Faculté d'Économie et de Gestion
Aix-Marseille Université

Optimization methods in finance
Master 2 MRF

Aix-Marseille Université 2022/2023 *Anthony Makarewicz*

PART I: Optimizing a function using the grid search, quasi-Newton and Differential Evolution optimization methods

On the one hand, a grid search optimization uses a grid (in the one dim case a sequence) for all values of all the decision variables of the objective function and test all possible combinations. Internally, it consists in using a nested for loop, where the number of for loops equals the number of decision variables, and the range of values of each loop the corresponding range of values for that variable. The advantage of this approach is that it guarantees the global min or max independently of the shape of the objective function. Nevertheless, the downside of such an approach is that it is not feasible in high dimensions due to limited computational power available. Indeed, for instance, if we had to test 100 values for 10 variables, we would end up with 100^{10}combinations to perform, what makes this approach more theoretical. Furthermore, this search approach is considered uninformed because it does not learn from the previous iterations.

On the other hand, in practice, one's would consider gradient based methods such as gradient descent or Newton-Raphson algorithm, or even more advanced informal searches with Bayesian optimization or differential evolution' heuristics method with genetic algorithms.
In our case, Newton Raphson is the preferred solution for gradient-based method because, unlike gradient descent, it allows to tweak the step size at each iteration (even though optimized techniques such as learning rate decay could also be possible for gradient descent).

The Newton-Raphson method is a root finding algorithm that was initially used to find the root of a function (for which $f(x) = 0$) based on the first derivative. The updating rule to find the root of a function is as follows:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

This algorithm would require an initial starting point x_0 (the better, the faster the algorithm will converge) and would find the optimal value $x_{n+1}$when $f(x_n) = 0$.
Moreover, if the derivative is also differentiable, we can apply this same principle to the first derivative of the objective function to find its roots, that is to locate critical points. As such, the function needs to be twice differentiable.
It attempts to approximate the objective function by a quadratic function, which is nothing but the 2nd order Taylor expansion of this same function.

$$f(x_k+h) \approx f(x_k) + f'(x_k)h + \frac{1}{2} f''(x_k)h^2$$

Then after each iteration, it tries to find the critical point of this quadratic function (for which $\frac{d(f(x_k+h))}{dh} = 0$). This step is achieved by finding the best step between each iteration.

$$\frac{d\left(f(x_k) + f'(x_k)h + \frac{1}{2} f''(x_k)h^2 \right)}{dh} = 0$$
$$\rightarrow f'(x_k) + f''(x_k)h = 0$$
$$\rightarrow h = -\left[f''(x_k) \right]^{-1} f'(x_k)$$

The updating rule is similar to the one used for finding roots of a function, but this time consider the first and second derivatives which is for multivariate variables the gradient vector and the Hessian matrix respectively:

$$x_{k+1} = x_k - [Hessian]^{-1}[gradient]$$

Nonetheless, this method suffers from many caveats. Indeed, the function does not distinguish between minima, maxima or saddle points (a point where the derivatives in orthogonal directions are 0 but which is not a local extrema).
Additionally, it may not converge at all and the Hessian matrix for dimensions higher than one might be non-convertible.
To remedy this, we can use a modified version such as quasi-Newton method that approximate the gradient and the Hessian numerically. Despite, it may be slow for the Hessian, and it usually requires that the objective function is (strongly) convex and that the Hessian is globally bounded or Lipschitz continuous.

Finally, more recent heuristic methods such as the genetic algorithms that follow the same philosophy than NR, namely, it uses the previous iteration and adjusts it. Basically, it mimics the principle we observe in genetic evolution in the real world and follows this workflow.
Firstly, there are many existing creatures ("offspring").
Then, the strongest creatures survive and there may be some "crossover" as they form offspring.
Afterwards, there are some random mutations in some of the offspring, as it could give them some advantages.

Lastly, the previous steps are repeated.
Consequently, we can follow the same principles with our optimization routine.
At first, we choose random values (it can be a hyperparameter to tune, for instance the offspring size). Afterwards we select only the k-best values and apply some random noise to these ones. Finally, we repeat the previous steps a certain number of times (also a hyperparameter to tune, for instance the number of generations).

PART II: Characterizing the return/risk profile of a single security

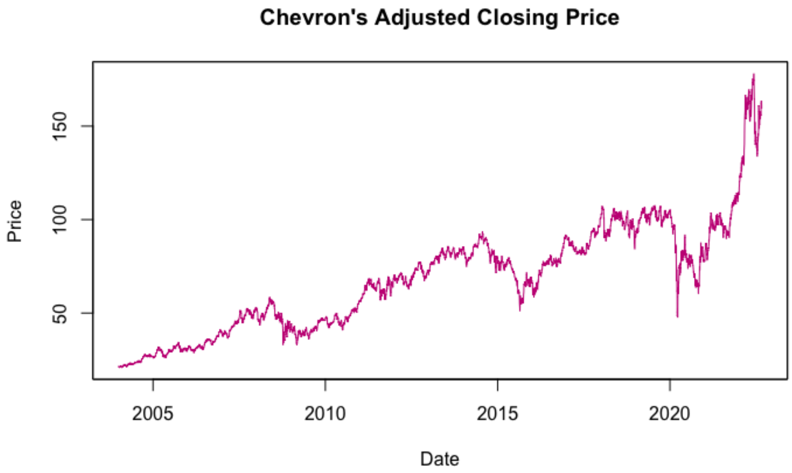
Initial analysis:

1) Chevron is a multinational American oil company. Its activities are spread out into two market segments:
First, an upstream segment comprising the company's exploration, development and production of crude oil and natural gas.
Second, a downstream segment consisting of the refining of crude oil into petroleum products.

As an oil company, Chevron's stock performance is mainly driven by oil prices. Consequently, we can display a chart of the oil future price on the future market and Chevron's price to corroborate our thought.

From the facet grid, we can observe that there are 3 major busts in the oil price. The first during the subprime crisis on August 2008, another in 2015 and the last during the covid-19 crash on March 2020. Lastly, since the Russia-Ukraine war oil prices have skyrocketed, and oil companies have benefited from this boom.

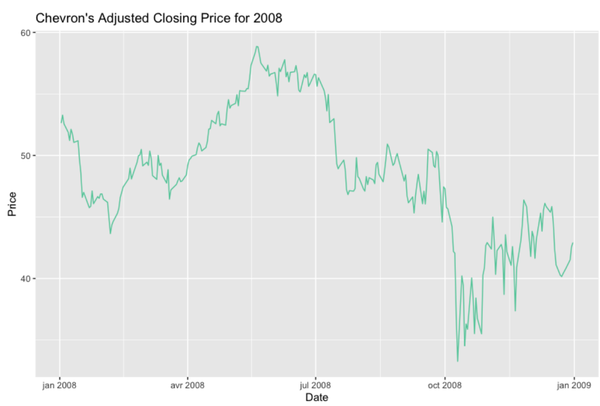
3) A more accurate chart for the Chevron's adjusted closing price can be obtained to emphasize volatility.

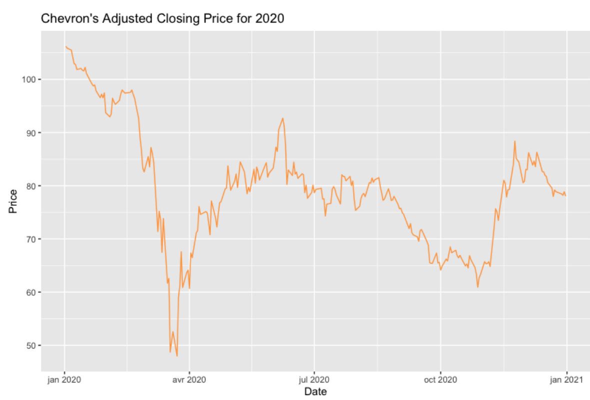


According to this chart, we can see much clearly the price evolution from 2004, and we can observe that there is a positive trend with some exceptionally large price movements. This long trend could be explained by the performance of the company itself, yet it appears excessive returns are essentially explained by macroeconomic events.

4) Moreover, we can thus identify the 2 worst 6-month periods, the first being during the year 2008 and the latter at the beginning of 2020 (we could have chosen the year 2015 instead of the year 2008, but this year shows extreme volatility as we are going to see from the rolling volatility window).

In order to select the worst 6 months, we need to plot those 2 particular years.

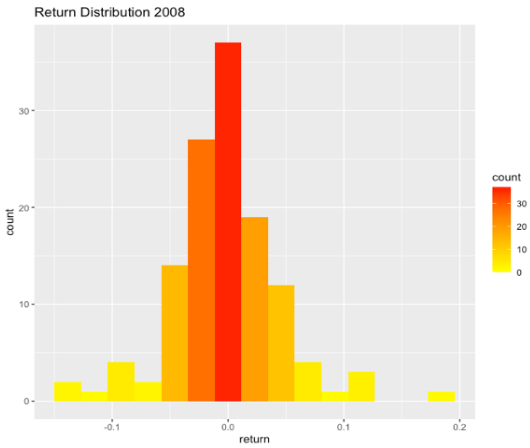


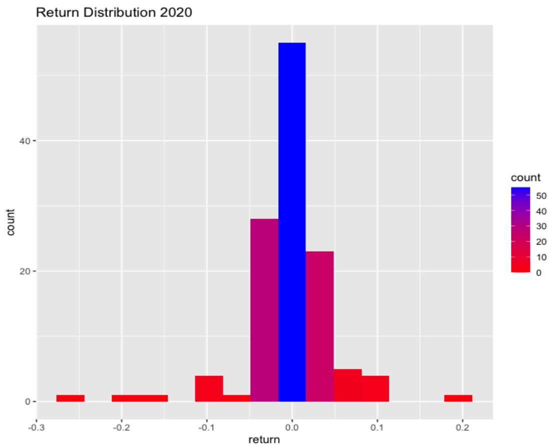


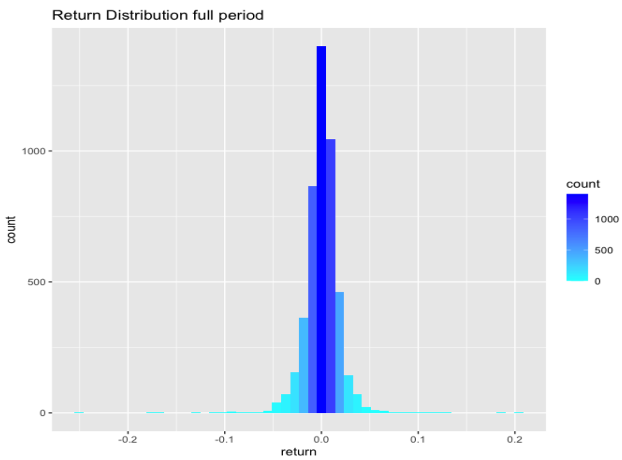
For the year 2008, we can notice that the crash happened at the very beginning of October and for 2020 it emerged at the start of the year.

We can notice that Chevron was essentially affected by macroeconomic events. The first one being the oil crash combined with the subprime crisis, the second being the oil crash in 2015 and the last one being the covid-19 again emphasized by the oil crash (the oil future price became negative because the production kept going and the demand plummeted).

To compare the return distributions, we have at our disposal several tools.
First, we can consider two popular graphical tools, namely the histogram and the boxplot. The histogram is a particular kind of plot that counts the number of observations in each bin of equal size, with each bin corresponding to a small range of the value of the variable to be explored. As we have much more observations in the full period and less in the sub-periods, we use 30 and 15 bins respectively.

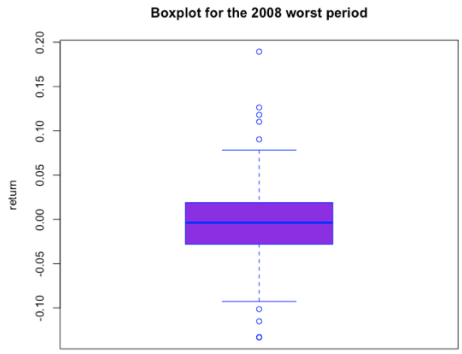


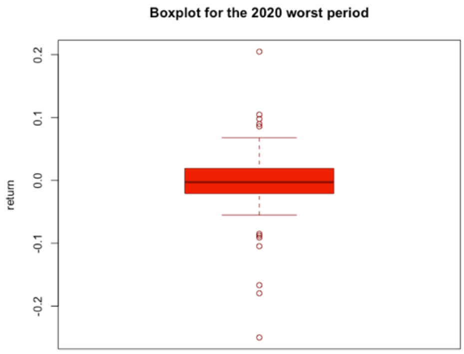


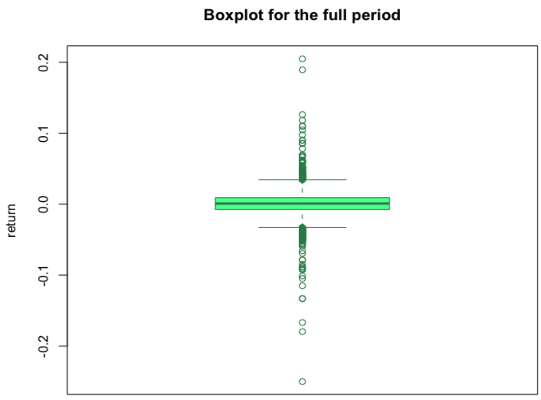


The return distributions for the 3 distinct periods are all centered at 0, in other words, the most frequent return observation (the mode) is 0 and they resemble the normal distribution.

The previous visualization with histograms could also be made with boxplots.
Boxplots are useful visualization tools that allow us to easily highlight the quartiles as well as the outliers.







The bottom of the box represents the first quartiles (Q1), the band in the middle the median and the upper bound of the box the third quartile (Q3).
The points correspond to outliers, and they are spotted outside the whiskers (the horizontal bars outside the box) which are generally computed as $+ 1.5 \times IQR$ for the top whisker and $- 1.5 \times IQR$ for the bottom where IQR denotes the Inter Quartile Range (Q3 – Q1).
Based on statistical properties described below, we can bring some description to these boxplots.
The 2008 sub-period displays a positive skewness as we observe more outliers on the top of the whisker rather than below, and we can see the inverse for the full and sub-2020 periods.
Additionally, the kurtosis property is described by the presence of excessive outliers which is considerably larger in the full period as it has the highest excess kurtosis.
Next, we have summary statistics and central moments such as the mean, standard deviation, skewness and excess kurtosis that are convenient for a numerical analysis. Indeed, for instance, we can compute the Jarque-Bera test statistic from the skewness and excess kurtosis to test for normality.

We can display the daily summary statistics along with the first four central moments.

Daily statistics	*Full period*	*2008 worst*	*2020 worst*
Min:	-0.2500	-0.1334	-0.2500
1st Quartile:	-0.0077	-0.0281	-0.0204
Median:	0.0008	-0.0037	-0.0028
Mean:	0.0004	-0.0021	-0.0022
3rd Quartile:	0.0091	0.0190	0.0183
Max:	0.2050	0.1894	0.2050

Central moments	*Full period*	*2008 worst*	*2020 worst*
Mean (Ann.):	0.1082	-0.5479	-0.5729
Standard dev (Ann.):	0.2859	0.7358	0.7938
Skewness:	-0.5192	0.3959	-0.9363
Excess kurtosis:	22.71	2.80	7.63

An important stylized fact about financial returns can be observed from these results, namely the return distribution is not stationary, it changes over time. In fact, the annual mean and standard deviation as well as the third and fourth central moment exhibit different results. For the average daily return, it is logical that the average return computed over 10 years of a positive trend is higher than the ones calculated from the worst sub-periods.

Nonetheless, the standard deviation is considerably different for the full and the sub-periods, with a value of 28% for the full period, whereas a more than 70% annual standard deviation can be reported for the sub-periods. As a matter of fact, during periods of market stress, the volatility tends to skyrocket generating large negative returns. This statistical property is described by the skewness, which underline the asymmetry of the distribution, and in this case the full period as well as the 2020 sub-period showcase negative skewness, that is they have a left-skewed distribution (large negative returns are more frequent than large positive ones). Furthermore, extremes returns are more frequent than those implied by the normal distribution, which is described by a positive excess kurtosis characterizing a leptokurtic return distribution (especially for shorter horizons, then as the return horizon increases the kurtosis tends to decrease because the return distribution looks more than the normal distribution).

These two last central moments highlight the deviation from normality, as the normal distribution has a skewness of 0 (symmetric around the mean) and excess kurtosis of 0 (mesokurtic distribution).
As they are different from 0, we can test if this result is statistically significant with a hypothesis test under the null hypothesis of normality using a Jarque-Bera test. This econometric test calculates a test statistic which is increasing in both skewness and excess kurtosis. The formula can be obtained as follows:

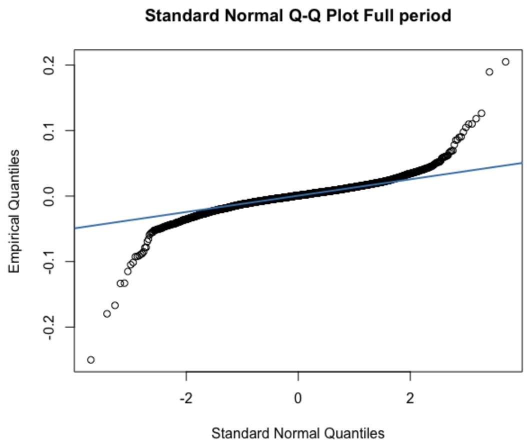
$$JB = \frac{n}{6}(S^2 + \frac{1}{4}(K-3)^2)$$

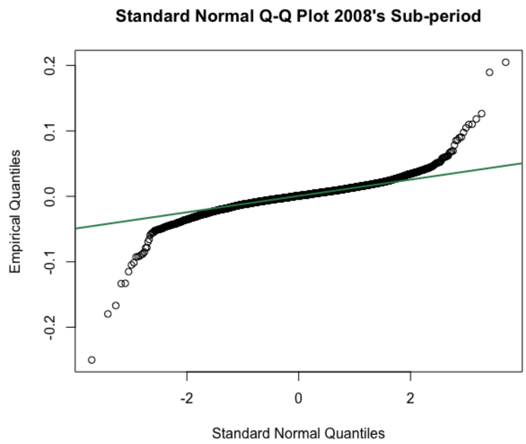
This test statistic is assumed to converge asymptotically (as the sample size goes to infinity) to a Chi-squared distribution with 2 degrees of freedom which correspond to the skewness and excess kurtosis.
From the result of this test, we are interested in the p-values which is probability of rejecting the null hypothesis whereas she is true.

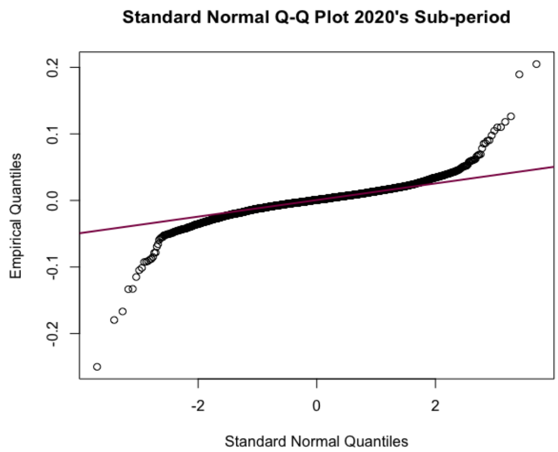
	Full period	*2008 worst*	*2020 worst*
JB:	101143	44.954	318.84
P-value:	< 2.2e-16	1.7e-10	< 2.2e-16

Here, we report p-values less than 1% for the full period and the two worst periods, so we cannot accept the null hypothesis of normality, so a better modeling choice for these returns would be a skewed student's distribution with two more parameters (d1 for the degree of fait-tails and d2 for the asymmetry) that we could estimate (even though cumbersome) with the Maximum Likelihood.

So as to corroborate our though graphically, we can display the return quantiles against those of the standard normal using Q-Q plots and see the inverse S-shape stressing deviations from the tails.







5) From now on we have considered a static approach to compute the mean and standard deviation and we have seen that because of the non-stationarity of the return distribution these measures will change accordingly. A more appropriate approach is to compute these moments with a moving window discarding the last observation and adding the most recent one.

Based on the rolling window estimates, we can clearly see the evolution of these two descriptive statistics, highlighting the non-stationarity of the distribution. During periods of market stress when the volatility shoots up, the expected return tends to decrease sharply, which is called the leverage effect.
Besides, the annualized standard deviation demonstrates 3 important features: the time variation, the persistence and the mean reversion aspect.

Firstly, for the time variation aspect, we can unequivocally see that it is not a straight line, with periods when the volatility is many times higher than its long-term value (particularly during the busts in 2008 and 2020).

Secondly, the persistence is explained by volatility clusters, that is periods of low and high volatility tend to be followed by periods of low and high volatility respectively.
Lastly, we can see from this chart that when volatility tends to be higher than its long run value (approximately 30%), it is expected to revert to this value.

According to these 3 features, it seems reasonable to consider using a simple econometric model such as GARCH (1,1) (could use more sophisticated models such as GJR-GARCH to model the leverage effect and/or skewed student's t distribution for the underlying distribution of the shocks), that could allow us to capture most of the characteristics of the time-varying variance.

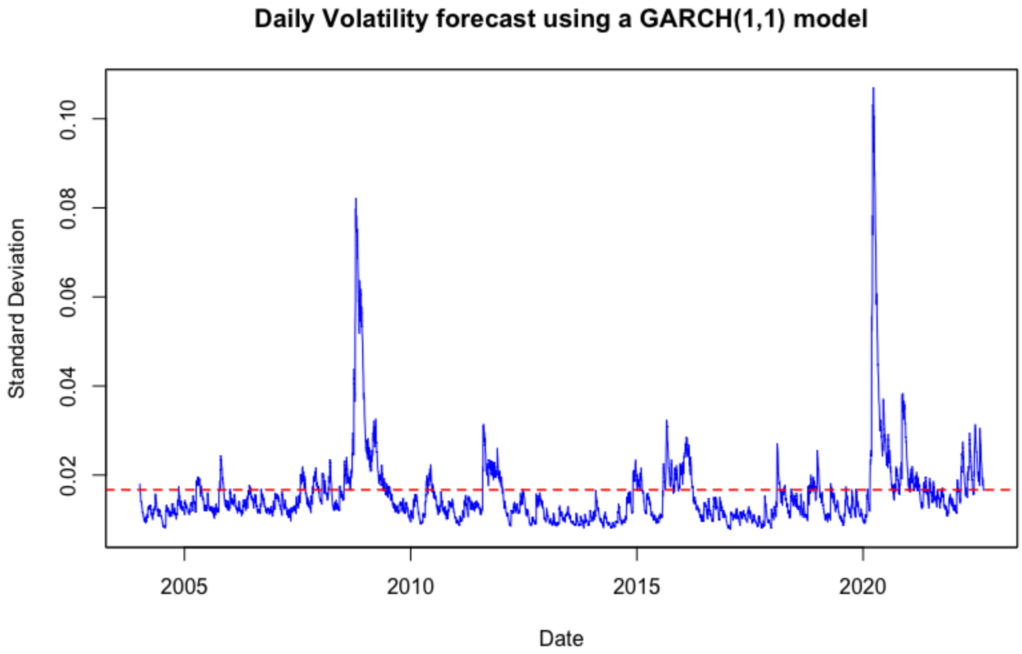
6) To this end, we can use the rugarch package that enable to quickly estimate such a model by specifying the model characteristics we want to use. A simple GARCH (1,1) model with constant mean (ARMA (0,0) for the return dynamics) and normally distributed shocks. Moreover, we can plot the estimated volatility in-sample forecasts (that is using data the model has already been trained on, so we cannot use these predictions to assess the performance our model, we would prefer to use an out-of-sample to compare the realized and predicted return).

By extracting fitted coefficients from the coef function, we can break down the mean variance model:

$$R_t \sim \mathcal{N}(6.93 \times 10^{-4}, \hat{\sigma}_t^2)$$

$$\hat{\sigma}_t^2 = 4.42 \times 10^{-6} + 0.091 \times (R_t - 6.93 \times 10^{-4})^2 + 0.89 \times \hat{\sigma}_{t-1}^2$$

$$\overline{\sigma^2} = \frac{4.42 \times 10^{-6}}{(1-0.091-0.89)} = 2.3 \times 10^{-4}$$



Based on this chart, we can notice that periods of extreme volatility are explained by major macroeconomic events such as: the 2008 subprime crisis, the oil crash in 2015 and the covid 18 in 20202. More recently, due to the Ukraine war, oil prices have skyrocketed as we can see at the end of the chart where volatility is higher than its long-term average.
Nevertheless, a period of above average was observed in 2012, which is explained by a refinery fire that happened on August 6, which seems to be the only major company event that affected the volatility.
Likewise, most of the time the volatility is less than its long-run value, and it fits what we could expect.
Indeed, unless there is an event that could lead to the bankruptcy of a company, periods of extremely high volatility are most of the time driven by macroeconomic events.

From the last day of our data (August 30, 2022), the predicted 1-day 5% Value-at-risk is -2.6370%, that is in the 5% worst cases we could expect a minimum capital loss of 2.6370% on the next trading day (as we are using daily returns). Conversely, in the other 95% cases, one's could expect a maximum percentage loss of 2.6370%.
Furthermore, if volatility doubles, then the VaR does the same, we could expect accordingly a 5.3434 % capital loss for the next trading day.
Up until now, we have computed the VaR as a percentage capital loss which is the standard approach to compute the VaR, as it allows to have a relative performance between other assets. In contrast, we can calculate the $VaR which is expressed directly in portfolio value terms.
Therefore, for a portfolio value of $1,000,00 and assuming the first volatility assumption, one's could expect that in the 5% worst case scenarios a portfolio loss of at least $26,370.46

7) Chevron exhibits a PER of 10 which suggest that its stock price is worth 10 times what is generates of net income which is slightly higher than the PER of the energy sector which is 8. Furthermore, it only has a beta of 1.20 ($\beta = 1.2$) which means that it amplifies movements in the stock market by only 20% more. Therefore, it is highly correlated to the stock market (as we have seen from annualized volatility) but is not relatively high when compared to riskier sectors such as technologies where the beta could attain 2.0, either for the Oil/Gas (Integrated) industry which highlight of beta of 1.47.
Moreover, Chevron's stock price "generates alpha" ($\alpha = 0.3$), that is it outperforms its theoretical price predicted by the Capital Asset Pricing Model (CAPM) which is characterized econometrically by a positive intercept when regressing the CVX return onto the market return.

Next, the liquidity aspect is a very important factor to consider when investing in the stock market. As a matter of fact, during periods of market stress it can be very costly to unbind its position leading to significant capital loss.
For that, Chevron stocks are exchanged on a daily basis on the New-York Stock Exchange (NYSE) and the %float is 91.6%.
The float corresponds to the number of shares that are not kept from investors and exchanged on the stock exchange, the more its value is, the more interesting it is for the investor because it means that if we must exit a position there will be enough liquidity in the market to unwind our position.
Moreover, Chevron's stock is a blue-chip stock, as it is included in the Dow Jones Industrial Average which incorporates the 30 largest companies in the US in terms of economic activity.
Despite financial crashes, the annualized volatility is relatively stable with a maximum of only 30% during the refinery in 2012 and the oil crash in 2015, which is approximately the same than Oil/Gas (Integrated) industry in which Chevron is.

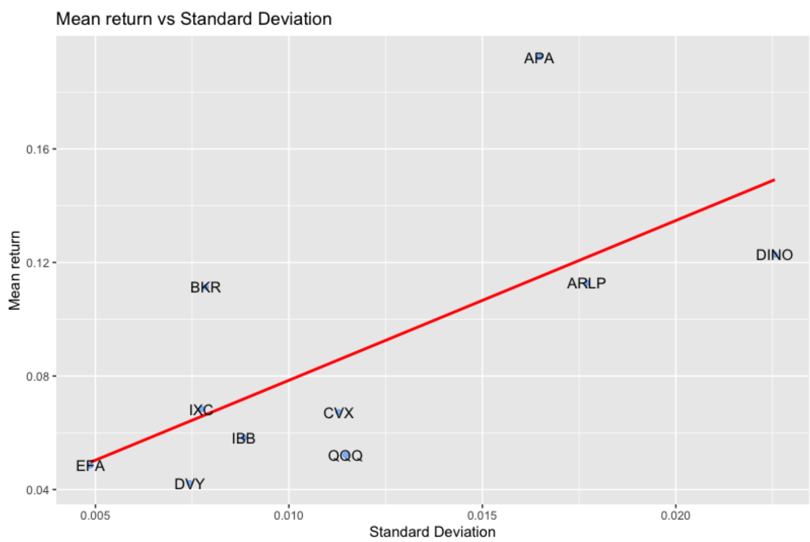
However, investing in the stock market is associated with a higher risk compared to a risk-free investment, due to the variation and extreme variation of the value of the investment. Central moments computed for the full period, suggest that investing in Chevron is associated with a higher risk which is characterized by a high volatility, negative skewness and positive excess kurtosis.
Accordingly, Chevron's investors should have a relatively low risk-aversion and would be willing to take a high risk in order to achieve a mean annualized return of approximately 10%.

From financial analysts' recommendation, it appears that the energy sector occupies the 6Th position with a recommendation of 2.23 relative to the 11 sectors in the US, and that Chevron achieves a 2.40 recommendation which is higher than the energy sector.

Additionally, Berkshire Attaway the famous fund of the investor Warren Buffet has opened significant positions in Chevron in the last few months, which indicate that it would be appropriate for a long-term perspective. Furthermore, Chevron is a high-dividend stock (3.05%) compared to other sectors, despite being less than the energy sector which display a 5.86%'s dividend yield.
Nevertheless, for short-term horizons, it could also be possible to trade Chevron's stock despite it would not be recommended for fundamental investors and would require more market data.

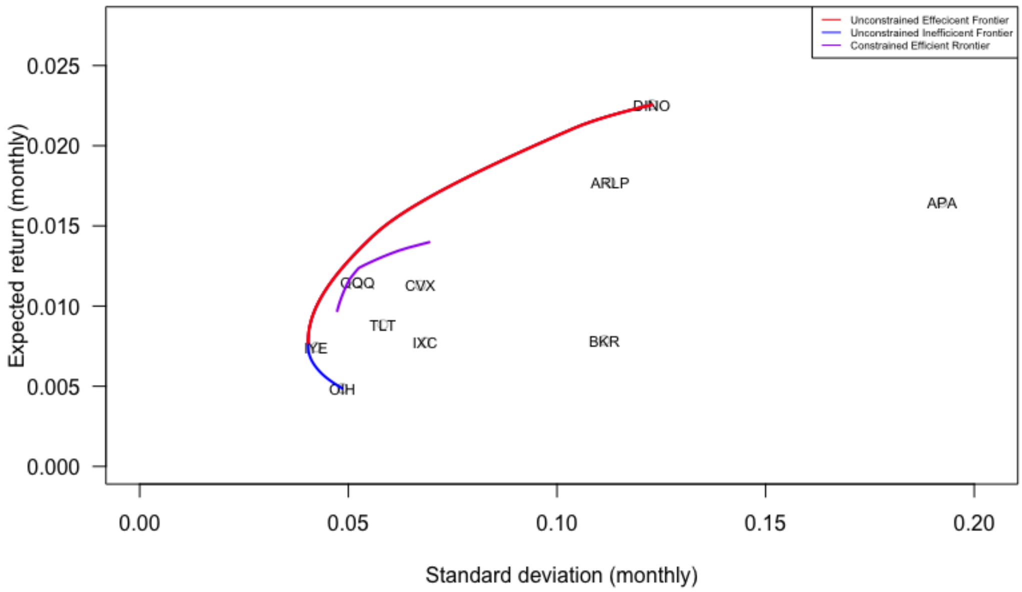
Portfolio analysis:

10) From a scatterplot of mean return vs. standard deviation we can see the risk-reward tradeoff in finance, it basically means that a higher return is associated with a higher risk.



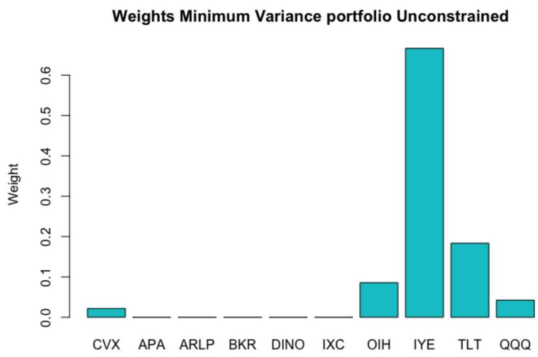
Nevertheless, as the asset universe is extremely large, thanks to diversification and correlation coefficients, it is possible with an optimization to construct a portfolio of assets that enable for a given amount of risk to maximize the expected portfolio return and reciprocally for a given expected return minimize the portfolio risk.
Indeed, based on the plot we can see that for a given amount of risk we can choose the asset that achieves the highest expected return (for instance BKR compared to IXC), and that for a given expected return we can choose the one with the lowest risk (also BKR compared to ARLP).
However, finding all portfolio combinations is not satisfactory due to investment mandate that describe how the fund manager should allocate funds based on the needs of its clients.

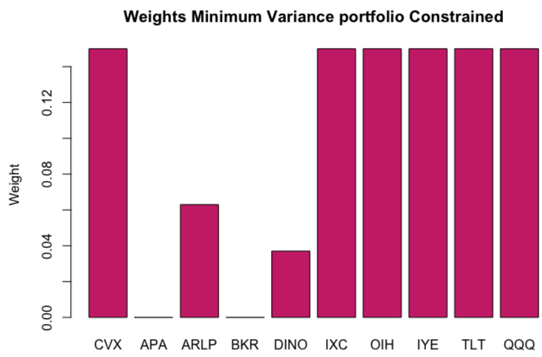
Therefore, we must add specific constraints such as: full investment, long-only, minimum and maximum box weights, group weights by sectors and geographical regions ect …
In our problem, we consider a long only and full investment constraints.
Nonetheless, by adding more constraints to the optimization (in this case switching a maximum box weight of 100% by 15%), the space of all possible solutions is restricted so less optimal values will be found. By the way, at a given standard deviation value the maximum expected return would be achieved at lower value than the less restricted weights.
To highlight this effect, we can visualize the efficient frontier with weight constraints between 0 and 100% and weights between 0 and 15% (will be named “constrained” efficient frontier and we will use this hyperbole to stress this effect):



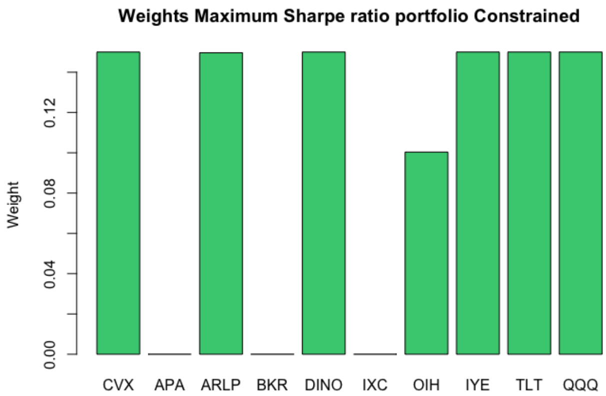
From this plot, we can observe that replacing a maximum box weight of 100% by only 15%, there are less optimal solutions for the constrained efficient frontier in purple.
Furthermore, for all given standard deviation, the “unconstrained” efficient frontier (the one with weights between 0 and 1) achieves a higher expected return.

11) Yet, simply restricting weights between 0 and 1 may not seem optimal from a diversification standpoint, although we avoid short selling and leverage effect by restricting weights between 0 and 1 respectively.
Indeed, by plotting the optimal allocation weights for the minimum variance portfolio with maximum weight constraint of 100% and 15%, we can see the weight distribution.





According to these two bar charts, we can point out that for the weight constraints between 0 and 100% there are only 5 assets that have been selected with little weight for 4 of them, except IYE that represents more than 60% of the portfolio. On the contrary, the weights for the maximum weights of 15% are more spread out, which indicate a more diversified portfolio with only 2 assets that have not been retained.



Now by comparing the weights between the maximum Sharpe portfolio with that of the minimum variance portfolio, it comes out that the maximum Sharpe ratio weights are less distributed between the assets, while the minimum variance portfolio with only one weight at 10% that is less than the maximum weight constraint of 15%.

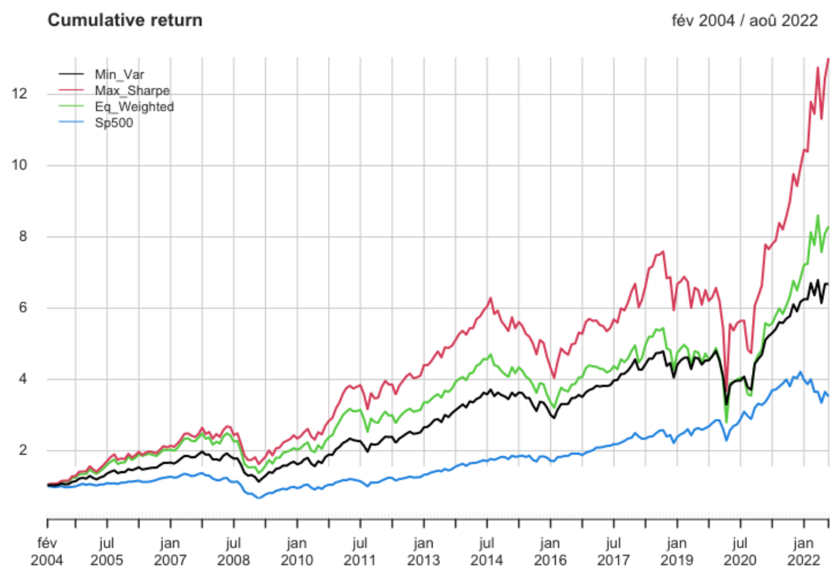
12) For the back-testing procedure, we need to reperform the optimization but this time using only the data available until the 1st January 2011 to avoid look ahead bias.
The only time data comes into the optimization is through the mean vector and covariance matrix, so we subset them to include the data for the estimation.

After estimating the weights and applied them to the estimation as well as the evaluation sample, we can merge these two sub-samples together and column bind them with the equally weighted portfolio and the S&P500 for the full period.

The first table mentioned below allows us to compare the annualized mean, standard deviation and the Sharpe ratio with a risk-free rate of 0% for these 4 strategies.
From this table, we can notice that the S&P 500 is the one that has the lowest annualized standard deviation and the lowest annualized return, which comforts what we could expect for the standard deviation as it encompasses the best 500 companies in the United States.
Then, the minimum variance portfolio is the one achieving the highest Sharpe ratio, which is based on this metric the best portfolio.
The maximum Sharpe ratio portfolio achieves the highest annualized mean, but at the cost of a larger annualized standard deviation.

	Min Var	Max Sharpe	Eq Weighted	Sp500
Annualized Return:	**0.107**	**0.148**	**0.121**	**0.070**
Annualized Std Dev:	**0.164**	**0.246**	**0.225**	**0.146**
Annualized Sharpe (Rf=0%):	**0.656**	**0.602**	**0.535**	**0.479**

Below, a chart displaying the cumulative return of these 4 strategies enable us to compare the compounded return, that is the evolution of the value of a $1 investment across the full time period:



From this chart, we can mention that the maximum Sharpe ratio generates the highest cumulative return and particularly at the beginning of the evaluation sample in July 2011, reaching $11 at the end of the full sample.
However, it loses its advantage during the covid-19 crash as it is the one with the highest annualized standard deviation (that we are going to see below) but skyrocket after that.
In contrast, the S&P500 accomplishes the worst cumulative performance reaching only $4, compared to $8 and $6.5 for the equally weighted strategy and minimum variance respectively.
Nevertheless, the S&P 500 is less responsive to shocks, especially during the covid 19 crash as it has the lowest annualized standard deviation.

Finally, we can plot the rolling performance of the annualized mean, standard deviation and the Sharpe ratio using a 24-month window to have a sense of how these metrics change over time.

From this chart, we can note that the annualized returns for the minimum variance, maximum Sharpe ratio along with the equally weighted portfolios seem to be very close, in contrast to the S&P 500 that sometimes underperforms them (subprime crisis) and outperforms them (covid-19 crash).
Just below, the annualized standard deviation appears to be relatively stable until the covid-19 crash, but with a relative order. Indeed, the maximum Sharpe ratio portfolio annualized volatility is always higher than the equally weighted portfolio, which itself higher than the minimum variance portfolio and itself higher than the S&P 500.
The annualized Sharpe ratio on the other hand follows the same trend as the annualized mean.

So far, we have been only focusing on the standard deviation as a risk measure, which is the standard metric used to analyze risk but this measure is sometimes inappropriate.
Indeed, the standard deviation penalizes equally deviations above and below the mean, and it is not focused on the left tail of the return distribution (which is what investors care about).
Therefore, we can include in our analysis charts and metrics that can counteract these drawbacks.
One popular alternative recognized by Markowitz as giving even better portfolios results, but due to limited computational resources at that time could not be used, is the semi-deviation. It is a lower partial moment that computes return deviations below the mean and comply with one important aspect of a good risk measure, namely it should be asymmetric.

To this end, we can compute the Sortino ratio which is an alternative risk-adjusted performance measure to the Sharpe ratio but using the semi-deviation instead of the standard deviation.
As there are less observations below the mean than the all returns distribution, we expect them to be less than the Sharpe ratio.

$$\text{Sortino Ratio} = \frac{r_p - r_f}{\sigma_d}$$

- $r_p \rightarrow$ Portfolio Return
- $r_f \rightarrow$ Risk-Free Rate
- $\sigma_d \rightarrow$ Standard Deviation of Negative Returns (Downside)

	Min Var	***Max Sharpe***	***Eq Weighted***	***Sp500***
Semi-deviation:	0.035	0.048	0.045	0.032
Sortino Ratio:	0.316	0.337	0.288	0.225

Here we can observe that the relative order is slightly different than the annualized Sharpe ratio, as now the maximum Sharpe ratio portfolio surpasses the minimum variance one.

The semi deviation is a useful metric to use, yet it is not concerned on the left tail of the distribution. For that, we could simply use the Value-at-risk (VaR) which is the standard tail measure used in finance, but it is simply a quantile and will not look beyond losses higher than that number.
Even though we are dealing with monthly returns, which highlight less excess kurtosis as excess kurtosis is decreasing with the return period, we can still use the non-normality aspects as skewness sometimes increases (in absolute value term).

Therefore, we can also compute the moments and check for normality using the same Jarque-Bera test but this time using directly the apply function on all columns.

	Min Var	*Max Sharpe*	*Eq Weighted*	*Sp500*
Skewness:	-0.34	1.06	0.40	-0.62
Excess kurtosis:	2.35	12.14	7.40	1.59
JB:	55.71	1411	515	37.7
P-value:	8.0e-13	< 2.2e-16	< 2.2e-16	6.3e-09

On the one hand, we report a left-skewed return distribution of portfolio returns for the minimum variance portfolio along with the S&P 500, which underline that large negative returns are more frequent than positive returns and so it decreases the utility for the investor.
On the other hand, the maximum Sharpe ratio as well as the equally weighted portfolios display positive skewness as a sign of a right-tailed return distribution which increases the utility for the investor.
Moreover, the positive excess kurtosis indicates a fat-tailed distribution with extremely large returns are more frequent than the normal distribution (which is mesokurtic), and therefore has a negative effect on the utility of the investor.
Again, as the p-values are all less than 10%, 5% and 1%, we can statistically reject the normality assumption and use this this information to compute more advanced tail metrics.

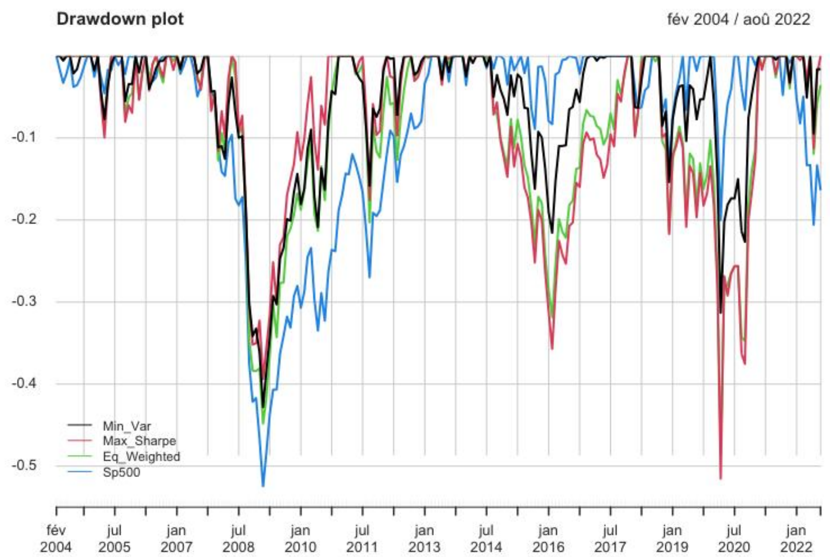
One very popular metric used, which is the standard approach for bank capital requirements for computing the market risk in the internal model approach is the Expected Shortfall (or Conditional VaR). Therefore, we can report it with a coverage rate of 2.5% that allows to take the mean of the 2.5% worst observations beyond the VaR. We use 2.5% instead of the traditional 1%VaR because we have much more observations to estimate accurately the most negative returns (which is similar in the Extreme Value Theory where the threshold from which we estimate the tail distribution must not be low, as we would obtain a bias estimate because we could no longer assume the asymptotic convergence to the generalized pareto distribution(GDP), but must not be to high so as to collect enough data to have an accurate estimate) .

Here, we report the ES using the Cornish Fisher extension which is a modified version of the gaussian ES, that enable to include the skewness and excess kurtosis into the calculation. Additionally, we can compare it with the VaR which must be less or equal than the CVaR and using the standard gaussian approach which should be lower than the modified version (in absolute value terms).

Tail Risk	*Min Var*	*Max Sharpe*	*Eq Weighted*	*Sp500*
VaR:	-0.097	-0.137	-0.135	-0.09
ES:	-0.10	-0.15	-0.14	-0.09
ES (modified):	-0.17	-0.44	-0.37	-0.14

The average loss that an investor could expect in the 2.5% worst cases for an investment horizon of one month is 17% and 44% for the minimum variance and the maximum Sharp ratio portfolios respectively.
Also note that the relative order of these risk measures with the ES always larger (absolute value) than the VaR and the modified ES larger than the gaussian ES.

So far, we have proposed extensions of the standard deviation as risk measures based on the return distribution but have neglected one important feature that causes pain to investor, namely the drawdown risk.
As a matter of fact, successive negative returns provoke more pain to investors than one-period negative returns considered so far from return distribution.
Therefore, it would not be exhaustive if we did not take into account this psychological characteristic.
Two approaches can be considered.
The first approach is to use a drawdown plot that shows the peak-to-though decline in percentage for a given investment.



From this chart, we can argue that the S&P 500 achieves the highest drawdown during the subprime crisis on September 2008 with a 52% drawdown, describing the capital loss incurred from investing at the peak and selling at the bottom.
Nonetheless, except the subprime crisis, the S&P 500 achieved relatively low drawdowns compared to the other portfolios. As a matter of fact, on January 2016 and during the covid-19 crisis, the other portfolios realized higher drawdowns compared to the S&P 500.
Even though this graphical representation gives us a sense of all the possible drawdowns, it does not give us a single number evaluation metric, which is very important when we must choose between a lot of different portfolios.

Hence, we can directly compute the maximum drawdown for all portfolios and obtain the Calmar ratio which is its return-adjusted version:

$$Calmar\ Ratio = \frac{\underline{R} - Rf}{Max\ Drawdown}$$

	Min Var	*Max Sharpe*	*Eq Weighted*	*Sp500*
Max Drawdown:	0.427	0.515	0.487	0.525
Calmar Ratio:	0.251	0.287	0.248	0.133

From the Calmar ratio, the worst portfolio would be the S&P 500 reaching a 13.3% Calmar ratio and the optimal portfolio would be the maximum Sharpe ratio portfolio.

Nevertheless, the maximum drawdown approach only considers one peak-to-through decline computed for the full period, it doesn't account for all the drawdowns and particularly extreme ones along with the length of these drawdown/s.
To this end, we can compute the Ulcer Index (UI) that can be considered as an average risk metric as it uses all drawdowns.

$$UI = \sqrt{\frac{1}{T}\sum_{t=1}^{T} {DD_t}^2} \ ,$$

where $DD_t = \frac{P_t}{\max\limits_{i \in [0,t]} P_i} - 1$ is the drawdown at date ***t***.

As investors are particularly sensitive to extreme losses, we could push even further and penalizing extreme losses by considering the DaR (Drawdown-at-Risk) or CDaR (Conditional Drawdown-at-Risk) which are the equivalent of the VaR and CVaR but using the drawdown distributions instead.
Here, we report the Ulcer Index ("UI") along with the CDaR which is nothing but the expected value of the 5% worst drawdowns.

Drawdown risk	*Min Var*	*Max Sharpe*	*Eq Weighted*	*Sp500*
Ulcer Index:	0.107	0.130	0.137	0.148
CDAR (5%):	0.172	0.212	0.243	0.155

The conditional drawdown in itself is relevant with respect to the particular amount risk the investor agreed to take on (risk-reward tradeoff) so we must be careful when using CDaR as a risk measure. Thus, we want to penalize large deviations from CDaR with respect to the standard risk measure used, that is the standard deviation. We use the Pitfall indicator which is defined as the average of the worst drawdowns expressed in units of volatility.

$$PI(\alpha) = \frac{CDaR(\alpha)}{Vol}$$

Consequently, we can use a penalized risk metric that combines the "average loss" Ulcer index along with the "extreme risk penalty" pitfall index.

$$\text{Penalized Risk} = \text{Ulcer} \times \text{Pitfall}$$

Finally, we can harness this penalized risk metric so as to compute the Serenity Ratio, and the higher it would be, the more serene the investor will be as she will expect less extreme risk than the one she agreed to take on.

$$\text{Serenity Ratio} = \frac{\text{Return}}{\text{Penalized Risk}}$$

	Min Var	*Max Sharpe*	*Eq Weighted*	*Sp500*
Pitfall Indicator:	1.052	0.860	1.080	1.056
Penalized risk:	0.172	0.212	0.243	0.155
Serenity Ratio:	0.950	1.330	0.814	0.449

By considering the Serenity ratio as our ultimate risk-adjusted performance metric, we would choose the highest that is the maximum Sharpe ratio portfolio.

Thus far, we have been using the row-merged version of our 2 optimized portfolio considering the full time period.
However, we have not checked the performance on these two samples to check how well we would perform in practice.
For that purpose, we can compare the absolute performance of our optimization task with annualized performance metrics.
As a single number evaluation metric, we could consider the Sharpe ratio which is the most used metric to choose between portfolios and choose the one with the highest annualized Sharpe ratio on the evaluation sample.

Minimum Variance	*Estimation*	*Evaluation*
Annualized Return:	0.108	0.107
Annualized Std Dev:	0.162	0.166
Annualized Sharpe (Rf=0%):	0.670	0.655
Serenity ratio:	1.635	1.534

Maximum Sharpe ratio	*Estimation*	*Evaluation*
Annualized Return:	0.182	0.128
Annualized Std Dev:	0.193	0.273
Annualized Sharpe (Rf=0%):	0.944	0.470
Serenity ratio:	2.309	1.607

We can notice that the minimum variance portfolio's performance on the evaluation sample is almost the same than the estimation sample, which informs us that our portfolio generalizes well to new data.
In contrast, the maximum Sharpe ratio portfolio is performing poorly using new data, which highlight the fact that we cannot use data that the optimization has already been trained on to evaluate the final performance.
Nonetheless, in practice, we would have re-estimated weights maybe every month or every quarter as market conditions change rapidly leading to significant estimation errors. Indeed, as the mean and variances are not stationary, we must estimate them, and these estimation errors will be transmitted to the estimated weights accordingly.

Remedies of this drawback will be re-estimating weights every month or every quarter, using robust estimators, factor models or shrinkage methods instead of the sample estimates.

13) Even though the maximum Sharpe ratio portfolio has a slightly higher Serenity ratio on the evolution sample, its performance from the estimation to the evaluation sample plummets.

In contrast, the minimum variance portfolio outperforms the maximum Sharpe ratio portfolio in the evolution sample using the annualized Sharpe ratio but underperforms using the Serenity ratio. When choosing between 2 portfolio we must be careful about the bias as well as the variance of our estimates.

The bias can be understood as the ability of our estimates to fit accurately the estimation sample, so choosing the best weight allocation possible.

In contrast, the variance of our estimates is the difference of performance from data the optimization has been performed and the new data to measure how well it would perform in the real world.

The estimation sample has been obtained using monthly data from approximately 10 years of data, so the performance is assessed for a 10-year investment.

However, long-term investors are looking for the most reliable estimate of their return at the end of the investment horizon.

For that purpose, despite the minimum variance portfolio has the highest bias (lowest performance on the estimation sample for the annualized Sharpe as well as the Serenity ratio), it achieves very low variance (performance between the two samples are similar) compared to the maximum Sharpe ratio portfolio switching from 0.944 and 2.309 for the annualized Sharpe ratio and Serenity ratio to 0.470 and 1.607 on the evaluation sample.

Compared with the benchmark investment of the S&P 500 which is a market capitalization-weighted portfolio, the S&P 500 is exposed to a concentration risk in US stocks compared to an international strategy (see 52% drawdown on the S&P during the subprime crisis and the highest Ulcer index).

Given the current period of high economic uncertainty, inflation and rising rates environment, the minimum variance portfolio seems to be the most reasonable choice and would be particularly adapted for a long-term investment as we have a good out-of-sample performance. Indeed, during periods of high volatility, investors are looking for the most secure investment as their utility function is convex when they lose money (see Prospect theory D.Khaneman and A.Tversky 1979), even though they will still be looking for an expected return higher than the risk-free rate while minimizing risk.