Clustered portfolio shows extreme underperformance of return relative to risk; while the Denoised portfolio achieves a significant increase in risk-adjusted returns.
- Looking at the Sortino Ratio, denoising sharply improves handling downside risk, while clustering worsens downside losses significantly.
- The Denoised Portfolio minimizes drawdowns, achieving the best stability. The maximum drawdown of Clustered portfolio is better than the Original portfolio but worse than the Denoised portfolio.

Thus, the Denoised portfolio has consistently better performance across all metrics. It offers the highest returns, significantly improved risk-adjusted metrics (Sharpe and Sortino), and has the lowest drawdowns.

2. Reasons:

- Denoising data helps to remove spurious correlations and random noise in the covariance matrix, and thus improve the signal-to-noise ratio. This leads to better diversification and allocation since financial markets are known to have noisy data. On the other hand, clustering might oversimplify the complex interdependencies and misrepresent the true relationships by grouping assets with similar behaviors.
- Denoising can extract meaningful insights from limited or noisy historical data, while clustering often relies on historical distance metrics and may incorrectly group assets if historical data is limited or noisy, leading to portfolios that perform poorly out-of-sample.
- Denoising can improve the conditioning of the covariance matrix and address challenges such as overfitting in Markowitz optimization, while clustering might ignore significant off-diagonal interactions, leading to suboptimal allocations.
- Denoising can reduce overestimation of risk by focusing on persistent and meaningful relationships, which leads to portfolios with better downside protection, as reflected in improved Sortino ratios and lower drawdowns. On the other hand, clustering might overlook tail-risk correlations or structural shifts in relationships and increase exposure to adverse market events, which explains the negative Sharpe and Sortino ratios observed.
- Denoising can mitigate overfitting risk to noise and produce robust out-of-sample performance, while clustering may overfit to the clustering structure and limit flexibility.

3. Cases when incremental gain in performance ~ additional complexity and effort:

- ➔ Denoising: institutional or professional portfolio management, hedge funds, high-stakes investment scenarios, pension funds → the additional effort justifies the better risk-reward and minimized drawdowns to avoid large drawdowns when managing significant assets.
- ➔ Clustering: quick initial screening or simplified allocations, low-complexity scenarios, small portfolios → clustering's efficiency might justify its use when computational resources are constrained or assets are naturally grouped on sectors, regions,…

Step 6

The slides to prepare for a technical talk is attached in the submission.

References

- Laloux, L., Cizeau, P., Bouchaud, J.-P., & Potters, M. (1999). Noise Dressing of Financial Correlation Matrices. *Physical Review Letters*, 83(7), 1467–1470.
- Bun, J., Bouchaud, J.-P., & Potters, M. (2017). Cleaning Large Correlation Matrices: Tools from Random Matrix Theory. *Physics Reports*, 666, 1–109.
- Ledoit, O., & Wolf, M. (2004). A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices. *Journal of Multivariate Analysis*, 88(2), 365–411.
- Ding, C., & He, X. (2004). K-means Clustering via Principal Component Analysis. *Proceedings of the Twenty-First International Conference on Machine Learning (ICML)*.
- Tumminello, M., Lillo, F., & Mantegna, R. N. (2010). Correlation, Hierarchies, and Networks in Financial Markets. *Journal of Economic Behavior & Organization, 75*(1), 40–58.
- Onnela, J.-P., Chakraborti, A., Kaski, K., Kertész, J., & Kanto, A. (2003). Asset Trees and Asset Graphs in Financial Markets. *Physica Scripta*, T106, 48–54.
- Sullivan, R., Timmermann, A., & White, H. (1999). Data-Snooping, Technical Trading Rule Performance, and the Bootstrap. *The Journal of Finance*, 54(5), 1647–1691.
- Bailey, D. H., Borwein, J., López de Prado, M., & Zhu, Q. J. (2014). The Probability of Backtest Overfitting. *Journal of Computational Finance*, 20(4), 39–70.
- Lopez de Prado, M. (2018). Advances in Financial Machine Learning. *Wiley Trading*.