Portfolio Analysis of Famous Investors

Team 13:

Khan Taskin Shireesha Allani Khanh Nguyen

Table of Contents

1. In	troduction	3
2. In	itial Hypothesis	3
3. Li	terature Review	3
4. Da	ataset and Features	4
4.1.	Data collection	4
4.2.	Data Cleaning	6
4.3.	Exploratory Data Analysis	6
5. M	ethodology	7
5.1.	Factor Regression	7
5.2.	K-means Clustering	8
6. Re	esults and Findings	9
6.1.	Investor Groups	9
6.2.	Analysis of Cumulative Returns	11
7. Co	onclusion	Error! Bookmark not defined.

1. Introduction

Do famous investors have similar investment styles? What stocks do they mostly own? How do their returns compare against each other, and which investors had the best return? In this report we will try to answer some of these questions by analyzing the portfolios of 8 famous investors using Fama-French Factor Analysis. The investor portfolios we will be analyzing are:

- Warren Buffett \$299 billion AUM (Assets Under Management)
- Bill Gates \$35.7 billion AUM
- Carl Icahn \$21.7 billion AUM
- Bill Miller \$1.5 billion AUM
- Mohnish Pabrai \$120 million AUM
- Howards Marks \$7.5 billion AUM
- Guy Spier \$186 million AUM
- Seth Klarman \$6.1 billion AUM

By analyzing the portfolios of these investors, we can identify the traits/strategies that have led to their success and apply those to our own stock portfolios. We will be able to determine the qualities of a stock that these famous investors target and use the same methodology to identify new potential investments for our own portfolio.

We used quarterly portfolio snapshots of investors to identify the stock holdings of the investors from 2014 to 2022. Using that we extracted stock price information for top holdings for each investor to determine portfolio return for the investor over time. In some cases, we utilized data imputation methods to fill in missing data, for instance, stock price not available before a certain date, or ticker information missing. Once we had the historical returns for the investor portfolios, we performed factor regression using factor dataset and monthly returns for the investor and market. Using the results of factor regression, we were able to determine the qualities of the investors' portfolio.

We then categorized investors with similar qualities/style of investment and determined which group of investors had the largest average return over the period. We also looked at which stocks resulted in the biggest return for the investors and what were the qualities associated with those stocks.

2. Initial Hypothesis

Our initial hypothesis is that all the 8 investors have the same investment style and have similar portfolio returns over the time period.

3. Literature Review

In Christian Koch's article "The Warren Buffet Project: A Qualitative Study on Warren Buffet", he observes that Warren Buffett rejects market efficiency theory and advocates for value

investing. Additionally, Buffett advises us to rethink our approach to risk in investing, as it is the permanent loss of capital instead of volatility (Koch, 2022). This study from Koch suggests that Warren Buffet is a conservative investor, and we hypothesize that his portfolio is tilted towards value stocks. In addition to Warren Buffett, we also reviewed literature discussing Howard Marks' strategy as an investor. In Money Blog ET's article "Top 10 Investing Lessons from Howard Marks," the author states that Howard Marks has an unconventional approach when it comes to picking stocks. In particular, he encourages investors to seek out unpopular, and even controversial stocks that are undervalued ("Top 10 Investing Lessons from Howard Marks," n.d.). As we can see, it takes an aggressive investor who is willing to take on risks to embrace Howard Mark's approach. Therefore, we hypothesize that Howard Mark is an aggressive investor, and his portfolio will consist of risky stocks from smaller companies.

4. Dataset and Features

4.1. Data collection

For this project we have 3 major sources of data:

 <u>Dataroma</u> – Portfolio information of super investors: Stocks/Tickers in portfolio, portfolio allocation. The website lists the Top 20 allocation of each investor by FY Quarter up to FY 2022 Q4.

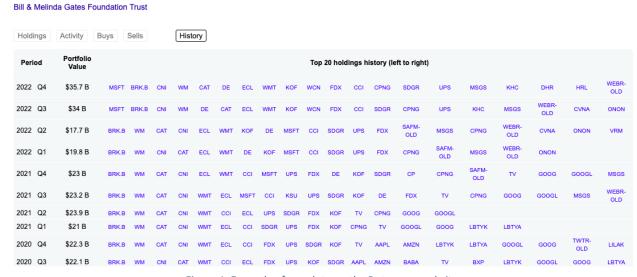


Figure 1. Example of raw data on the Dataroma website

The stocks and allocation information is extracted from Dataroma website using Excel's PowerQuery feature. The generated table was then read into a R dataframe.

	Portfolio																				
Period	Value		Top 20 holdings history (left to right)																		
																					WEBR-
		MSFT	BRK.B	CNI	WM	CAT	<u>DE</u>	ECL	WMT	KOF	WCN	FDX	<u>CCI</u>	CPNG	SDGR	<u>UPS</u>	MSGS	KHC	DHR	HRL	<u>OLD</u>
													Crown				Madison				
			Berkshir	Canadia	Waste					Coca-			Castle				Square				
			e	n Nati	Manage					Cola	Waste		Internati		SCHRO	United	Garden	Kraft		Hormel	
		Microsof	Hathawa	Railway	ment	Caterpill	Deere &	Ecolab	Walmart	FEMSA	Connect	FedEx	onal	Coupan	DINGER	Parcel	Sports	Heinz	Danaher	Foods	Weber
		t Corp.	y CL B	Co.	Inc.	ar Inc.	Co.	Inc.	Inc.	SACV	ions	Corp.	Corp.	g Inc.	Inc.	Service	Corp.	Co.	Corp.	Corp.	Inc.
		26.36%	21.35%	18.24%	15.47%																
				of			4.70% of														
2022 Q4	\$35.7 B	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio
																			WEBR-		
		MSFT	BRK.B	CNI	WM	<u>DE</u>	CAT	ECL	WMT	KOF	<u>WCN</u>	FDX	<u>CCI</u>	SDGR	CPNG	<u>UPS</u>	KHC	MSGS	OLD	CVNA	ONON
													Crown					Madison			
			Berkshir	Canadia	Waste					Coca-			Castle					Square			
			e	n Natl	Manage					Cola	Waste		Internati	SCHRO		United	Kraft	Garden			On
		Microsof		Railway			Caterpill	Ecolab	Walmart		Connect	FedEx	onal	DINGER	Coupan	Parcel	Heinz	Sports	Weber	Carvana	
		t Corp.	y CL B	Co.	Inc.	Co.	ar Inc.	Inc.	Inc.	SACV	ions	Corp.	Corp.	Inc.	g Inc.	Service	Co.	Corp.	Inc.	Co.	AG
		26.91%	23.32%	17.42%	16.61%																
		of	of	of			3.55% of														
2022 Q3	\$34 B	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio	portfolio

Figure 2. Extracted stock portfolios from Dataroma

This dataset allows us to capture changes in an investor's portfolio (buying new position or selling existing position) by each quarter and obtain a better a more accurate picture of the investor's portfolio returns for the time period. However, because the data is available by quarter, we will make the following assumptions for our analysis:

- o Top 20 holdings of the investor is used to determine portfolio and allocation
- Portfolio and weight allocation remains same for a given financial quarter for each investor. Buy/Sell activities of the investor are captured at a Quarter level and monthly stock prices and returns are calculated for the stocks the investor held during that quarter to capture investment returns.
- Any holding with no stock price information will be removed from the portfolio for that quarter. The weight allocation for the remaining stocks in the portfolio in that quarter are recalibrated so that the sum of weights equal to 1.
- <u>Kenneth French (Dartmouth)</u> Historical factor regression dataset up to December 2022. The data is downloaded as CSV and imported into a R dataframe. The factors we will be using for our analysis are:
 - o SMB (Small Minus Big) is the average return on the nine small stock portfolios minus the average return on the nine big stock portfolios.
 - HML (High Minus Low) is the average return on the two value portfolios minus the average return on the two growth portfolios.
 - QMJ (Quality Minus Junk) is the average return on the two robust operating profitability portfolios minus the average return on the two weak operating profitability portfolios.
 - CMA (Conservative Minus Aggressive) is the average return on the two
 conservative investment portfolios minus the average return on the two
 aggressive investment portfolios.

- o MOM (Momentum) is the average return on the two high prior return portfolios minus the average return on the two low prior return portfolios.
- <u>Tiingo</u> Monthly Stock price data. The data is extracted using an API key and using the tidyquant package.

4.2. Data Cleaning

The portfolio data from Dataroma is cleansed for formatting and prepared for data extraction. The Merged sales are unmerged, web tags, Header and Footer have been removed. A new Column called "Value" (Ticker, Name, Percentage) has been created and the column labels (X1 ..X20) have been added

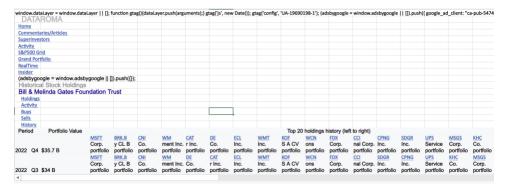


Figure 3. Extracted portfolio data

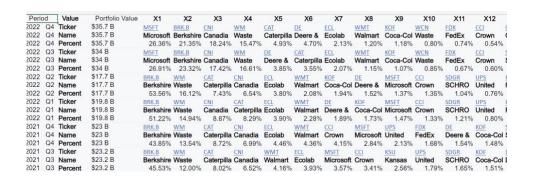


Figure 4. Cleansed portfolio data

4.3. Exploratory Data Analysis

The cleansed portfolio data is read into RStudio, and stock returns are calculated using tiingo package. The data has been analyzed and prepared to meet the tiingo dataset requirements. There were 2 major observations, One - data needs to be extracted using date range and the since the

'Period' from data source is a string, Year and Quarter has been extracted from 'Period'; first and last day for each quarter has been calculated for the given quarter and year.

```
> str(Input_File)
'data.frame':
                 84 obs. of 25 variables:
                           "2016
                                    01" "2016
                                                 01" "2016
                                                               01" "2016
                                                                             02"
$ Period
                   : chr
                           "Ticker"
                                     "Name" "Percent" "Ticker"
                    : chr
                           "$2.01 B" "$2.01 B" "$2.01 B" "$1.54 B"
$ Portfolio.Value: chr
                           "DAL" "Delta Air Lines Inc." "5.7200000000000001
"XON-OLD" "Intrexon Corp." "5.6399999999999999
                     chr
                     chr
                           "UAL" "United Airlines Holdings Inc." "4.8899999"
"AAL" "American Airlines Group Inc." "4.82E-2"
$ $ $ $ $ $ $ $
  Х3
                     chr
  X4
                    : chr
                           "LEN" "Lennar Corp." "4.6100000000000002E-2" "DA
  X5
                    : chr
                           "PHM" "PulteGroup Inc." "4.279999999999998E-2"
  X6
                     chr
                            "NXPI" "NXP Semiconductors NV" "3.7499999999999
  X7
                     chr
                            "OMF" "OneMain Holdings" "3.599999999999997E-2
  X8
                     chr
                            "QUOT" "Quotient Technology Inc." "3.5400000000
                     chr
                           "PAH" "Platform Specialty Products Corp." "3.420
"MTG" "MGIC Investment" "3.379999999999997E-2"
  X10
                     chr
  X11
                    : chr
                           "EIGI" "Endurance Int. Group Hldgs" "0.03" "BAC'
  X12
                    : chr
                           "BAC" "Bank of America Corp." "2.86E-2" "EIGI"
  X13
                    : chr
                           X14
                    : chr
  X15
                     chr
  X16
                     chr
                           "GILD" "Gilead Sciences" "2.3900000000000001E-2
  X17
                    : chr
                           "GME" "GameStop Corp." "2.29E-2" "GNW" ...
"C" "Citigroup Inc." "2.2800000000000001E-2" "GI
  X18
                    : chr
  X19
                     chr
                           "JPM.WS" "JPMorgan Chase WTS" "2.24999999999999
                     chr
  X20
                           $
  year
                      int
  Qtr_number
                           1 1 1 2 2 2 3 3 3 4 ...
                     int
```

Another observation is about stock data availability, some stock tickers are not supported in tiingo package, and for stock tickers that are supported but the stock data is not available in the date range. To handle missing data the tickers and corresponding weights are removed from the portfolio and the remaining ticker weights are adjusted to add up to 1.

5. Methodology

5.1. Factor Regression

To perform factor regression, we built a linear regression model for each investor with the portfolio's excess return above the risk-free rate as the response variable, and the factors as the six predictor variables, which are market beta, momentum, risk, value, quality, and size.

			Dependen	t variable:							
_	warrenbuffett_rf										
	(1)	(2)	(3)	(4)	(5)	(6)					
SMB		-0.153**	-0.245***	-0.165**	-0.163**	-0.180**					
		(0.076)	(0.075)	(0.078)	(0.079)	(0.078)					
HML			0.258***	0.238***	0.222***	0.155**					
			(0.055)	(0.054)	(0.070)	(0.074)					
QMJ				0.290***	0.288***	0.281***					
				(0.096)	(0.096)	(0.095)					
BAB					0.040	0.097					
					(0.114)	(0.115)					
MOM						-0.107**					
						(0.043)					
MKT_RF	0.949***	0.980***	0.966***	0.970***	0.974***	0.948***					
	(0.039)	(0.041)	(0.039)	(0.039)	(0.040)	(0.041)					
Constant	0.247	0.231	0.275	0.161	0.152	0.173					
	(0.185)	(0.183)	(0.174)	(0.175)	(0.177)	(0.175)					
Observations	195	195	195	195	195	195					
R^2	0.757	0.762	0.787	0.797	0.797	0.803					
Adjusted R ²	0.756	0.760	0.784	0.792	0.792	0.797					
Residual Std. Error	2.546 (df = 193)	2.526 (df = 192)	2.398 (df = 191)	2.348 (df = 190)	2.353 (df = 189)	2.321 (df = 188)					
F Statistic 6	02.000*** (df = 1; 193)	307.700*** (df = 2; 192)	235.000*** (df = 3; 191)	186.200*** (df = 4; 190)	148.300*** (df = 5; 189)	128.000*** (df = 6;					
Note:					*n/	:0.1; **p<0.05; ***p<					

Figure 5. Warren Buffett's factor regression results

To demonstrate our analysis from the factor regression results, we use Warren Buffet's outputs as an example. In Figure 5, we can see that the adjusted R-squared improves as we add more factors into the factor regression model. The intercept (constant) is positive in all six models indicating that Warren Buffett outperforms when taking each factor into consideration.

5.2. K-means Clustering

We extract the coefficients of each factor from each investor's factor regression results to use as features for the k-means clustering model. Hyperparameter tuning is performed using the elbow method. The plot shows k = 4 as the optimal number of clusters, as it yields the most marginal decrease in Total Within Sum of Squares.

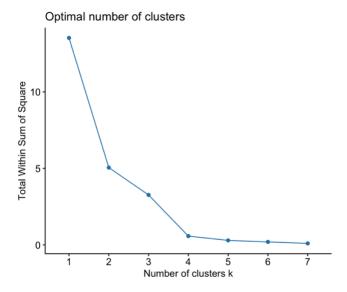


Figure 6. Hyperparameter tuning for K-means clustering model

6. Results and Findings

6.1. Investor Groups

Characteristics by Group

Ollui	deterioties by Gree	4P						
GROUP	INVESTORS	SIZE	VALUE	QUALITY	AGGRESSIVENESS	MOMENTUM		
1	Bill Gates, Warren Buffett	Small Cap	Value	Non-profitable	Conservative	Low momentum		
2	Bill Miller, Howard Marks	Small Cap	Growth	Non-profitable	Aggressive	Low momentum		
3	Carl Icahn, Seth Klarman	Large Cap	Value	Profitable	Aggressive	Low momentum		
4	Guy Spier, Mohnish Pabrai	Small Cap	Value	Non-Profitable	Aggressive	Low momentum		

Figure 7. Characteristics of investors by group

Based on what we have learned about the investment style of each investor in the literature review, the k-means clustering model has performed well in clustering the investors based on the factors. For instance, as mentioned in Section 3, Warren Buffet is an advocate for value investing, and our results from factor regression and clustering have confirmed that. Another example is Howard Marks. In Section 3, we have learned that he is an aggressive investor as he is comfortable taking on risks by investing in under-appreciated stocks and in sectors that are troubling.

Some key findings from Figure 7:

- All groups are tilted towards low momentum stocks.
- Group 1 with Bill Gates and Warren Buffett is the only group that is tilted towards volatile stocks.

- Group 2 with Bill Miller and Howard Marks is the only group that is tilted towards growth stocks.
- Group 3 with Carl Icahn and Seth Klarman is the only group that is tilted towards large cap and non-profitable stocks.

To gain a deeper insight into the characteristics of the four groups, we have computed the mean value of each factor for each group and compare them in Figure 8 and Figure 9. These are the key findings that we draw from the charts:

- Group 1 is the lowest performer. Out of the four groups, Group 1 is most tilted towards value, non-profitable, and low momentum stocks.
- Group 2 is the highest performer. Out of the four groups, Group 2 is the most aggressive, and is the most tilted towards small cap stocks.
- Group 3 and Group 4's investment styles are quite different, but in terms of performance, Group 3 only performs better than Group 4 to a small extent.
- Group 1 and Group 4 have similar characteristics except for their aggressive level, yet Group 4 performs twice as better compared to Group 1.
- Despite the difference in the returns of all groups, they have all outperformed.

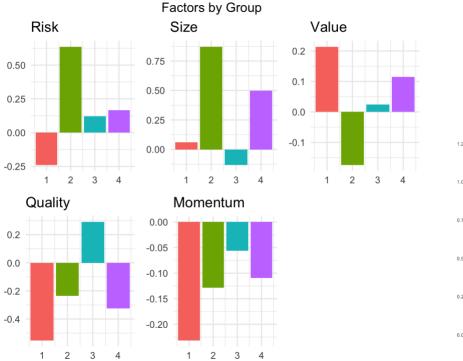


Figure 8. The degree of factors for each group

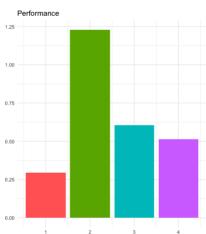


Figure 9. Performance of the four groups

6.2. Analysis of Cumulative Returns

Comparing the Cumulative returns of the 8 investors against the market returns in Figure 10, all of them beat the market in the period from 2014 to 2022. Howard Marks and Bill Miller have the best return during this period and had significantly high abnormal returns. Following them were Bill Gates and Warren Buffet and their cumulative returns closely followed each other's.

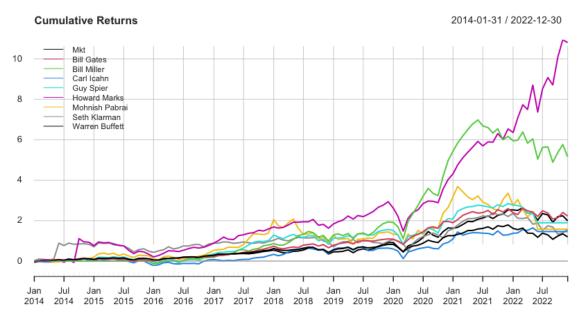
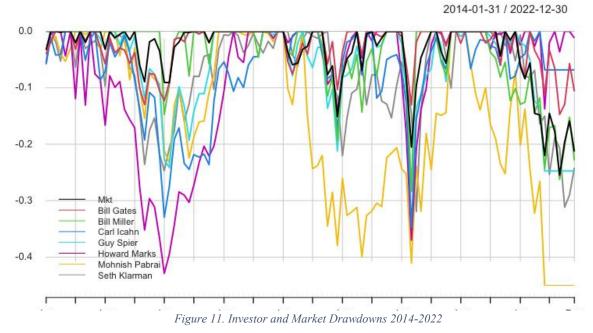


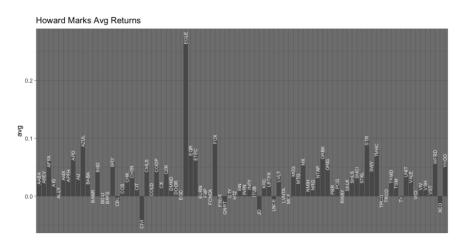
Figure 10. Cumulative Returns of Investors vs Market

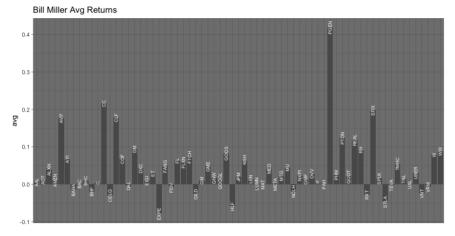
Looking at the drawdown chart for each of the investors and the market in Figure 11, Bill Miller and Howard Mark's portfolio were more sensitive to market declines in the earlier phase (2014-

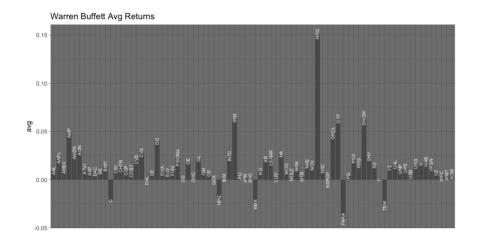


2018) with portfolio returns dropping significantly compared to the market. However, during the most recent market drawdown, Howard Mark's portfolio seems to be more resilient, and the declines were much less compared to the market. Bill Millers drawdowns were similar to that of the market in the latter half of the analysis period.

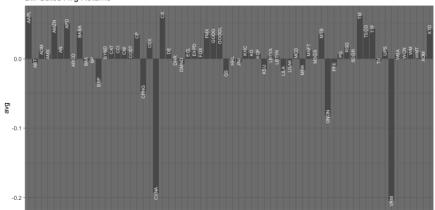
We also analyzed the average monthly returns of individual stocks that comprised the top 20 holdings of each of the investors. Since these stocks were potentially held by the investors over a long period of time, they will need to be investigated further in terms of their current profitability, future investment/revenue projection etc. to determine if they still meet our investment criteria to be included in our portfolio.



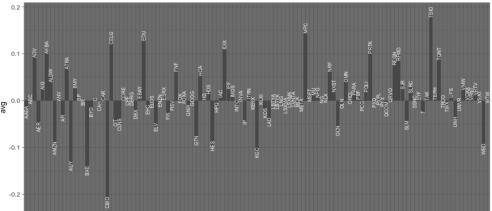




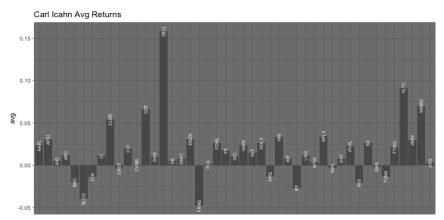
Bill Gates Avg Returns

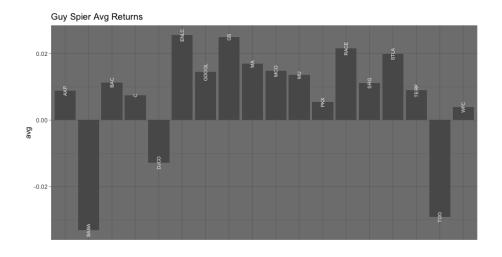


Seth Klarman Avg Returns









7. Conclusion

Based on our analysis of the 8 investor portfolios, we recommend investing in small cap stocks with high growth but have a conservative approach to investing in new projects. With the recent rise in interest rates, these companies will be able to preserve cash and only invest in highly profitable ventures, generating greater value for their shareholders. In addition, during times of economic growth these stocks are more likely to increase in value at a greater rate compared to large cap stocks resulting in high abnormal returns.

Reference

- [1] French, K. (2023). *Description of Fama/French 5 Factors (2x3)*. Kenneth R. French. Retrieved April 15, 2023, from https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library/f-f5 factors 2x3.html
- [2] DATAROMA. (2023). Retrieved April 15, 2023, from https://www.dataroma.com/m/home.php
- [3] Koch, Christian G., *The Warren Buffett Project: A Qualitative Study on Warren Buffett* (2020). USF Tampa Graduate Theses and Dissertations. https://digitalcommons.usf.edu/etd/8457
- [4] 10 Must Know Investing Lessons from Howard Marks. ET Money Blog. Retrieved April 15, 2023, from https://www.etmoney.com/blog/top-investing-lessons-from-howard-marks/#:~:text=Howard%20Marks%20considers%20risk%20as,result%20in%20significantly%20lower%20returns