

Appendix

Data Source:

Date for this research was obtained solely from 13F Filing Data provided by SEI Novus' [Public Ownership platform](#).

Every firm in the data set belongs to the Novus Hedge Fund Universe, a collection of over 1,500 Hedge Funds for which SEI Novus ingests quarterly 13F Filings. Using solely these filings SEI Novus is able to derive the metrics used in this research.

Metrics:

The metrics used for analysis are made up of pure output (e.g., sharpe ratio or up-capture), pure input (e.g., beg position count or average exposure to sectors), and a mix of the two (e.g., sector contribution). The combination provides a holistic look at the drivers of performance.

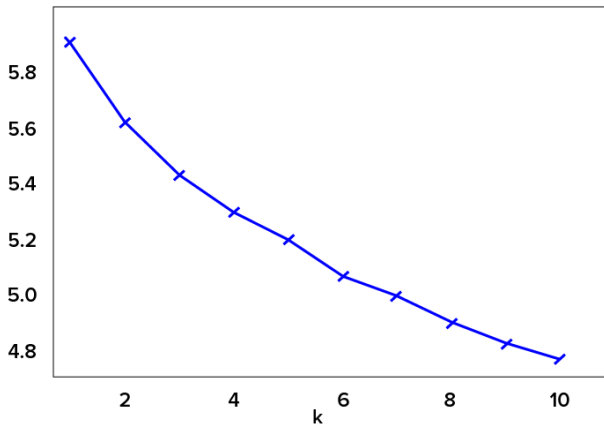
For definition of metrics below used please refer to the Novus University Glossary (available to SEI Novus Clients).

<u>Metrics Used</u>	<u>Exposure to Following Sectors:</u>	<u>Average Exposure to Following Geographies:</u>
Sortino Ratio	Communication Services	
Beg Position Count	Consumer Discretionary	United States
Market Cap Weighted Average	Financials	Canada
Ending Position Count	Industrials	Developed Europe
Crowdedness Weighted Average	Information Technology	Emerging Markets
S&P 500 TR	Consumer Staples	Asia-Pacific
Beg AUM	Energy	Emerging Asia
Up Capture - S&P 500	Health Care	Japan
Sharpe Ratio	Materials	
EVol Annualized Last	Real Estate	
Market Cap Median	Utilities	
Crowdedness Median		
EVol Annualized Avg		
Down Capture - S&P 500		
Performance		
Contribution	Mega Cap (>50B)	
Market Contribution	Large Cap (10B-50B)	
Sector Contribution	Mid Cap (2B-10B)	
Security Contribution	Small Cap (250MM-2B)	
Trading Contribution	Micro Cap (<250MM)	
Win Loss Ratio		
Batting Average		

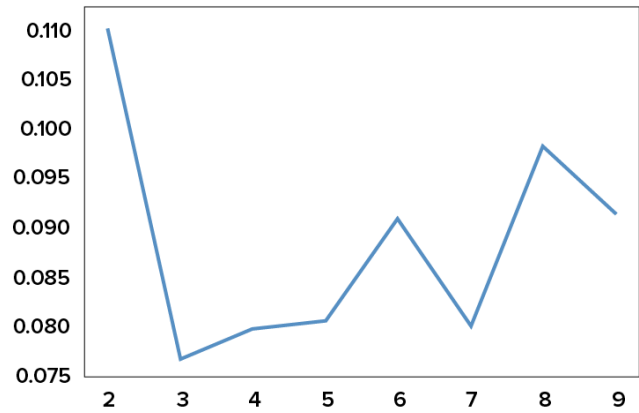
Group Creation:

To create the various groups, we employed ‘K-means clustering.’ K-means clustering is a machine learning method which aims to segment n observations into k clusters in which each observation belongs to the cluster with the nearest mean (cluster centers or cluster centroid), serving as a prototype of the cluster. This results in a partitioning of the data space in such a way that the within-cluster ‘Average Distortion’ is minimized. The Average Distortion is simply the average distance between an observation and its group’s center. Additionally, we employ the ‘Silhouette Score’ which is a metric that measures how similar an observation is to its group by measure its pair wise distance to other observations and their centers.

Average Distortion vs. Number of Groups



Silhouette Score vs. Number of Groups



Eight is the number of groups at which point the Average Distortion no longer declines significantly. Additionally, at eight groups the ‘Silhouette Score’ peaks, suggesting the tightest grouping with the lowest within-cluster ‘Average Distortion’. That is why we constrained the analysis to the first eight groups.