Individual Module Assignment 1

Matteo Sammut 0206002L

February 13, 2025

1 Assignment 1

This documentation addresses the written portion of Assignment 1 for the Financial Engineering (ARI5122) submission. This document should be accessed alongside the code portion *A1Q1.ipynb* and *A1Q2.ipynb* for tasks A and B respectively

1.1 Question A

1.2 1

Data for the required tickers was downloaded using the Yahoo finance package and saved to separate csv files.

1.3 2a

After visually plotting the closing prices of these tickers in figure figure 1, several assertions could be made about their stationarity.

1.4 2b

As shown in plot figure 1 the mean of the S&P 500 plot is clearly rising. Taking chunks in different periods will result in different means, compare S&P 500 2015 - 2016 vs 2017 - 2018. It is also irregularly volatile when comparing 2017-2018 vs 2018-2019. This long run of positive values is a clear indication of a non-stationary time series. This also holds true for the FTSE plot which seems to indicate both a gradual increase in mean as well as irregular periods of variance typical of non-stationary time series. For the SPDR plot it is difficult to visually asses if the time series is stationary or not. There does not appear to be a mean trend, with the time series regularly fluctuating around 115. It appears to experience different levels of volatility but this cannot be asserted without appropriate statistical analysis. From the plot alone, it would be classified as stationary.

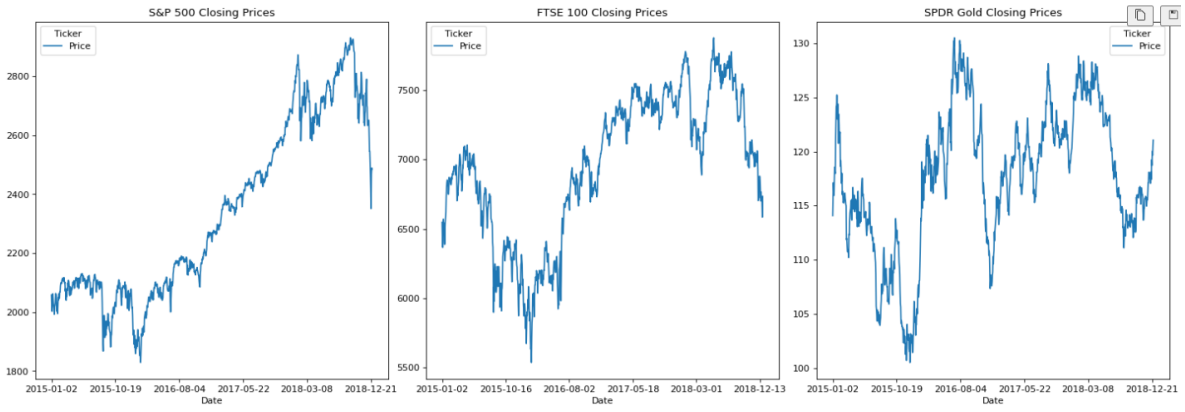

Figure 1: Closing Prices of the 3 tickers

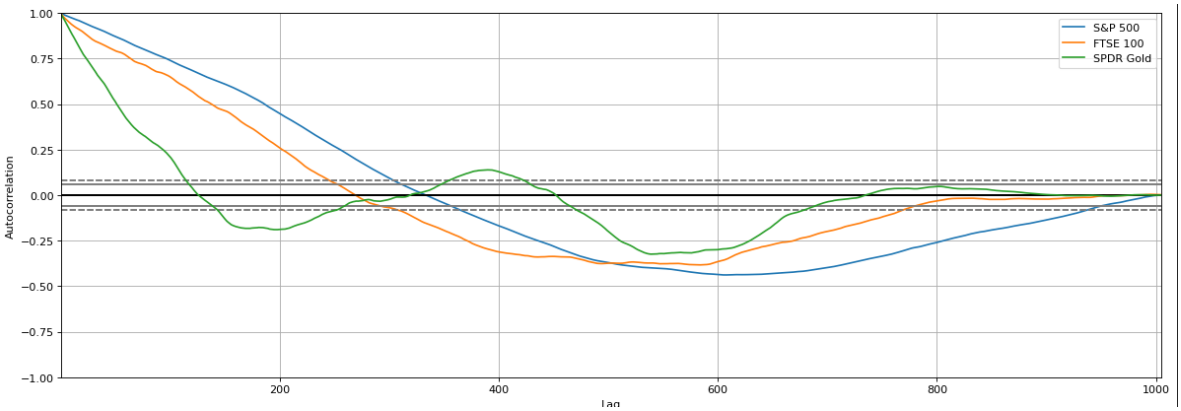

Figure 2: Auto-correlation of the 3 tickers

1.5 2c

	Mean Former	Mean Latter	Variance Former	Variance Latter	p-value	Stationary
SPDR Gold	115.28	119.90	56.13	18.94	0.177	No
FTSE 100	6531.14	7372.18	139505.01	52904.60	0.432	No
S&P 500	2077.20	2596.42	6822.52	33469.52	0.666	No

Table 1: Descriptive Statistics for SPDR Gold, FTSE 100, and S&P 500

The results from the Augmented Dickey-Fuller test were tabulated in table 1. As may be noticed from these calculations there are noticeable mean shifts for FTSE 100 & S&P 500, meanwhile that of the SPDR Gold is less pronounced. When it comes to variance, significant differences in variance are noticeable for all three plots. Once the p-values were calculated and tabulated it became clear that none of the time series are stationary at the 5% significance level because all p-values are greater than 0.05.

1.6 3a

The autocorrelation values of the tickers were plotted in 2 to see if they form a random walk.

1.7 3b

We may notice that SPDR Gold shows the closest behaviour to that of a random walk, as its autocorrelation quickly diminishes and oscillates around zero. The other two have stronger autocorrelation, indicating less randomness and suggesting the presence of trends in price movements.

1.8 4a

Logarithmic returns $\left(\ln\left(\frac{P_t}{P_{t-1}}\right)\right)$ offer several advantages over arithmetic returns $\left(\frac{P_t - P_{t-1}}{P_{t-1}}\right)$, particularly in financial analysis. One key advantage is time additivity, log returns over multiple periods can be summed to yield the total log return, this makes them more convenient for long-term analysis such as portfolio return calculations. While arithmetic returns can also be used for multi-period analysis, the process involves a more complex formula[HG15].

Additionally, prices are often assumed to follow a log-normal distribution, though this may not hold true for every price series. Transforming prices into log returns results in values that approximate a normal distribution, which is beneficial in contexts where normality is assumed. This property is particularly valuable in machine learning and statistical analyses, where normal distributions are commonly used.

While arithmetic returns are often more intuitive for comparing individual price changes, this advantage are limited as these returns tend to exhibit higher volatility compared to logarithmic returns.

1.9 5a

Index	Mean	Variance	Skewness	Kurtosis
S&P 500	0.000188	0.008619	-0.4922	3.9125
FTSE 100	2.8×10^{-5}	0.009004	-0.1746	2.4141
SPDR Gold	5.9×10^{-5}	0.008043	0.2315	2.5807

Table 2: Calculated Moments of the S&P 500, FTSE 100, and SPDR Gold

1.10 5b

For the scope of this question it was important to calculate and utilise the daily log returns. The following equation 1 was used to calculate the log returns. It represents the natural logarithm of the ratio between the adjusted closing price at time t and the adjusted closing price at time t - *1*.

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \tag{1}$$

The first moment of the distribution is the mean, calculated using equation 2. Simply put it is the summation of daily logs divided by the number of daily logs and represents the average of the computed log returns, providing the expected return and indicating the central tendency of the distribution.

The second moment, shown in equation 3, is the standard deviation of the log returns. This measures the volatility of the index or asset, reflecting the degree of dispersion around the mean.

The third moment is the skewness, calculated as shown in equation 4. Skewness measures the asymmetry of the distribution, resulting in a positive or negative distribution with a tail of outliers. It is obtained by summing the cube of the deviations of each log return *Ri* from the mean, dividing by the total number of observations(daily logs), and then standardizing by the cube of the standard deviation.

The fourth moment is the kurtosis, presented in equation 5. It measures the tailedness or shape of the distribution, indicating whether the distribution has heavy or light outliers relative to a normal distribution. Kurtosis is calculated similarly to skewness but with the deviations raised to the fourth power. A kurtosis value close to 3 suggests a distribution similar to the normal distribution.

These metrics have become key metrics in stock risk assessments[LYW06].

$$\text{Mean} = \frac{\sum_{i=1}^{N} R_i}{N} \tag{2}$$

$$\text{Variance} = \sqrt{\frac{\sum_{i=1}^{N}(R_i - \mu)^2}{N}} \tag{3}$$

$$\text{Skewness} = \frac{\sum_{i=1}^{N}\left(\frac{R_i-\mu}{s^3}\right)^3}{N} \tag{4}$$

$$\text{Kurtosis} = \frac{\sum_{i=1}^{N}\left(\frac{R_i-\mu}{s^4}\right)^4}{N} \tag{5}$$

1.11 5c

Index/Asset	Mean	Variance	Skewness	Kurtosis
S&P 500	0.000188	0.008619	-0.4922	3.9125
FTSE 100	2.8e-05	0.009004	-0.1746	2.4141
SPDR Gold	5.9e-05	0.008043	0.2315	2.5807

Table 3: Statistical Metrics for S&P 500, FTSE 100, and SPDR Gold

When examining the results in Table 4, the S&P 500 stands out as an immediate outlier. As previously mentioned, the kurtosis of a normal distribution is 3. In contrast, the S&P 500 has a kurtosis of approximately 3.9, indicating a taller, more peaked distribution—known as leptokurtic. Additionally, the distribution is negatively skewed, suggesting a higher likelihood of extreme negative outliers. This all points towards the stock being especially risky at this time.

In contrast to the S&P 500, the FTSE 100 and SPDR Gold exhibit different distribution characteristics. The FTSE 100 has a kurtosis of approximately 2.41, which is lower than the normal distribution's benchmark of 3. This indicates a flatter (platykurtic) distribution, with thinner tails suggesting fewer extreme outliers. While it is negatively skewed it only exhibits a mild tendency towards negative outliers, and is relatively insignificant.

SPDR Gold performed the best in our analysis for risk and return. While the skewness is larger, it is positive, slightly positive skewness may be seen as less risky because the probability of extreme negative losses is reduced.

1.12 6a

Index	Annualized Return (%)	Annualized Volatility (%)
S&P 500	4.7373	13.6829
FTSE 100	0.7003	14.2938
SPDR Gold	1.4906	12.7682

Table 4: Annualized Return and Volatility for S&P 500, FTSE 100, and SPDR Gold

1.13 6b

Using the same formula above (eq. 2 and 3) we calculate the first and second moment which are the mean and volatility respectively. To annualize these moments we multiply them by the number of trading days in a year, this being 252 as weekends and holidays are excluded. For volatility we multiply it with the square root of trading days as it scales with the square root of time.

After tabulated the results of these calculations we were presented with a nice performance spread. The S&P 500 had the highest annuitization return indicating very strong growth, while SPDR Gold had the lowest volatility indicating that it is a safe-haven asset which is seen as a store of value indicative of gold. The FTSE 100 particularly had the highest volatility, yet delivered the lowest return. This implies poor risk-adjusted performance.

2 Question B

2.1 1a

We got a number of tickers and grouped them by sector, they were chosen out of personal interest in their performance over the COVID-19 period. The following table provides a view of the tickers and under which sectors they were grouped under:

2.2 1b

The data was normalized to account for different scales among stocks, this data was plotted and visualised over three graphs to avoid crowding as shown in figure 3.

2.3 2a

After applying PCA we were given the contribution of each feature to every principle component(PC). These values were plotted into a heat-map for easy analysis. We then plotted the individual and cumulated explained variance to see how much of the variation was explainable by each PC. Since the first three PC's accounted for over 80% of the variation it was opted to use those as the main components.

Ticker	Company Name	Sector
ABT	Abbott Laboratories	Healthcare
JNJ	Johnson & Johnson	Healthcare
DUK	Duke Energy	Energy
XOM	Exxon Mobil	Energy
ABNB	Airbnb	Hospitality
BKNG	Booking Holdings	Hospitality
DAL	Delta Air Lines	Hospitality
WMT	Walmart	Retail
COST	Costco	Retail
BAC	Bank of America	Financials
JPM	JPMorgan Chase	Financials
AAPL	Apple	Technology
MSFT	Microsoft	Technology
AMZN	Amazon	Technology

Table 5: Company Tickers, Names, and Sectors

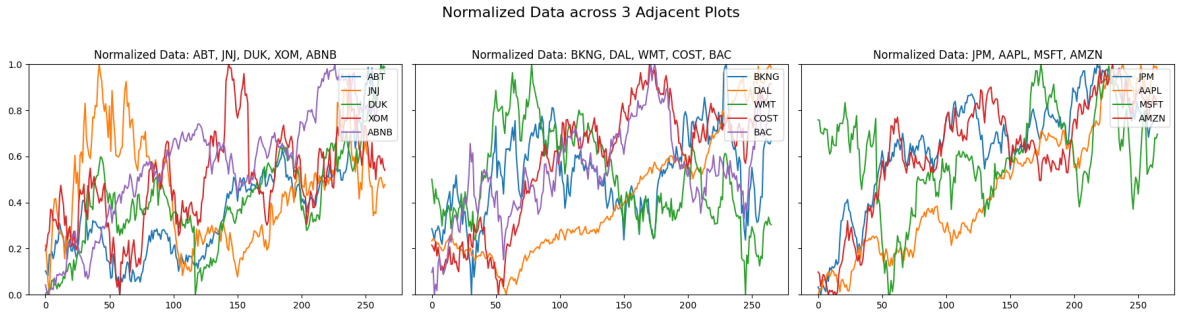

Figure 3: Normalized Data

2.4 2c

After calculating the loading scores and weighing them by the explainability of each PC the following results were obtained and tabulated.

Sectors	Value
Retail	0.075956
Healthcare	0.121799
Energy	0.129565
Financials	0.141560
Technology	0.165539
Hospitality	0.187166

Table 6: Contribution by sector

Stocks	Value
WMT	-0.004947
BAC	0.072599
ABT	0.145895
COST	0.151911
DUK	0.170858
ABNB	0.238276
JNJ	0.243598
AAPL	0.256683
XOM	0.259129
MSFT	0.266964
JPM	0.283120
BKNG	0.364742
AMZN	0.496616
DAL	0.561497

Table 7: Contribution by stock

As was expected the tickers that contributed the most to variance were those in the hospitality sector such as Delta Airlines and Booking, while those in the retail and healthcare sector were resilient. This can be easily attributed to the fact that people still need to consume groceries during a pandemic

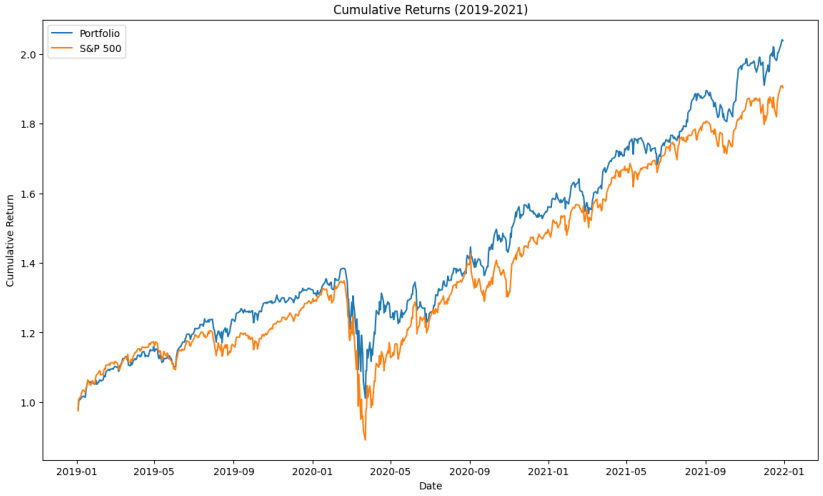

Figure 4: Cumulative Returns of our portfolio and the S&P500

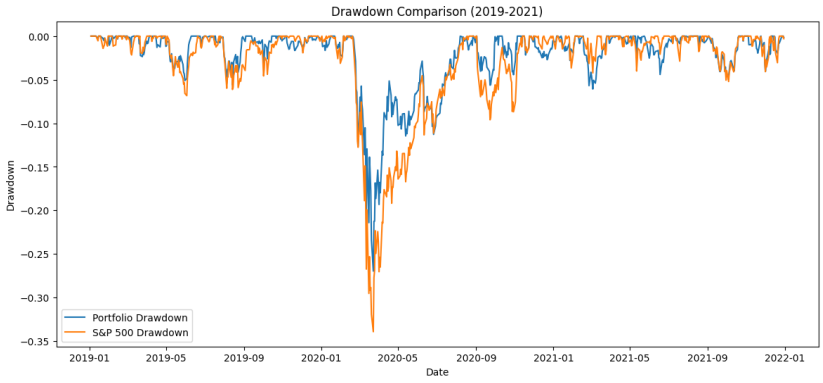

Figure 5: Drawdown Comparison of our portfolio and the S&P500

at similar rates as before, similarly people would still be actively participating in banking.

2.5 3a

We selected the top 5 most resilient stocks from table 7 to construct our portfolio. Therefore ensuring above average and stable performance to the S&P 500 during the COVID 19 period.

2.6 3c

With our resilient performance we were able to receive above average results compared to the S&P 500 which is our baseline. While our portfolio reflected the baseline before the COVID period, during the COVID downturn (2020-01 till 2020-09) it can be observed from figure 4 that despite both experiencing a sharp decline our portfolio recovered outperformed the broader index in the following months with a steep recover. This performance continued for the remainder of the observed period.

This is further illustrated in Figure 5, which depicts the drawdowns, with lower values indicating more significant declines from previous peaks. The sharp drop likely corresponds to the period of widespread lockdowns, during which industries across all sectors faced disruptions. However, the recovery phase reveals that certain sectors rebounded more quickly, particularly those able to resume operations sooner, unlike industries such as hospitality, which faced prolonged restrictions.

References

[HG15] Robert S Hudson and Andros Gregoriou. Calculating and comparing security returns is harder than you think: A comparison between logarithmic and simple returns. *International Review of Financial Analysis*, 38:151–162, 2015.

[LYW06] Kin Keung Lai, Lean Yu, and Shouyang Wang. Mean-variance-skewness-kurtosis-based portfolio optimization. In *First International Multi-Symposiums on Computer and Computational Sciences (IMSCCS'06)*, volume 2, pages 292–297. IEEE, 2006.

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY

Declaration

Plagiarism is defined as "the unacknowledged use, as one's own, of work of another person, whether or not such work has been published, and as may be further elaborated in Faculty or University guidelines" (University Assessment Regulations, 2009, Regulation 39 (b)(i), University of Malta).

I / We*, the undersigned, declare that the [assignment / Assigned Practical Task report / Final Year Project report] submitted is my / our* work, except where acknowledged and referenced.

I / We* understand that the penalties for committing a breach of the regulations include loss of marks; cancellation of examination results; enforced suspension of studies; or expulsion from the degree programme.

Work submitted without this signed declaration will not be corrected, and will be given zero marks.

* Delete as appropriate.

(N. B. If the assignment is meant to be submitted anonymously, please sign this form and submit it to the Departmental Officer separately from the assignment).

Matteo Sammut
Student Name

Signature

Student Name

Signature

Student Name

Signature

Student Name

Signature

ARI5122
Course Code

Assignment 1
Title of work submitted

27/01/2025
Date