**Submission Number: 2**

**Group Number: 3**

**Group Members:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Full Legal Name** | **Location (Country)** | **E-Mail Address** | **Non-Contributing Member (X)** |
| Shubham Shyamal Kotal | India | [shubhkotal@gmail.com](mailto:shubhkotal@gmail.com) |  |
| Ishaan Narula | India | [ishaan.narula@outlook.com](mailto:ishaan.narula@outlook.com) |  |
|  |  |  |  |

**Statement of integrity:** By typing the names of all group members in the text box below, you confirm that the assignment submitted is original work produced by the group (*excluding any non-contributing members identified with an “X” above*).

Shubham Shyamal Kotal, Ishaan Narula

Use the box below to explain any attempts to reach out to a non-contributing member. Type (N/A) if all members contributed.

N/A

*\* Note, you may be required to provide proof of your outreach to non-contributing members upon request.*

**Answer 8.1: Report to Portfolio Manager**

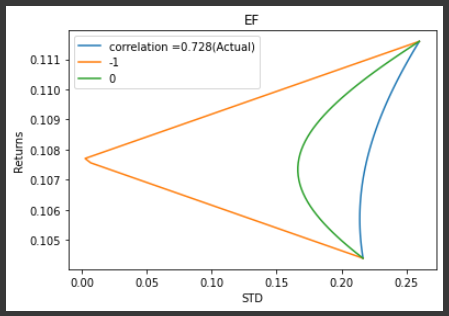
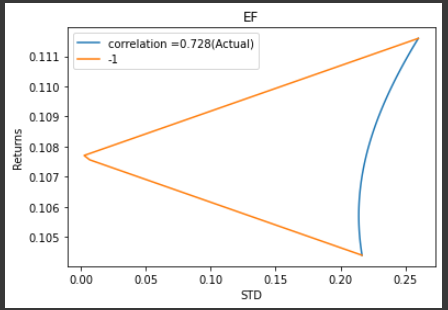
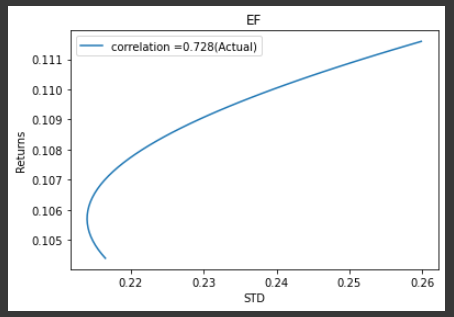
**Introduction**

Daily adjusted closing price data is imported for ETFs covering 11 US sectors. Based on this, we analyse the portfolio statistics and efficient frontier for a 2-ETF portfolio (XLK, XLI) and a 3-ETF portfolio (XLK, XLI and XLY) of our choice. We then extend this analysis to all 3-ETF portfolios which can be formed from our 11-ETF dataset. Thereafter, the economic indicator (LEI, CEI, LAG) and positional (long + / short -) categories for each ETF in each portfolio are specified, to look for indicator buckets which tend to be positively associated with returns. Finally, we construct a new portfolio from the first 3 principal components of our 11-dimensional dataset and compare its return performance over 2019 and 2020 with our 3-ETF portfolios.

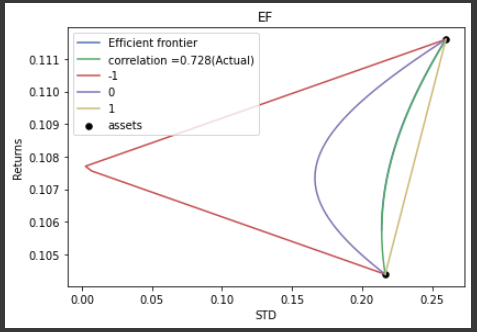
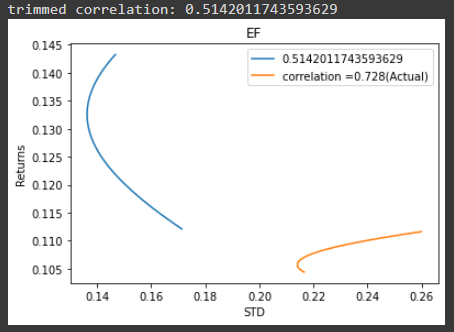
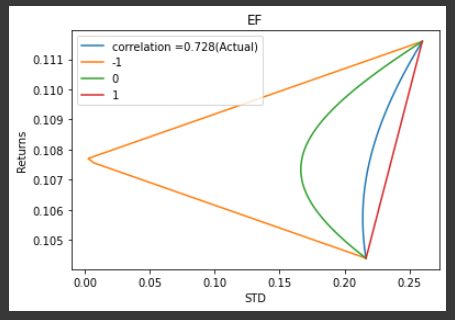
**2-sector (XLK, XLI) and 3-sector (XLK, XLI, XLY) Portfolios**

**2-sector (XLK, XLI)**

**Fig 1 Fig 2 Fig 3**



**Fig 4 Fig 5 Fig 6**

  
  
The sectors when compared shows the correlation of 0.72 which is been seen in the 1st figure above. When pretended the correlation between the securities is -1. The shape obtained show the perfect relationship, straight line which can be seen on the 2nd figure. When pretended the correlation between the securities is -1, zero-correlation EF is located between the actual-correlation 0.728 EF and the perfectly negative correlation EF seen on figure 3rd. When pretended the correlation between the securities is 1, negative correlation EF is connected by a perfect correlation of EF which can be seen on the figure 4th. When the data was trimmed by 5% and the correlation and EF was obtained again, the correlation of the actual was seen much better than the correlation of the trimmed which can be seen on figure 5th. The figure 6th shows the efficient frontier along with actual correlation, 0, 1 & -1.

**3-sector Portfolios – All Combinations**

*Efficient Frontier Estimation*

The Critical Line Algorithm (CLA) along with a covariance shrinkage has been used to estimate efficient frontiers for all 165 3-ETF portfolios which can be constructed from our 11-ETF dataset. This has been done for 2019 returns CLA was developed by Harry Markowitz to optimise general quadratic functions subject to linear inequality constraints. CLA solves any portfolio optimisation problem that can be represented in such terms, like the standard Efficient Frontier problem.

Once these are estimated, we assume volatility to be 0.08% and find the expected return and ETF weights for the portfolio lying on each of the 165 efficient frontiers. Since the efficient frontier is given as a set of discrete points instead of an equation, we use the expected return and weights available for the volatility level nearest to our assumed volatility level. The resulting portfolios are our trained portfolios.

*Predictive Power of Trained Portfolios*

We then apply the weights in the above table to estimate the 2020 expected return for all 165 3-ETF portfolio combinations as a weighted average of the individual ETFs' 2020 expected returns (calculated based on 2020 daily return data).

|  |  |
| --- | --- |
| Chart, scatter chart  Description automatically generated |  |

The above two plots show that the expected return/ rank for 2019 is not a good predictor of the expected return/ rank for 2020 respectively.

**3-sector Portfolios – Analysing Impact of Exposure to LEI, CEI and LAG Indicators on Portfolio Returns**

Using the 3 supervised learning models (Linear, Lasso regressions and Regression tree) used in submission 1, we placed each of the 11 ETFs into the 3 economic indicator buckets (LEI, CEI, LAG). With this, we specify the indicator (LEI, CEI, LAG) and positional (long + / short -) categories for each ETF in each of the 165 portfolios.

Thereafter, 27 unique indicator and positional combinations resulting from the above categories are identified, returns for each of these are aggregated and then ranked for 2019 and 2020. However, this does not reveal much about which indicators tend to perform better.

So, we calculate the weights which each portfolio attributes to LEI, CEI and LAG. This is done by looking at the indicator bucket for each asset in the portfolio, and assigning the weight for that asset to its indicator bucket. If more than one asset in the portfolio are assigned to the same indicator bucket, the weights for those two assets are summed and the result is assigned to the indicator bucket. If no asset is assigned to an indicator bucket, the weight of that bucket in the portfolio is taken to be 0.

We then visualise how the 2019 and 2020 expected returns evolve as the weight allocated to a certain indicator bucket increases.

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |

* Increasing a portfolio's exposure to LEI tended to increase returns over 2019 and decrease returns over 2020
* Increasing a portfolio's exposure to CEI tended to decrease returns over 2019. However, the effect of CEI on 2020 returns appears to be ambiguous upon visual inspection
* Increasing a portfolio's exposure to LAG appears to have a marginal negative impact on 2019 returns and a relatively stronger positive impact on 2020 returns

Chart, line chart

Description automatically generated**3-principal components Portfolio**

Finally, a 3-dimensional dataset using Principal Component Analysis (PCA) is used to estimate 2019 expected return and covariances, which are then fed into the Critical Line Algorithm (CLA) to obtain the efficient frontier (chart alongside) resulting from a portfolio of the 3 PCs.

We then use the 0.8% volatility to get the expected return and weights for the 3-PC portfolio lying on the efficient frontier for this assumed volatility level. The PCA portfolio return of 0.176% is much lower than the average of the 2019 returns across 165 portfolios, which stands at 23.566%.

The historical estimates of expected returns given above are used with the 2019 weights calculated in 7.3 to get the expected return for the 2020 3-PC portfolio. This stands at 3.89%, higher than the expected return for the 2019 3-PC portfolio.

**Conclusion**

In the 2-sector portfolio analysis, the actual correlation was much better than trimmed data correlation. The correlation relationship was been successful been validated comparing with 0, 1 and -1 correlation.

In the 3-sector portfolio analysis, we find that expected return/ rank for 2019 is not a good predictor of the expected return/ rank for 2020 respectively. LEI, CEI and LAG tended to have different effects on returns over 2019 and 2020, so no one indicator clearly stood out. Finally, the return of the PCA portfolio for both 2019 and 2020 was below the average return for the 3-sector portfolio.

**Answer 8.2: Work Split Report**

* Shubham worked on the code and report covering parts 1-4
* Ishaan worked on the code and report covering parts 5-7