An Aproach to Using K-means Clustering to Uncover Latent Relationship Amongst Asset Performance

Gabriel Rambanapasi[a]

[a]*Stellenbosch University, Stellenbosch, South Africa*

Abstract

This study employs a K-Means clustering algorithm coupled with technical factors, to create "cluster portfolios" to uncover latent relationships in asset performance. Initial assessments show that these portfolios exhibit strong short-term performance with elevated risk levels. A rolling backtest reveals that Cluster 1, representing three sectors, displays strong positive performance measured by the Sharpe Ratio over short horizons. This integrated approach provides a valuable toolkit for portfolio managers navigating market shifts, offering insights into stock perfromance.

Keywords: K-Means, Clustering, Price Momentun, Volatility

1. Introduction

Consider the prevailing market consensus that anticipates a specific direction for market conditions. For instance, in the current interest rate environment, the majority of practitioners believe that in 2024, the Federal Reserve will initiate interest rate cuts. This could trigger a chain reaction in valuations across capital markets, affecting sectors and industries disproportionately. For investors, this underscores the importance of carefully scrutinizing different parts of the stock market to pinpoint companies that can enhance the risk/return profile of a portfolio. In this study, we employ a combination of K-Means clustering and technical factors to uncover latent relationships wih group of stock. Essentially, we reduce the dimensions of our sample data, to create "cluster portfolios" making it less challenging to uncover latent relationships in asset performance and group characteristics.

Our initial assessment involves the utilization of a simple look-back period to observe performance and risk over 3, 6, and 12 months. This analysis reveals that, on average, cluster portfolios exhibit strong performance over short investment horizons. However, they also display significant risk levels when compared to a market index. To further evaluate the robustness of the portfolios formed through our factor filter, we conduct a rolling backtest. The findings suggest that, despite the elevated levels of risk,

*Corresponding author: Gabriel Rambanapasi
Email address: rambanapasi44@gmail.com (Gabriel Rambanapasi)

risk-adjusted performance, as measured by the Sharpe Ratio (SR) in Cluster 1 that was represented by three sectors, returns strong positive performance over short time horizons. Thus buttressing, our anlaysis from the lookback analysis. However, as the investment horizon extends, the superior risk adjusted return flatuates and diminishes with low positive SR.

This integrated approach of clustering and factor analysis provides a valuable toolkit for portfolio managers seeking to identify potential out performers admidst anticipated market shifts. The focus on reducing the complexity of the data allows for a more nuanced understanding of asset performance, with the potential for enhanced risk-adjusted returns in the short term. Nevertheless, the strategy's limitations in the context of number of factors underscore the importance of continual refinement and adaptation to better suit evolving market conditions.

The next section 2 offers a brief discription of the various unsupervised machine learning algorithms. Section 3 details the methodology and data used along with its preparation procedure to be utilized by the K means clustering algorithm. Section 4 discusses the results and sectiom 5 offers the conclusion.

2. Clustering and Appliactions to Asset Management

Clustering is the process of grouping objects based on similar characteristics. The algorithms designed to cluster, achieve this function by connecting observation through distances, density of data points, graphs, or various statistical distributions. This is a form of unsupervised machine learning, that searches for patterns in datasets with no pre-existing labels and a minimum of human intervention. For a cluster to have meaning an algorithm has to maximize intra-cluster similarity and minimize inter-cluster similarity, such that each cluster contains information that's as dissimilar to other clusters(Kassambara, 2017). There exists various forms of cluster algorithms, each that addresses a broader task of analysis. The algorithms can be divided into two main types being partitioning clustering and hierarchical clustering. The major difference between the divisions of clustering is the partition clustering aims to specify a predetermined number of clusters whilst hierarchical clustering does not (Kassambara, 2017). Within partition clustering, for data with a small set of variables, K-means clustering and partitioning around medoids (PAM) are the most frequently used due to their fast compuation and simplicity. With K-means, each cluster is represented by the center or means of the data points belonging to the entire dataset. This makes the algorithm sensitive to outliers. However with PAM, each cluster is represented by one of the objects in the cluster. The other partition clustering algorithm used for datasets with a very large number of variables is Clustering Large Applications (CLARA).

In asset management, key to funds generating superior risk adjusted returns is efficient portfolio diversification, thus presenting a great application for partition clustering. Stocks would be separated into groups through a clustering algorithm to maximize similarity within groups and minimizes similarity between groups. Thus allowing managers to select handpick stock to construct a diversified portfolio. Marvin (2015) use fundamental ratios (turnover and profitability ratios) weighted equally and K-means clustering to group US technology stocks listed on the NASDAQ and NYSE. A diversified portfolio is then constructed based on within cluster stock performance i.e. stock selected are those that possess the highest Sharpe ratio. Results over a period of 15 years that included the dot com bubble and the global financial crises showed that cluster portfolios exhibited more volatility than the benchmark (S&P 500), however returns to investors were above the benchmark at multiples ranging from 3.5 to 5.7 times when earnings are reinvested into the cluster portfolios. Bin (2020) uses a similar approach to Marvin (2015), however employing a combination of market ratios and fundamental ratios (price to earnings ratio, return on assets ratio and asset turnover ratio). From this study, compared to the S&P 500, portfolios constructed using market ratios under performed those that used fundamental ratios.

3. Data and Methodology

This section describes how we obtain the data set used in the study, details the clustering process and validating metrics employed, to obtain the results in 4. For the methodology results, see Appendix 7.1 and 8.

3.1. Obtaining and prepariing the dataset

The data employed in this paper is based on the constituent list of the Johannesburg All Share Index (ALSI) from January 1, 2014 to January 29 2024. The historical price and volume data is retrieved from Yahoo Finance. From historical price data obtained from Yahoo Finance, we filter stock that have trading volume that exceed 1 000 000 shares traded per year and exclude stock that have less than 90 percent of observations in the historical price dataset.

To avoid large oscillations in the data, we transformed the price series to include end of month data points thus returns calculations are based on from the monthly data. Monthly historical prices are transformed using simple returns and we assume that embedded in the price action are cooperate events such as stock splits or consolidations of the shares. Therefore there is no need to make additional transformations on the return series to reflect corporate actions.

The measures of similarity used in this study are volatility and price momentum. To cluster stock based on the two measures, we apply a percentile ranking criterion on stock scores during a time period. For price momentum, describes the causality between relatively strong performance and high future return and vice-versa. Ranking highly implies that strong performers and thus higher returns than weak performers. This study defines cross sectional momentum as the trailing 12 month cumulative return Jegadeesh & Titman (1993), a modication to capture some longer term effects in the return series. For volatility, using a 12 month lock back period compute the standard deviation. The results of the ranking are shown in Table 7.1

3.2. Stocks clustering

3.2.1. Lloyd's algorithm

We employ the K-means that partitions n observations into k clusters. The goal is to minimize the within cluster sum of squares or analytically, $\arg\min_{s} \sum_{i=1}^{k} \sum_{x \in S_i} \|x - \mu_i\|^2$ where x are the observations, $S = S_1, S_2, \ldots, S_k$ are the sets of observations, and μ_i is the mean of the points in S_i. To arrive at the optimum number of clusters, we utilize the most popular algorithm called the Lloyd's algorithm that is closely followed by Marvin (2015), Bin (2020) & Xu, Xu & Wunsch (2010).

Analytically, given a set of points $\{x_1, \cdots x_n\}$ $(x_i \in \mathbb{R}^m)$, Initialize the K clusters with $\{C_1, \cdots C_K\}$ with centers. $\{m_1, m_2, \cdots m_K\}$ $(m_i \in \mathbb{R}^m)$. The centers are picked using the silhouette method discussed in 3.3 For all points $x_i (i \in \{1, \cdots n\})$, find the center that closest based on a euclidean distance d. Following this, assign x_i to the cluster corresponding to the closest centre. Then $x_i \in C_j$ if $d(x_i, m_j) \leq d(x_i, m_l) \quad (\forall l \in \{1, \cdots K\})(j \neq l)(\forall i \in \{1, \cdots n\})$. Recalculate the center for each cluster $C_l (l \in \{1, \cdots K\})$. The new cluster centers are the mean of the points in the same cluster.

$m_l = \frac{1}{|C_l|} \sum_{x_p \in C_l} x_p \quad (\forall l \in \{1, \cdots K\})$.

Repeat processes two and three until no cluster has any change in point assignment.

3.3. Silhoutte index

To evaluate the goodness of fit of partitioning using K means clustering the silhoutte index is used (Rousseeuw, 1987). Given n data points $\{x_1, \cdots x_n\}$, a partitioning result of K cluster $\{C_1, \cdots C_K\}$ and distance metric d, for each x_i in cluster C_l, define

$a(x_i) = \frac{1}{|C_l| - 1} \sum_{\forall x_j \in C_l, i \neq j} d(x_i, x_j)$

$a(x_i)$ is the mean dissimilarity between x_i to all other points within the same cluster.For each point x_i in cluster C_l, define

$b(x_i) = \min_{\forall p \in \{1, \cdots K\}, p \neq l} \frac{1}{|C_p|} \sum d(x_i, x_j)$

$b(x_i)$ is the minimum dissimilarity between x_i and all points in some C_p which does not contain x_i. For each point x_i in cluster C_l, their silhouette index is defined as

$$s(x_i) = \begin{cases} 1 - \frac{a(x_i)}{b(x_i)} & \text{if } a(x_i) < b(x_i) \\ 0 & \text{if } a(x_i) = b(x_i) \\ \frac{b(x_i)}{a(x_i)} - 1 & \text{if } a(x_i) > b(x_i) \end{cases}$$

where $s(x_i)$ ranges between $[-1, 1]$. For $s(x_i)$ that approaches 1, it means that $a(x_i)$ needs to be significantly smaller than $b(x_i)$, implying that within-cluster mean dissimilarity is much less than the smallest between-cluster mean dissimilarity, and thus the model does a good job clustering similar points together.

For the $s(x_i)$ that approaches 0, $a(x_i)$ needs to be significantly greater than $b(x_i)$, implying that within-cluster mean dissimilarity is much greater than the smallest between- cluster mean dissimilarity, and thus the model does a poor job clustering similar points together.

For this study, we choose K with the highest silhoutte index/value in Figure3.1 and gives the clusters in Figure 8.8

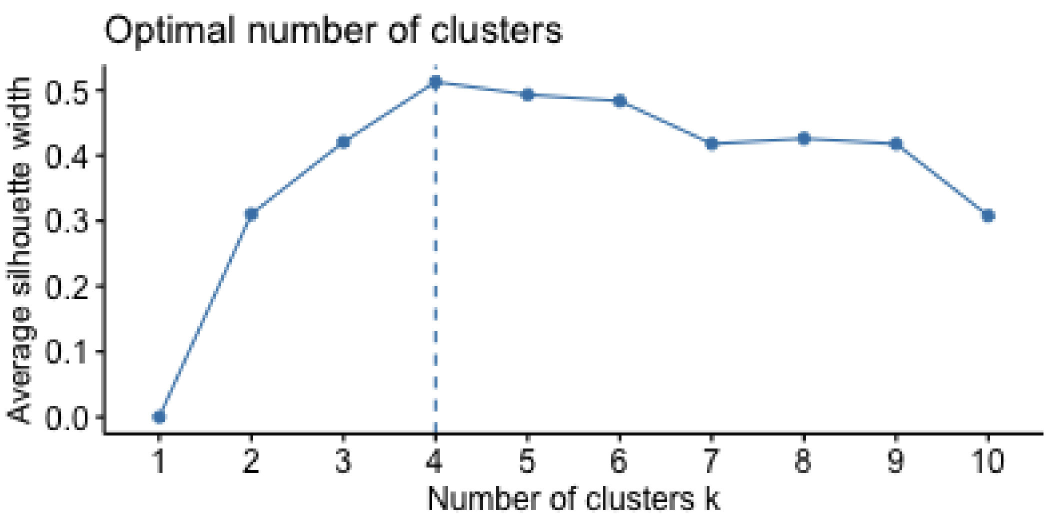

Figure 3.1: Silhoutte Indexes for Clusters

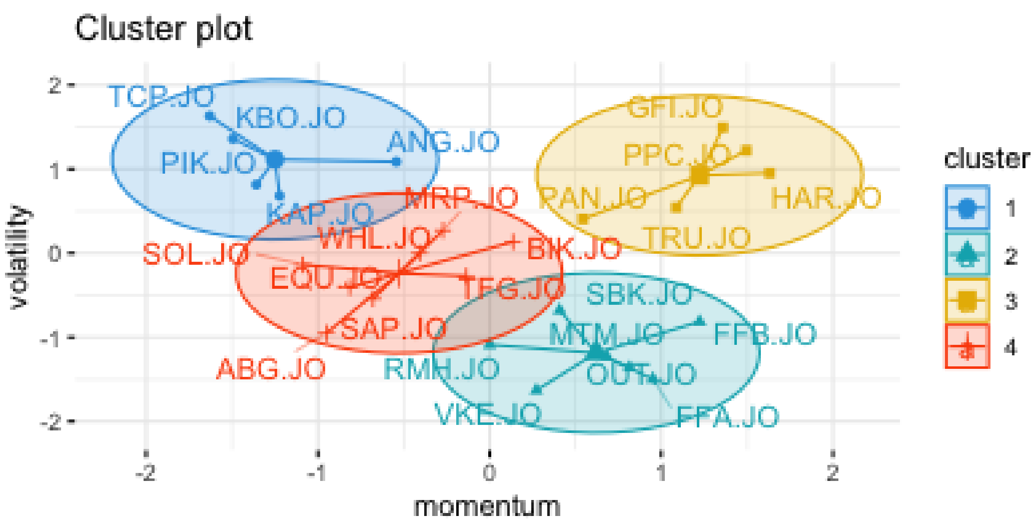

Figure 3.2: Clusters Results from Highest Silhoutte

3.4. Portfolio Backtest

The out-of-sample performance of cluster portfolios is compared to the benchmark, the JSE All Share Index. To manage risk exposure to a single asset or industry, we use a cap on each asset's allocation. Thus use a single company methodology similar to Standard & Poor Capping Methodology [1]. In a single company capping methodology, no company in an index (cluster in our case) is allowed to breach a certain pre-determined weight as of each rebalancing period [2]. Theoretically, this should preserve the within cluster diversification benefits and allow the portfolio value to either increase or decrease depending on relative stock performance during the quarter however capping limit asset risk expsoure. We re balance the portfolios once every three months, similar to the frequency of rebalancing conducted by the JSE on the JSE/FTSE indices, that is, re balance on the last daty of March, June, September and December.

4. Results

This section will present and analyze the clustering results of the K-means algorithms. We will discuss constituent and cluster charcateristics, hence show how we can uncover latent relationship within stocks that based on factors that otherwise would be challenging.

[1] see S&P (2023) for a breakdown of the process used to cap an index

[2] The maximum set for each stock in each cluster is its equal weight for that cluster i.e if there were 10 stock in cluster 1, the maximum weight would be 10%

4.1. Constiutent Description in Clusters

In table 7.1, our ranking criterion and filtering lead to the inclusion of three sectors: Basic Materials, Consumer Services, and Financials. The ranking criterion assigns the highest score to stocks that perform the best based on the specified factors. For instance, a score close to 1 for the Price Momentum Rank indicates that a stock has exhibited the strongest price movement over the past 12 months, while a lower score suggests weaker performance. Similarly, for the Volatility Rank, a score close to 1 signifies a stock with lower volatility, and a lower score indicates higher volatility.

Examining Table 7.1, we observe that most stocks in the Financial sector rank high in return volatility, Consumer Services stocks rank low, and Basic Materials stocks rank even lower. Price Momentum ranks Basic Materials stocks across the spectrum of Price Momnetum Rank, this is similar for Financials and Consumer Services. Consequently, from the ranking, there is no clear grouping by sector in terms of both price and volatility.

Moving on to Table 7.2, where we apply K-means clustering and determine the appropriate number of clusters to be K = 4, as suggested by Figure 3.1. Notably, Cluster 2 predominantly comprises stocks from the Financial sector, while Cluster 1, Cluster 3, and Cluster 4 exhibit a mix of representation from all sectors except Financials. Additional details, including plots of rolling 12-month returns and rolling 12-month standard deviation, can be found in Appendix 8 [3]. From tables 8.1, 8.2 and 8.3 we notice that risk adjusted returns for the clusters are highest on the shortest look back period of 3 months. Within that period SR is highest for Cluster at 1.37, emphasizing that investor recieve more reward for undertaking a certain level of risk.

4.2. Results from Rolling Backtest

Appendix 8 from 8.1 to 8.13 give a detailed results of our rolling backtest. We aim to evaluate the performance of these clusters over varying investment horizons, specifically focusing on relative risk and return perspectives in comparison to the JSE All Share Index. We add a modification to the methodology discussed in 3.4 designing a weight column constructed from relative volume weighting. This approach ensures a thorough consideration of liquidity, emphasizing the most liquid stocks within each cluster and have weights to calulate execute our capping function.

Upon analyzing the final results, a noteworthy trend emerges — the clusters consistently demonstrate positive excess returns relative to the benchmark. When holding the same clusters across investment horizons ranging from 1 to 10 years, it becomes apparent that excess returns are maximized in the

[3] These plots illustrate similar movement in returns and volatility, providing insights into stock behavior over a one-year rolling period.It is important to note that these findings do not make a definitive case for a direct relationship among the securities but rather highlight the observed behavior over a specified time frame.

shorter time frames. Specifically, over the short-term (1 year from table 8.2), clusters generate superior excess returns ranging from 30% to 57%, followed by 8.2% to 10.8% in the medium term (5 years from table 8.7, and 4.8% to 5.6% over the ultra-long term (10 years from table 8.13).

We notice consistency once we study the tracking error. Over the short term, tracking errors range from 1.8% to 7.9%, followed by 10.2% to 10.9% in the medium term, and 9.9% to 11.6% in the long term. This highlights the varying degrees of risk associated with the clusters across different investment horizons. However, this tracking error largely remains in tight ranges through the investment horizons, thus displaying ability for cluster portfolios to take consistent amount of risk relative to the benchmark.

Examining risk-adjusted metrics, we observe that clusters achieve a high Sharpe Ratio (SR) of 2.3 in the short term. As the investment horizon extends, the SR remains generally low but positive, suggesting that, while the clusters may exhibit lower risk-adjusted returns over longer periods, they continue to outperform in terms of risk efficiency in the shorter term.

Lastly, Cluster 1 displays the least drawdowns and exhibits quick recovery to previous from its maximum drawdown, particularly in very short investment horizons. On the other hand, Cluster 3 consistently experiences a more moderate level of drawdowns and recovery rates. Notably, Clusters 1, 2, and 4 undergo a doubling of drawdowns from year 6 to year 7. However, over the long term, all clusters experience a comparable level of drawdowns, with Cluster 4 showcasing the most protracted recovery period from its down levels.

5. Conclusion

Our objective was to develop a methodology utilizing factors with theoretical foundations to unveil latent relationships among stocks and/or sectors through the application of the K-means algorithm. Additionally, we delved into return and risk metrics from a retrospective standpoint, constructing a rolling backtest to assess the robustness of the clustering results.

The obtained results indicate that cluster portfolios not only provide positive excess returns but also demonstrate consistency in risk relative to the benchmark. Over short periods, these portfolios generate above-average risk-adjusted returns, albeit with the exception of the lone high value observed in Cluster 1 over a one-year investment holding period. A nuanced examination of the formed clusters reveals several important discoveries from our study. We found that as the investment horizon extends, maximum drawdowns for all clusters, in general, remained the same. However, at year 6 to 7, all but Cluster 3 had their maximum drawdowns increase. A deeper investigation could reveal valuable insight into risk drivers for the clusters.

Second, a single industry can make up an entire cluster, while a mix of industries can compose another.

These findings could be interpreted in various ways; for instance, clusters with sector concentration and large drawdowns could signal great sensitivity to systematic risk attached to that sector. On the other hand, diverse clusters could imply greater opportunities for sectoral rotation, depending on investor capital expectations.

We used price momentum and volatility factors, but other factors could be used to study group characteristics, thus showing the power of clustering and factor analyses. In conclusion, to enhance the comprehensiveness of this study, the inclusion of additional factors, particularly fundamental ones, in our K-means clustering framework could be considered. Furthermore, testing various partition algorithms for their ability to form clusters with high orthogonal risk drivers would contribute to a more thorough analysis.

6. References

Asness, C. 2011. Momentum in japan: The exception that proves the rule. *The Journal of Portfolio Management*. 37(4):67–75.

Bin, S. 2020. K-means stock clustering analysis based on historical price movements and financial ratios.

Jegadeesh, N. & Titman, S. 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*. 48(1):65–91.

Kassambara, A. 2017. *Practical guide to cluster analysis in r: Unsupervised machine learning*. Vol. 1. Sthda.

Marvin, K. 2015. Creating diversified portfolios using cluster analysis. *Princeton University*.

Rousseeuw, P.J. 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*. 20:53–65.

S&P. 2023. *Index mathematics methodology*.

Xu, R., Xu, J. & Wunsch, D.C. 2010. Clustering with differential evolution particle swarm optimization. In IEEE *IEEE congress on evolutionary computation*. 1–8.

7. Appendix

7.1. Appendix A

Table 7.1: Stock Ranking By Factors

	Ticker	Price Momentum Rank	Volatility Rank	Sector
1	ANG	11.76	100.00	Basic Materials
2	GFI	82.35	94.12	Basic Materials
3	PAN	5.88	52.94	Basic Materials
4	SAP	52.94	64.71	Basic Materials
5	MRP	100.00	17.65	Consumer Services
6	PIK	35.29	70.59	Consumer Services
7	TFG	94.12	88.24	Consumer Services
8	TRU	23.53	82.35	Consumer Services
9	WHL	70.59	76.47	Consumer Services
10	ABG	76.47	47.06	Financials
11	EQU	17.65	23.53	Financials
12	FFA	29.41	5.88	Financials
13	MTM	47.06	41.18	Financials
14	OUT	88.24	35.29	Financials
15	SBK	41.18	29.41	Financials
16	VKE	58.82	11.76	Financials
17	KAP	64.71	58.82	Industrials

Table 7.2: Stock Classification by Sector and Cluster

	Ticker	Cluster	Sector
1	ABG.JO	1	Financials
2	MRP.JO	1	Consumer Services
3	OUT.JO	1	Financials
4	EQU.JO	2	Financials
5	FFA.JO	2	Financials
6	MTM.JO	2	Financials
7	SBK.JO	2	Financials
8	VKE.JO	2	Financials
9	ANG.JO	3	Basic Materials
10	PAN.JO	3	Basic Materials
11	PIK.JO	3	Consumer Services
12	TRU.JO	3	Consumer Services
13	GFI.JO	4	Basic Materials
14	KAP.JO	4	Industrials
15	SAP.JO	4	Basic Materials
16	TFG.JO	4	Consumer Services
17	WHL.JO	4	Consumer Services

8. Appendix B

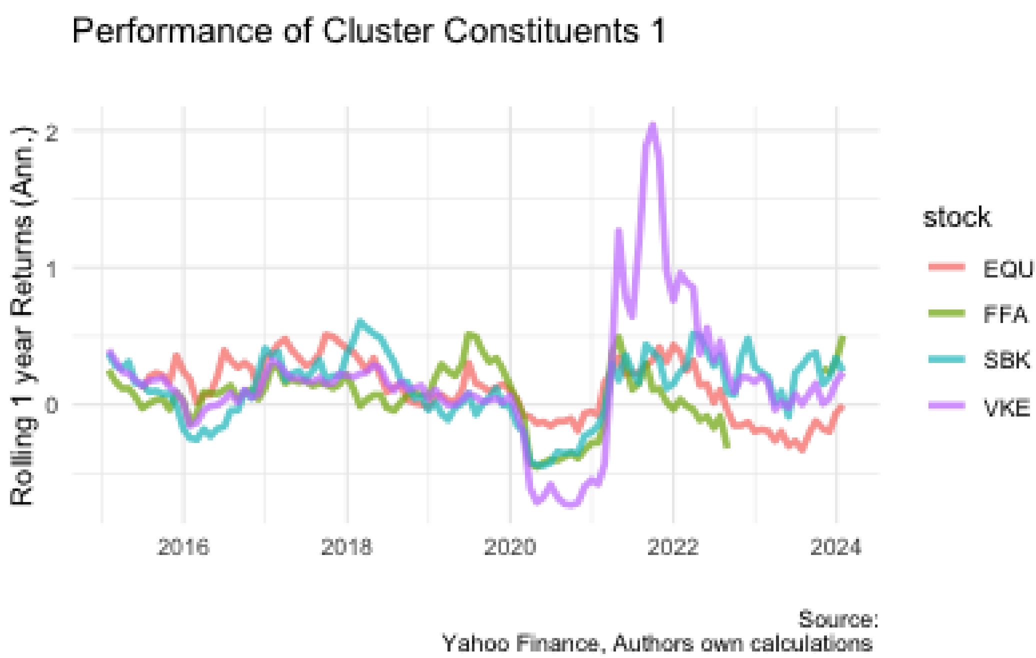

Figure 8.1: Clusters Results from Highest Silhoutte

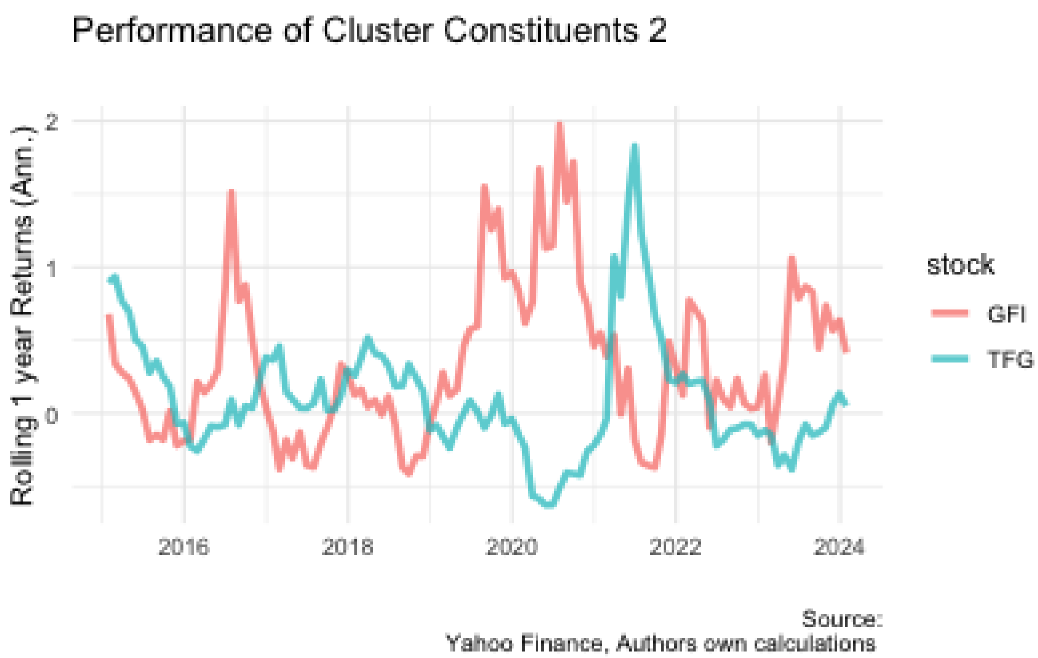

Figure 8.2: Clusters Results from Highest Silhoutte

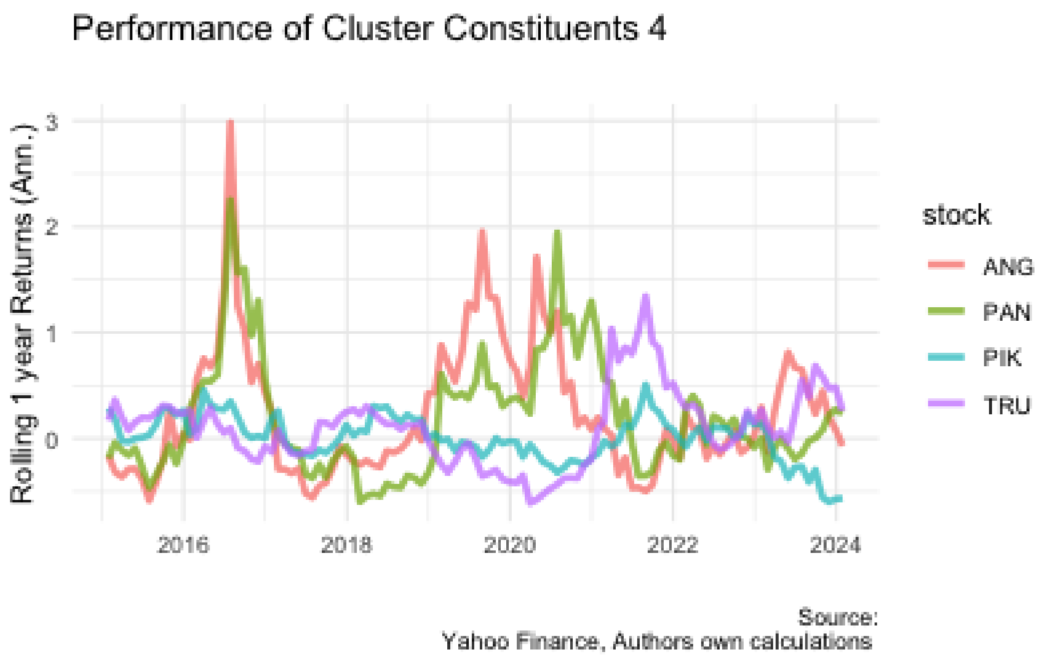

Figure 8.3: Clusters Results from Highest Silhoutte

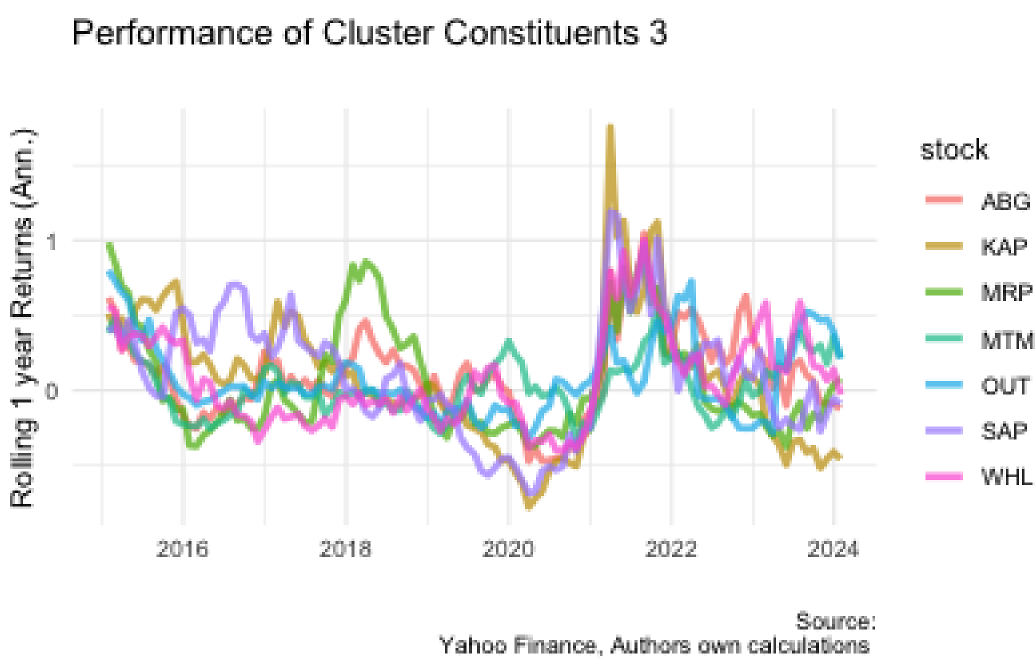

Figure 8.4: Clusters Results from Highest Silhoutte

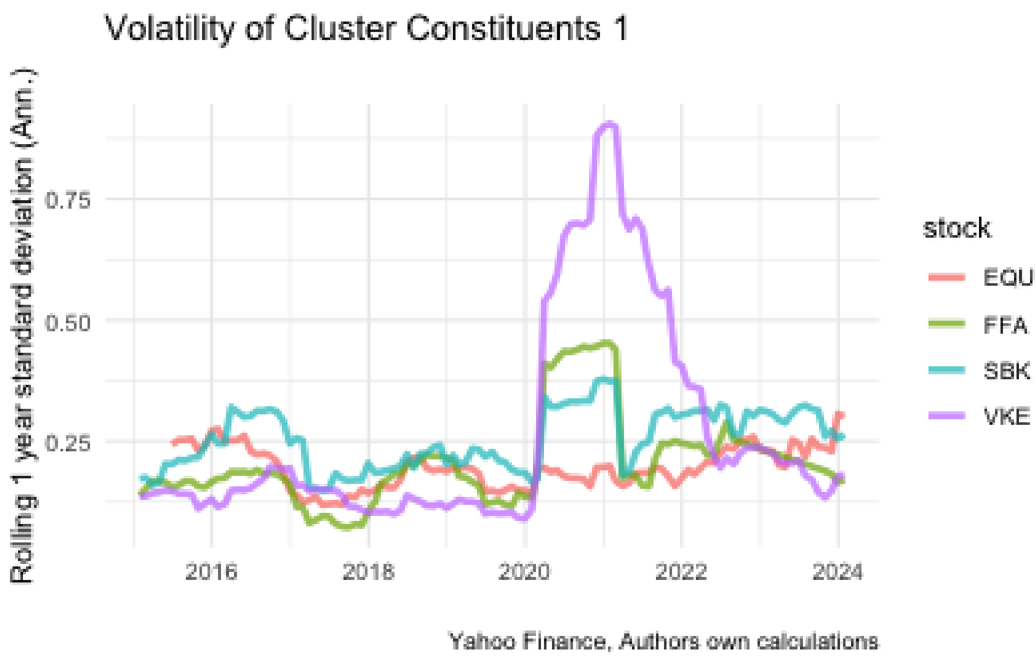

Figure 8.5: Clusters Results from Highest Silhoutte

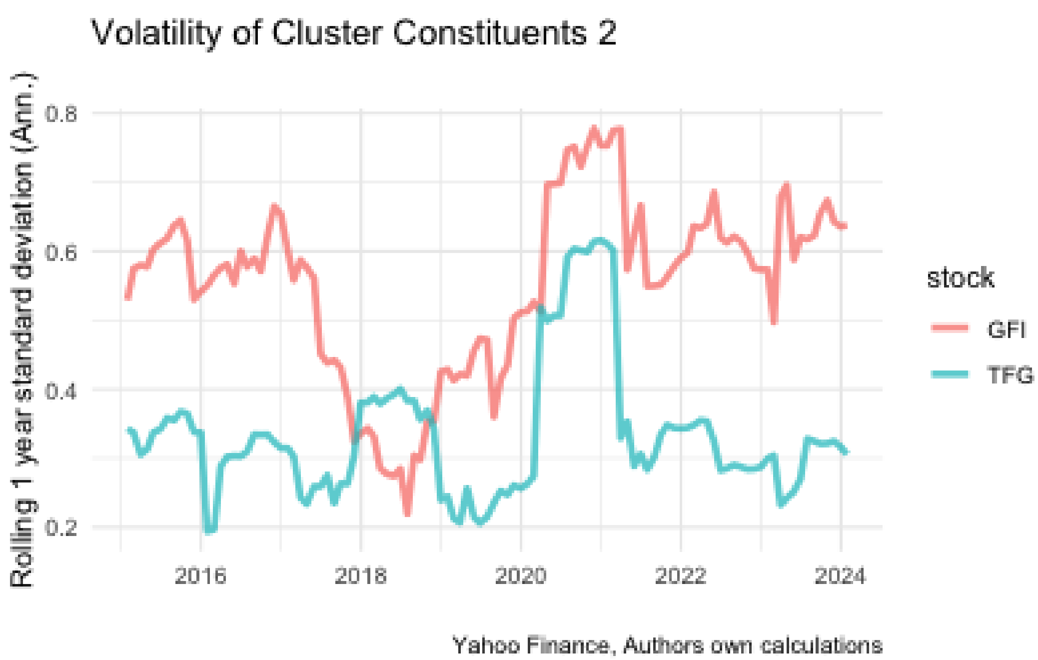

Figure 8.6: Clusters Results from Highest Silhoutte

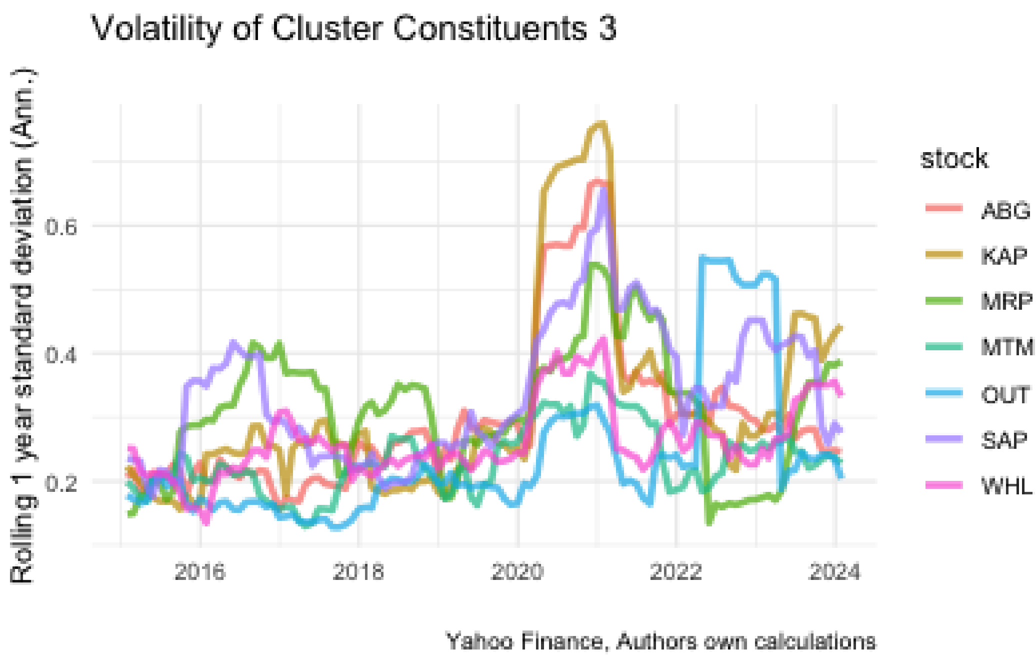

Figure 8.7: Clusters Results from Highest Silhoutte

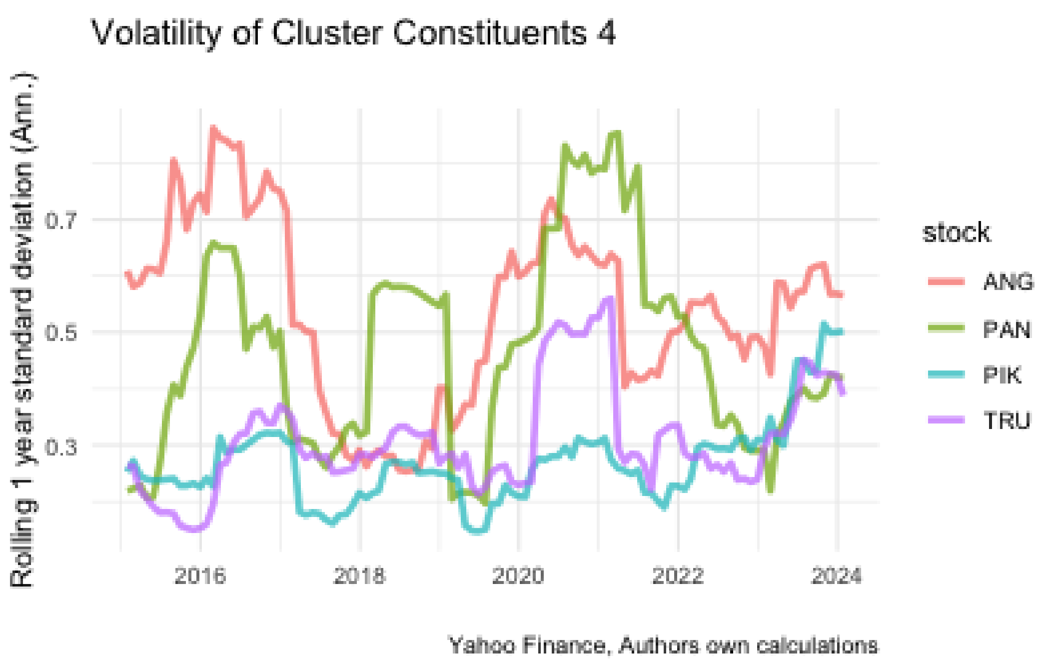

Figure 8.8: Clusters Results from Highest Silhoutte

8.1. Look Back

Table 8.1: LookBack Performance

	Period	Info	X1	X3	X4	X2
1	Last 12 Months	Adj. Sharpe Ratio	0.17	0.18	0.23	0.12
2	Last 12 Months	Avg DD	0.13	0.15	0.17	0.12
3	Last 12 Months	Beta	1.11	0.69	0.97	1.20
4	Last 12 Months	Beta Bear	1.97	0.01	1.30	2.05
5	Last 12 Months	Beta Bull	1.05	1.43	0.52	1.41
6	Last 12 Months	Returns Excess (Ann.)	0.06	0.05	0.06	0.05
7	Last 12 Months	Tracking Error	0.20	0.19	0.20	0.17

Table 8.2: LookBack Performance

	Period	Info	X1	X3	X4	X2
1	Last 3 Months	Adj. Sharpe Ratio	0.43	0.17	1.37	0.25
2	Last 3 Months	Avg DD	0.10	0.07	0.03	0.09
3	Last 3 Months	Beta	1.11	0.46	0.53	1.29
4	Last 3 Months	Beta Bear	1.19	0.55	1.20	1.87
5	Last 3 Months	Beta Bull	0.55	-0.15	-2.46	-0.24
6	Last 3 Months	Returns Excess (Ann.)	0.13	0.09	0.21	0.11
7	Last 3 Months	Tracking Error	0.16	0.17	0.17	0.13

Table 8.3: LookBack Performance

	Period	Info	X1	X3	X4	X2
1	Last 6 Months	Adj. Sharpe Ratio	-0.06	0.55	0.23	0.07
2	Last 6 Months	Avg DD	0.35	0.06	0.10	0.17
3	Last 6 Months	Beta	1.16	0.57	1.06	1.19
4	Last 6 Months	Beta Bear	2.20	0.37	1.56	2.38
5	Last 6 Months	Beta Bull	1.09	1.35	0.77	0.70
6	Last 6 Months	Returns Excess (Ann.)	0.03	0.10	0.08	0.05
7	Last 6 Months	Tracking Error	0.20	0.17	0.20	0.18

8.2. Rolling Backtest Results

Table 8.4: Rolling Performance Performance

	Investment Horizon	Info	Cluster_1	Cluster_3	Cluster_4	Cluster_2
1	1 Year	Ann Excess Return	0.58	0.31	0.46	0.36
2	1 Year	Ann Tracking Error	0.08	0.13	0.12	0.02
3	1 Year	Adj. Sharpe Ratio	2.33	-0.17	0.65	0.30
4	1 Year	DD Length	2.00	3.00	3.00	3.00
5	1 Year	Max DD	0.00	0.12	0.06	0.02

Table 8.5: Rolling Performance Performance

	Investment Horizon	Info	Cluster_1	Cluster_3	Cluster_4	Cluster_2
1	2 Year	Ann Excess Return	0.18	0.10	0.14	0.10
2	2 Year	Ann Tracking Error	0.10	0.12	0.10	0.08
3	2 Year	Adj. Sharpe Ratio	0.20	-0.38	-0.11	-0.63
4	2 Year	DD Length	3.00	7.00	7.00	4.00
5	2 Year	Max DD	0.12	0.24	0.16	0.14

Table 8.6: Rolling Performance Performance

	Investment Horizon	Info	Cluster_1	Cluster_3	Cluster_4	Cluster_2
1	3 Year	Ann Excess Return	0.17	0.13	0.14	0.15
2	3 Year	Ann Tracking Error	0.10	0.13	0.12	0.09
3	3 Year	Adj. Sharpe Ratio	0.59	0.18	0.25	0.35
4	3 Year	DD Length	4.00	6.00	11.00	4.00
5	3 Year	Max DD	0.12	0.24	0.16	0.14

Table 8.7: Rolling Performance Performance

	Investment Horizon	Info	Cluster_1	Cluster_3	Cluster_4	Cluster_2
1	4 Year	Ann Excess Return	0.16	0.05	0.11	0.11
2	4 Year	Ann Tracking Error	0.11	0.12	0.11	0.09
3	4 Year	Adj. Sharpe Ratio	0.69	-0.19	0.25	0.27
4	4 Year	DD Length	4.00	8.00	15.00	5.00
5	4 Year	Max DD	0.12	0.24	0.16	0.14

Table 8.8: Rolling Performance Performance

	Investment Horizon	Info	Cluster_1	Cluster_3	Cluster_4	Cluster_2
1	5 Year	Ann Excess Return	0.11	0.07	0.09	0.08
2	5 Year	Ann Tracking Error	0.11	0.13	0.10	0.10
3	5 Year	Adj. Sharpe Ratio	0.41	0.08	0.22	0.19
4	5 Year	DD Length	4.00	10.00	19.00	4.00
5	5 Year	Max DD	0.12	0.24	0.16	0.14

Table 8.9: Rolling Performance Performance

	Investment Horizon	Info	Cluster_1	Cluster_3	Cluster_4	Cluster_2
1	6 Year	Ann Excess Return	0.08	0.06	0.06	0.09
2	6 Year	Ann Tracking Error	0.11	0.12	0.11	0.10
3	6 Year	Adj. Sharpe Ratio	0.28	0.09	0.04	0.39
4	6 Year	DD Length	6.00	12.00	23.00	5.00
5	6 Year	Max DD	0.16	0.24	0.16	0.14

Table 8.10: Rolling Performance Performance

	Investment Horizon	Info	Cluster_1	Cluster_3	Cluster_4	Cluster_2
1	7 Year	Ann Excess Return	0.06	0.07	0.04	0.06
2	7 Year	Ann Tracking Error	0.12	0.12	0.11	0.11
3	7 Year	Adj. Sharpe Ratio	0.08	0.19	-0.03	0.09
4	7 Year	DD Length	6.00	13.00	27.00	5.00
5	7 Year	Max DD	0.35	0.24	0.30	0.26

Table 8.11: Rolling Performance Performance

	Investment Horizon	Info	Cluster_1	Cluster_3	Cluster_4	Cluster_2
1	8 Year	Ann Excess Return	0.06	0.04	0.04	0.06
2	8 Year	Ann Tracking Error	0.12	0.11	0.11	0.10
3	8 Year	Adj. Sharpe Ratio	0.16	0.05	0.03	0.15
4	8 Year	DD Length	8.00	10.00	31.00	6.00
5	8 Year	Max DD	0.35	0.24	0.30	0.26

Table 8.12: Rolling Performance Performance

	Investment Horizon	Info	Cluster_1	Cluster_3	Cluster_4	Cluster_2

1	9 Year	Ann Excess Return	0.05	0.04	0.05	0.04
2	9 Year	Ann Tracking Error	0.12	0.11	0.11	0.10
3	9 Year	Adj. Sharpe Ratio	0.12	0.05	0.10	0.04
4	9 Year	DD Length	8.00	12.00	35.00	5.00
5	9 Year	Max DD	0.35	0.24	0.30	0.26

Table 8.13: Rolling Performance Performance

	Investment Horizon	Info	Cluster_1	Cluster_3	Cluster_4	Cluster_2
1	10 Year	Ann Excess Return	0.06	0.05	0.06	0.05
2	10 Year	Ann Tracking Error	0.12	0.11	0.11	0.10
3	10 Year	Adj. Sharpe Ratio	0.17	0.18	0.23	0.12
4	10 Year	DD Length	10.00	10.00	18.00	6.00
5	10 Year	Max DD	0.35	0.24	0.30	0.26