[PDP] Final assignment Group-04
Professors: Guillermo Vigueras / Belén Ríos

Data-driven Portfolio Management
Part 2: Investment Strategies Analysis

Academic year 2022/2023

Nicolás Amigo Sañudo
Gema Díaz Ferreiro
Rafael Sojo García
Stephan Wolters Eisenhardt

Contents

1. Introduction

Smallville Asset Management have obtained all 126 portfolio options with the corresponding return and risk (volatility) ratios from the data generation of the preceding tasks.

A series of important factors have been considered related to the calculations:

a) The management costs of the invested capital have been included in the return calculation for both stocks, corporate bonds, public bonds, and gold.
b) For the cash asset, the reduction of the purchasing power from 1st of January 2020 (buy date) to 31st of December 2020 has been considered.
c) Some of the data visualizations have a normalized data input (min-max scaling as there a no extreme outliers) for return and volatility to account for scale invariance of the two features, easier interpretation, and comparison.

The provided formulas have been slightly adapted to calculate the correct return rate. As for volatility, the standard deviation over the mean has been maintained as suggested in the assignment.

Although many different pricing and risk models can be applied to the different assent types, we adhere to the typical return-risk tradeoff from a backward-looking perspective based in historical data from 2020.

The following sections provide insights into the two topics Smallville Asset Management is most concerned about: the probability of a positive or negative return and the relationship between risk and return.

2. Return

The first query refers to whether it is more probable to obtain a positive or negative return. We assume that the question pertains to the expected return over all 126 possible portfolios.

As can be observed in *figure 1: Average Positive and Negative Returns*, out of 126 portfolios 104 yield a positive return with an average of 7.00%, meanwhile there are only 22 portfolios yielding a negative return with an average of -2.30%, the invested capital being constant across all portfolios.

Therefore, the expected return across all 126 portfolios is 5.38%, which already answers the first question as the expected value is clearly positive and hence, it is more probable to achieve a positive return.

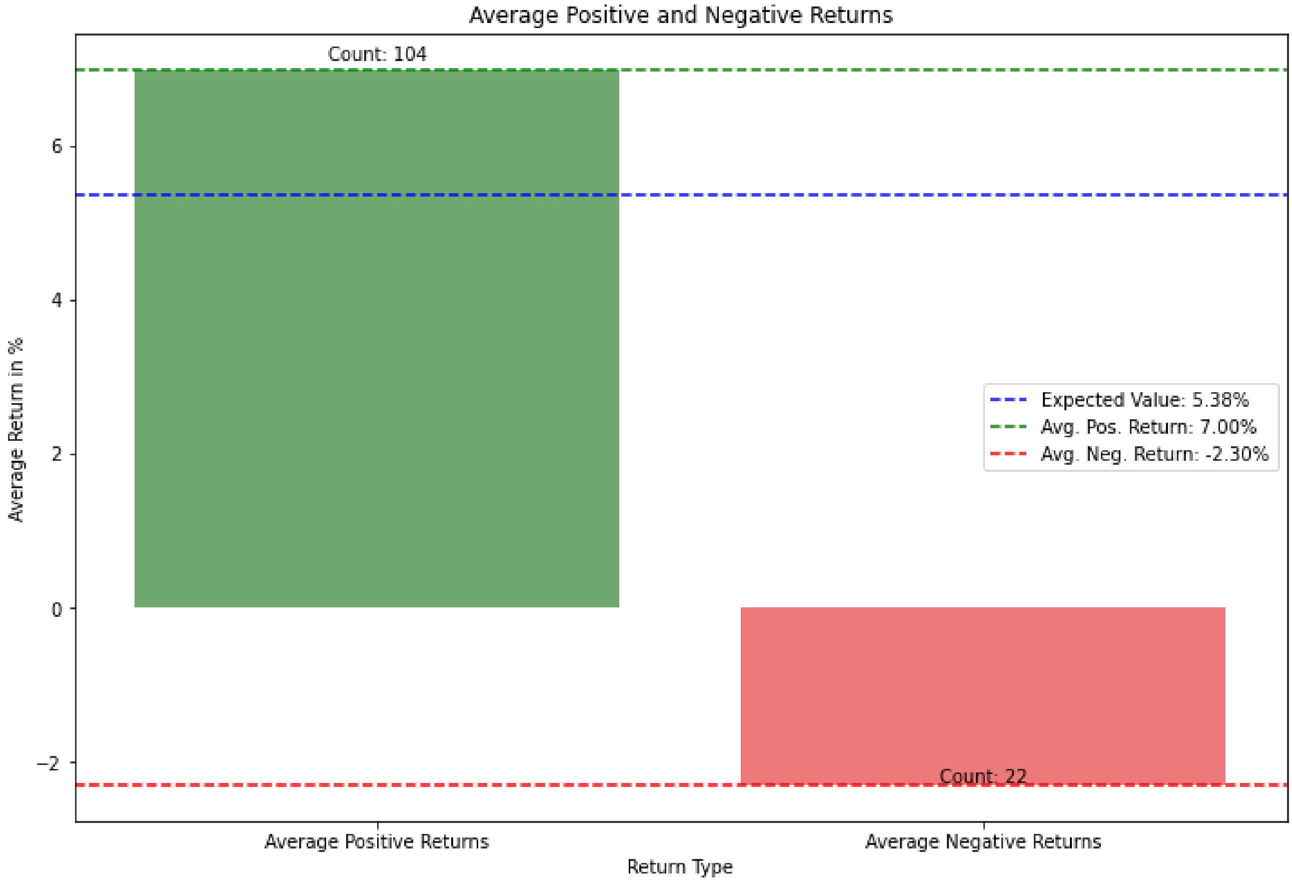


Figure 1: Average Positive and Negative Returns

Another way of looking at the query is a simple histogram and density distribution as shown in *figure 2: Return Distribution*. Once again, the different bins of the histogram and the density distribution (here: KDE kernel density estimate based on the datapoints to approximate the PDF) show a clearly right-skewed distribution towards the positive returns greater than zero.

The KDE integrates to one, i.e., the area under the smoothed curved represents the probability of either positive or negative returns. *Figure 3: Return Distribution Area* shows a probability of 80% for positive returns and 19% for negative returns. The difference of 1% is due to the choice of the bandwidth of the KDE, but the message is straightforward: There is a far higher probability for positive than for negative returns, which in turn concludes part 1 of the analysis.

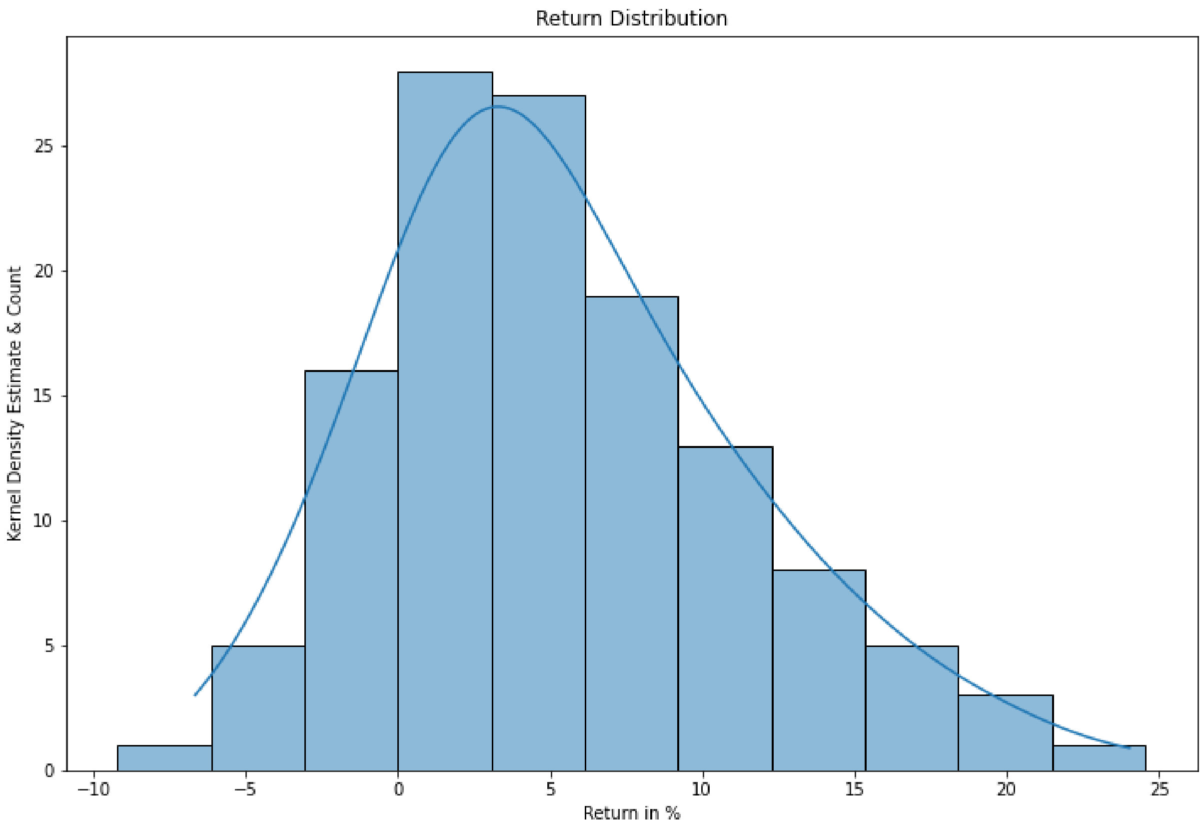


Figure 2: Return Distribution

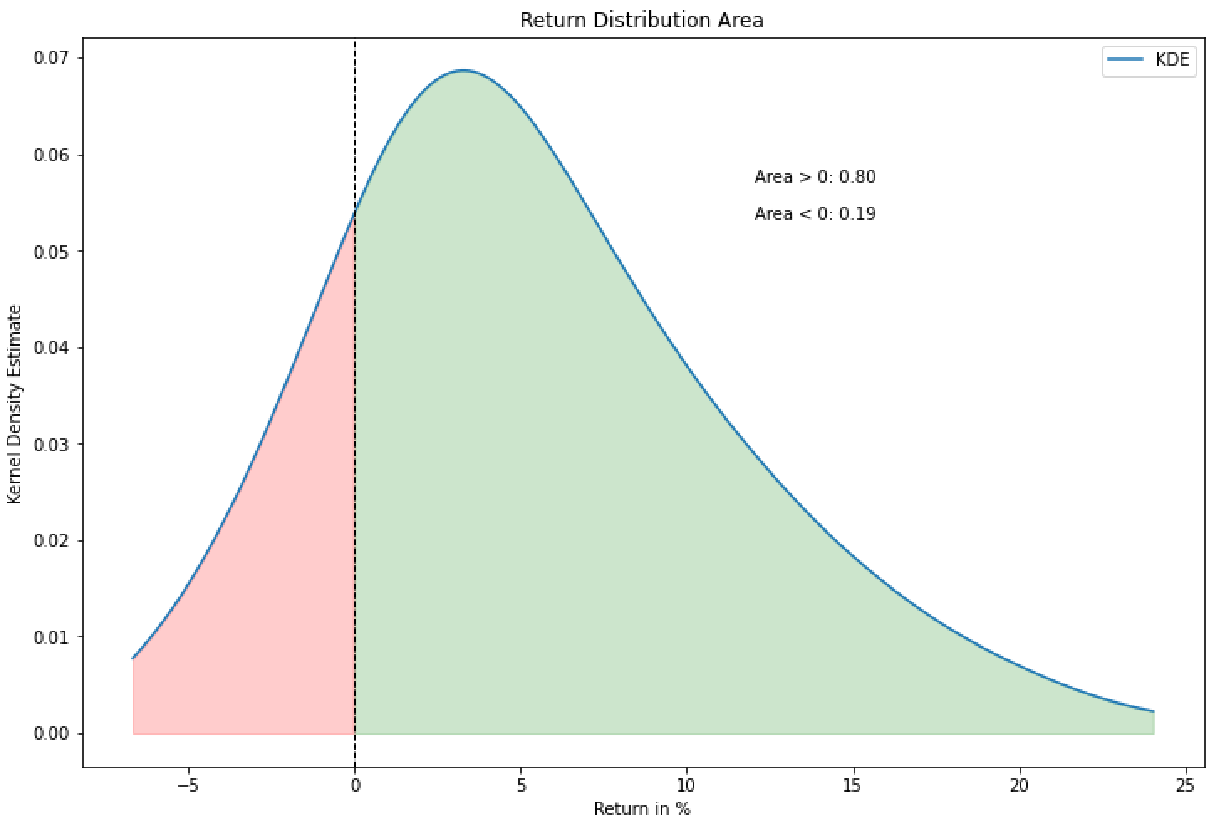


Figure 3: Return Distribution Area

4. Return vs. Risk

Generally spoken in asset pricing models, there tends to be a tradeoff between risk and return for investment assets. This concept is known as the risk-return tradeoff, which implies that higher potential returns come with higher levels of risk, while lower risk investments typically offer lower potential returns. The rationale behind this tradeoff is that investors require a higher return as compensation for taking on more risk.

However, it is important to note that the risk-return tradeoff is not always a perfect relationship. There can be situations where some investments may offer relatively high returns with lower risk, or conversely, low returns with higher risk. This can be due to market inefficiencies, mispricing, or other factors that impact asset pricing.

In this backward-looking investment strategy of the 126 portfolios of Smallville Asset Management, this statement cannot be easily confirmed, although one needs to consider that asset risk is not only influenced by its volatility, but many other factors would usually need to be taken into account (cf. the conclusions section).

We start the risk-return analysis with *figure 4: Normalized Return vs. Risk (Volatility) for All Portfolios*, in which both return (y-axis) and volatility or risk (x-axis) have been normalized in the range of [0, 1] for an easier interpretation. The horizontal line reflects the mean return of 0.39 and the vertical line the mean volatility of 0.58, creating four quadrants where Q1 shows the portfolios with a return and volatility above the mean and a total of 24 portfolios, Q2 higher return and lower volatility with 30 portfolios, Q3 with return and volatility below the mean and 42 portfolios and finally, Q4 with higher volatility and lower return and 30 portfolios, i.e., according to the typical tradeoff between risk and return we would expect most of the portfolios to be located in quadrants 1 and 3 and less in quadrants 2 and 4. However, as we can easily observe, this is hardly the case as quadrants 1 and 3 account for 66 portfolios and quadrants 2 and 4 for 60, which is almost the same number of portfolios. Therefore, the statement that the higher the risk, the higher the obtained return cannot always be maintained. The historical data for the Smallville Asset Management portfolios in 2020 confirms the statement is not always true, or only with reservations. As a matter of fact, most of the portfolios are concentrated around the mean return and mean volatility forming two clusters, as we will also see with the following data visualizations.

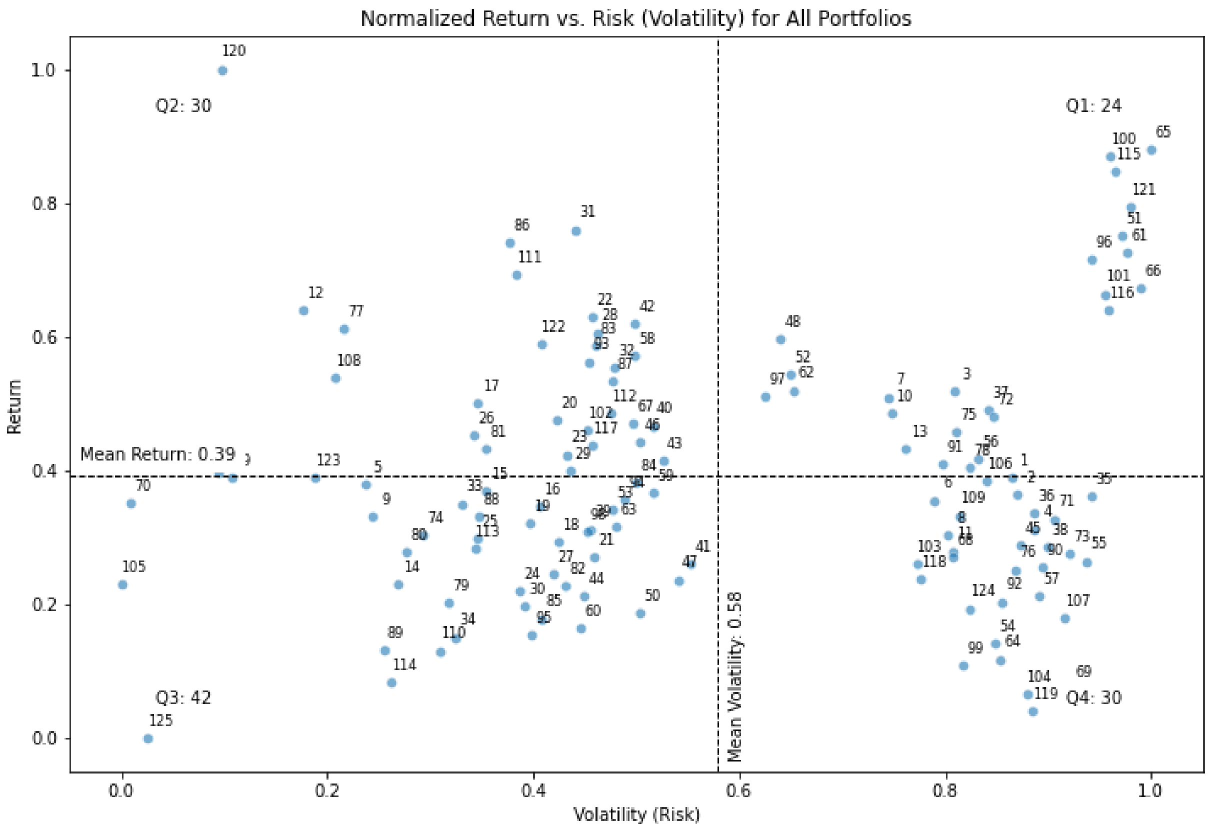
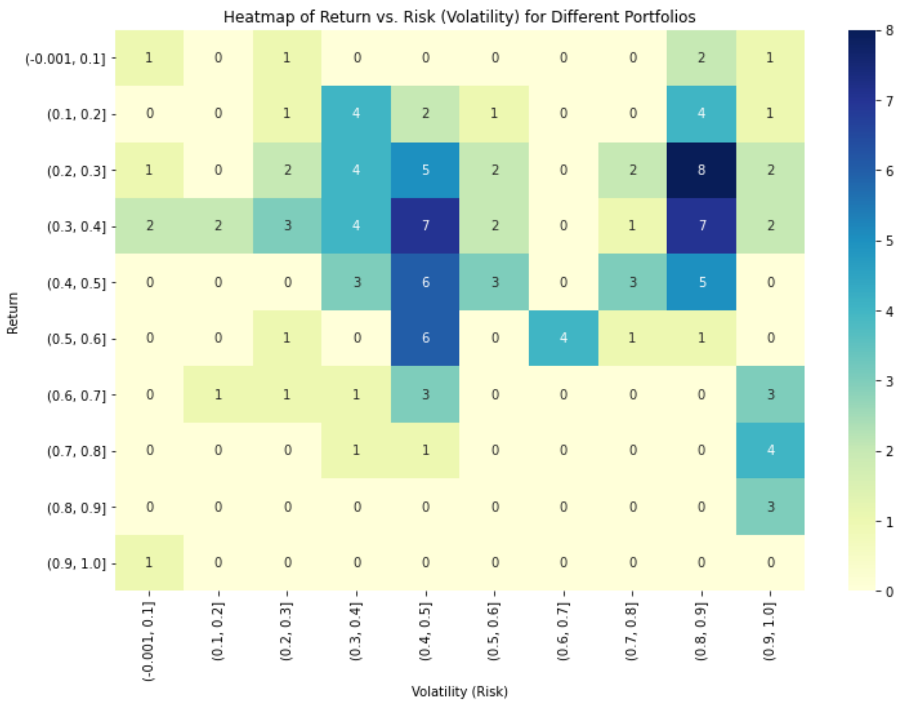
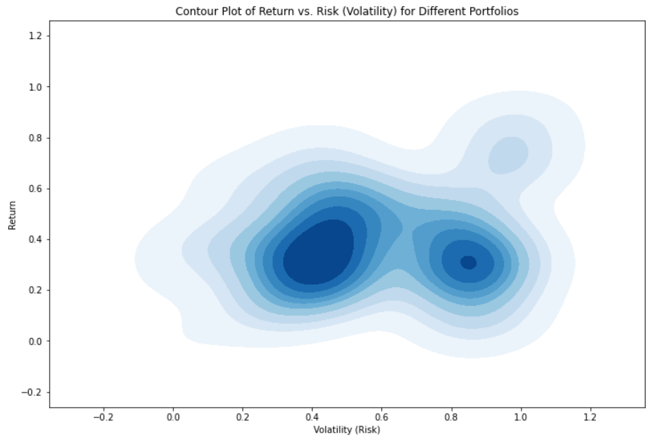


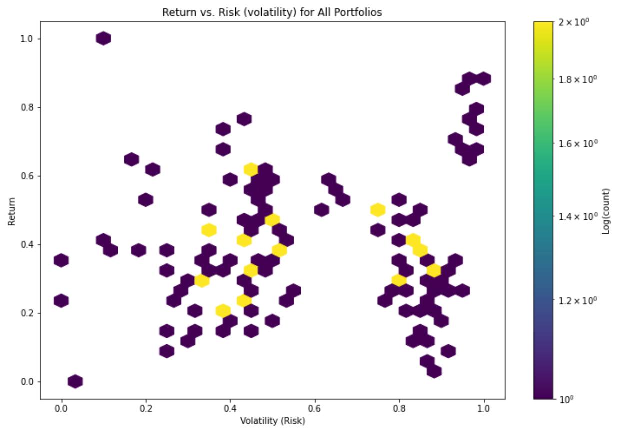
Figure 4: Normalized Return vs. Risk (Volatility) for All Portfolios

Figure 5: Heatmap of Return vs. Risk (Volatility) for Different Portfolios shows most of the portfolios are located around a low to mean return and mean to high volatility in two clusters:



Figures 5 and 6 use other data visualization techniques such as contour plots and hexbin plot and are simply included to complete the analysis of the risk-return tradeoff.





A contour plot shows the density of portfolios in the return-volatility space. The contour lines connect points with the same density. In other words, the contour lines represent regions where portfolios have the same probability of being found.

A hexbin plot is a way to visualize the density of data points using hexagonal bins. The plot is divided into a grid of hexagons, and the color of each hexagon represents the number of portfolios that fall within its boundaries. The color intensity is determined by the count of portfolios in each hexagon, with darker colors indicating higher counts and lighter colors indicating lower counts.

The aforementioned statements are, once again, confirmed: a high density of portfolios around the respective means in two main cluster, but no predominant relations of high return and high risk or low return and low risk.

5. Conclusions

From a backward-looking perspective, the data analysis is descriptive, and the two relevant queries have been answered in the previous sections, however, we would like to emphasize that an in-depth analysis would require additional features and apply different models, e.g., the following ones are just a few to mention.
For **stocks**, the Capital Asset Pricing Model is a widely-used model to estimate the expected return on a stock, taking into account its risk relative to the overall market. It uses the stock's beta (a measure of its sensitivity to market movements), the risk-free rate (usually a short-

term government bond rate), and the expected market return to calculate the expected return on the stock.

For **corporate bonds**, Merton's structural model treats a corporate bond as a combination of a risk-free bond and a put option on the company's assets. The bond's value depends on the company's asset value, its debt level, and the volatility of its assets.

For **government or public bonds**, the yield curve is a graphical representation of interest rates (yields) on government bonds with different maturities. It's often used as a benchmark for pricing other fixed-income securities, including corporate bonds. The yield curve is influenced by factors such as monetary policy, inflation expectations, and economic growth prospects.

Gold, on the other hand, does not have a single widely-accepted pricing model like stocks or bonds, but inflation (gold as hedge against inflation), currency movements of the dollar (gold expressed in dollars), interest rates (lower interest rates make gold more attractive) and geopolitical risks (gold as a safe haven) play a major role in determining the price of gold.

Cash cannot really be considered an investment asset, but of course, inflation, geopolitical risks and monetary policies have a great influence on cash holdings.

On the other hand, equalizing risk and volatility is not the full story and thus, the statements in section 4 need to be interpreted with a grain of salt. Other risk factors have been neglected in the analysis according to the assignment's instructions, however, the following list contains a series of risk factors which could be considered in future works:

- Market risk: overall market movements are measured using an asset's beta, which represents its sensitivity to market fluctuations.
- Credit risk: bond issuers default on their debt obligations
- Interest rate risk: interest rates negatively impact the value of fixed-income investments, such as bonds.
- Inflation risk: inflation erodes the purchasing power of an asset's returns.
- Liquidity risk: investor not being able to buy or sell an investment quickly.
- Currency risk: changes in exchange rates negatively impact on the value of investments in foreign currencies.
- Concentration risk: significant portion of an investment portfolio allocated to a single asset or a small group of assets.

- Political and regulatory risk: changes in government policies, regulations, or political instability.
- Reinvestment risk: investor not able to reinvest cash flows (such as interest or dividend payments) at the same rate of return as the original investment.
- Operational risk: company or investment vehicle's internal processes, systems, or people fail.

These are only a few ideas for future analyses at Smallville Asset Management as our team of proficient data scientists is always keen to improve the company’s processes and data analysis approaches.