Lancaster University Management School

Group No: 10

Department of Accounting and Finance

Group Coursework Cover Sheet

Coursework submitted for assessment will only be marked if the following signed declaration sheet accompanies it.

Assessed Coursework Declaration

I confirm that I have read and understood the University's regulations relating to Plagiarism (as stated in the courses handbook). I declare that this submission is my own work. I have not submitted it elsewhere in substantially the same form towards the award of a degree or other qualification. It has not been written or composed by any other person and all sources have been appropriately referenced and acknowledged. I agree that this work may be interrogated for plagiarism software.

SURNAME	FORENAME	Library Card Number(s)	Signature(s)
Tam	Ka Wai (Kelly)	38774054	
Li	Yin Hang	38860279	
Binti Mohd Nazlan	Irdina	39164748	
Jiao	Rongyu	39051345	Rongyu Jiao
Yeung	Ching Yan	38776022	
Howley	Sean	38866668	

Module Number: ACF324 **Tutor:** Manh Pham

Submission Date. 5 Dec 2024

Students may be asked to submit themselves to an oral examination on the basis of any item of coursework. Students should note that a mark of zero may be given to any coursework that is found to be in breach of university regulations. Any cases of plagiarism may be brought to the attention of the Board of Examiners. Breaches of regulations may be referred to in any references which members of the Department are asked to provide.

All late submissions (without pre-approved extensions) will be awarded a zero mark.

FOR OFFICE USE ONLY

Extension: **Yes / No** Date Received:

Penalty (if applicable): **Extension Date:**

Final Mark % (after deduction of penalty if applicable): 78/100

1a) Q1a) 18/20

1. **Overview**

Financial time series, such as stock prices, exchange rates, and market indices, play a crucial role in understanding the dynamics of financial markets.

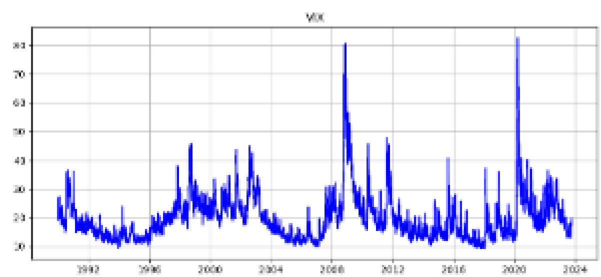
In this report, we mainly focus on Volatility Index (VIX) series and stock index price series (FTSE100). According to CFI (2024) the VIX measure the market expectation based on the near-term volatility of the S&P500 Index. Meanwhile, the FTSE100 includes the 100 largest companies in the London Stock Exchange. It is a benchmark of the United Kingdom stock market. The report below is analysis on the statistical properties of financial time series between the Volatility Index (VIX) and the FTSE100 Index.

We will discuss through different aspect such as stationarity, normality (if there is heavy tail), autocorrelation, partial autocorrelation. By having a better understanding of the key patterns and relationship of the VIX and the FTSE100 Index, allow us to forecast future movement in the financial market.

2.Data Description

2.1 Key characteristics from raw data (FTSE100 and VIX Series)

(figure1.1)



(figure1.2)

Firstly, we can generate the plot of FTSE100 and VIX data through python.

Let's discuss about FTSE100 with the certain important financial events (Forsdick, S.).

In figure 1.1, we can observe the share price of FTSE 100 displays an upward trend from 1992 to 2004.

The first peak is in 2000, with share price reaches £7000. This can be explained by the dot-com bubble which peaked on 10 March 2000.

It then experienced recession and reaches its all time low, share price dropped to £3500 in 2003. This can be explained by the fear of Iraq war in the financial market and In 2006, it reaches its second peak (index price over £6000) due to the strengthening of mining stocks. However, in 2007-2008, there was great recession and the index price dropped below 4000. This rapid price decline can be explained by the failure of the US financial firm Lehman Brothers and the collapse of Northern Rock which is owned by the British government.

From 2008 to 2015, the FTSE100 index price is gradually increasing, showing an expansion in the economy.

In 2016, the FTSE100 index price slightly dropped to £5000, which can be explained by the Brexit.

In 2020, we experienced an economic downturn during Covid19 .

In 2023, the FTSE100 price index reaches over £8000, showing the recovery from Covid 19.

Does FTSE100 series look stationary? How about the VIX series?

Now, we will discuss about VIX series.

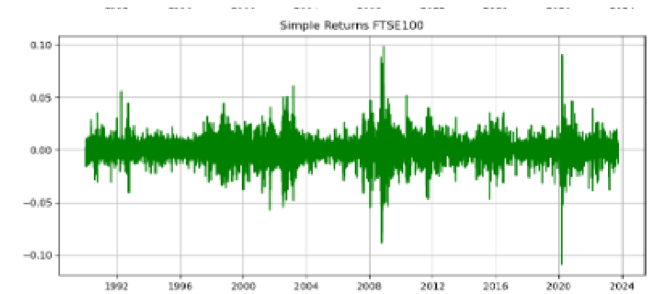
The VIX series (figure1.2) remain stable in general from 1992 to 2006, slightly fluctuates between 1998 to 2003. It reaches two peaks (in 2008 and around 2020 to 2021) during the financial crisis and Covid 19 pandemic.
There is a negative relationship between the FTSE100 index price and the vix. The high vix indicates the investors' uncertainy in market expectation. They were more cautious, decided to pull their market out of the market into less risky assets eg. Bonds and cash.

2.1 Key characteristics of transformed data

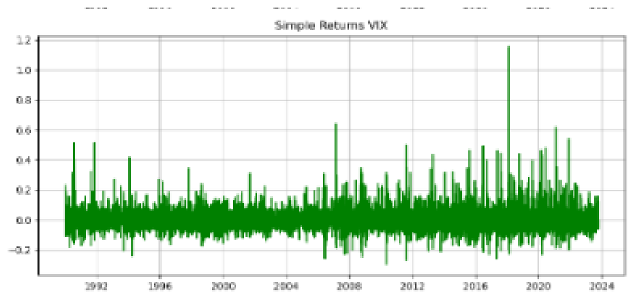
We also investigate the simple returns and log returns of FTSE100 and VIX.

- ***2.1.1 Simple returns***

The simple returns of both indices tend to be less fluctuate compared to the time series (figure1.1 and 1.2) . It has a mean of zero overtime although some large spikes still occur during the volatility periods in 2008 and 2020. Volatilty clustering? -1



(figure1.3)

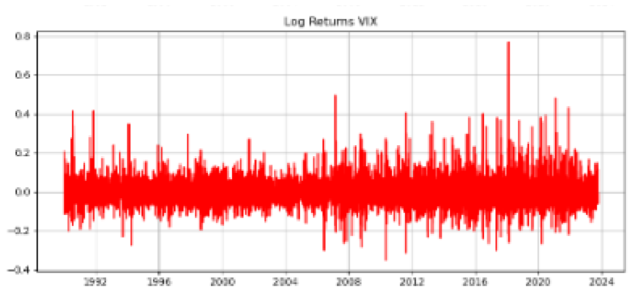


(figure1.4)

- ***2.1.2 Log returns***

Moreover, log turns of both indices exhibit similar behaviour and pattern to simple returns however with slightly smoother variations.

(figure 1.5)



(figure1.6)

3.Statistical properties and analysis

In this session, we will discuss the statistical properties of both series and their transformations.

3.1 Descriptive Data

Descriptive data

```
from statsmodels.tsa.stattools import acf
descriptive_stats = data[['FTSE100', 'VIX', 'simple_returns_FTSE100',
                 'log_returns_FTSE100', 'std_log_returns_FTSE100',
                 'abs_log_returns_FTSE100', 'squared_log_returns_FTSE100']].describe()
skewness = data[['FTSE100', 'VIX', 'simple_returns_FTSE100',
                 'log_returns_FTSE100', 'std_log_returns_FTSE100',
                 'abs_log_returns_FTSE100', 'squared_log_returns_FTSE100']].skew()
kurtosis = data[['FTSE100', 'VIX', 'simple_returns_FTSE100',
                 'log_returns_FTSE100', 'std_log_returns_FTSE100',
                 'abs_log_returns_FTSE100', 'squared_log_returns_FTSE100']].kurtosis() + 3
descriptive_stats.loc['skew'] = skewness
descriptive_stats.loc['kurtosis'] = kurtosis

for columns in ['FTSE100', 'VIX', 'simple_returns_FTSE100',
                 'log_returns_FTSE100', 'std_log_returns_FTSE100',
                 'abs_log_returns_FTSE100', 'squared_log_returns_FTSE100']:
    acf_values = acf(data[columns].dropna(), nlags=1, fft=False)
    lag_1_acf = acf_values[1]
    descriptive_stats.loc['acf_lag_1'] = lag_1_acf
descriptive_stats.round(3)
```

	FTSE100	VIX	simple_returns_FTSE100	log_returns_FTSE100	std_log_returns_FTSE100	abs_log_returns_FTSE100	squared_log_returns_FTSE100
count	8807.000	8807.000	8807.000	8807.000	8807.000	8807.000	8807.000
mean	5369.992	19.594	0.000	0.000	0.012	0.007	0.000
std	1567.466	7.901	0.011	0.011	1.000	0.008	0.000
min	1990.200	9.140	-0.109	-0.115	-10.650	0.000	0.000
25%	4200.050	13.880	-0.005	-0.005	-0.460	0.002	0.000
50%	5756.850	17.800	0.000	0.000	0.007	0.005	0.000
75%	6572.730	23.015	0.006	0.006	0.516	0.010	0.000
max	8014.300	82.690	0.098	0.094	8.682	0.115	0.013
skew	-0.482	2.131	-0.132	-0.290	-0.290	3.036	14.488
kurtosis	2.151	11.180	10.733	10.884	10.884	21.188	345.058
acf_lag_1	0.213	0.213	0.213	0.213	0.213	0.213	0.213

(figure1.7)

The descriptive data table (figure 1.7) shows several characteristics of the FTSE100 index, VIX and their returns. Firstly, the mean represents the average value of the index during the observed period, in which FTSE100 has a mean of 5369.99 while the VIX has a mean of 19.594, indicating that FTSE100 has a higher central tendency. Regarding the standard deviation, it measures the degree of dispersion of the mean, reflecting volatility. The standard deviation of FTSE100 is 1567.466, meaning significant variability in index levels, whereas the VIX has a standard deviation of 7.901, implying less fluctuation.

The skewness measures the asymmetry of the distribution. The FTSE 100 index, simple returns, and log return of the FTSE 100 index show negative skewness (-0.482), indicating a longer tail on the left side, while VIX is positively skewed (2.131), implying a longer tail on the right. As for the kurtosis, it measures the tailedness of the distribution. The case of FTSE 100 has a kurtosis value of 2.151 (close to 3), indicating a relatively moderate tail. However, VIX has a kurtosis value of 11.180, meaning the data have heavier tails and a higher probability of extreme outcomes than a normal distribution (i.e. leptokurtic distribution).

We also perform the Jarque-Bera Test to test whether the series is normally distributed. The results suggest that none of the series follows a normally distributed. This conclusion is supported by the kurtosis values, in which none of the kurtosis values equal 3 (the normal distribution kurtosis value is 3) from the table.

3.2 Stationarity

3.2.1 Observation from plot

In figure1.1, the FTSE100 index series has an upward trend. This is a non-stationarity time series as it violates one the properties in stationarity-constant mean.

3.2.2 Augmented Dickey-Fuller (ADF) Test

To assess stationarity, we perform Augmented Dickey-Fuller (ADF) Test on two time series and their corresponding simple returns and log returns. The ADF test evaluates whether a series is non-stationary by testing the null hypothesis H0 : the presence of a unit root (non-stationarity). If the p-value is less than the critical value, we reject the null hypothesis of non-stationarity. On the other hand, if the p-value is greater than the critical value, we fail to reject the null hypothesis.

In the ADF test on FTSE100, the p-value is 0.38429 which is greater than 0.05, therefore we fail to reject the null hypothesis of non-stationarity. This confirms that the FTSE 100 time series is non-stationary.

In the ADF test on VIX, p-value is 0 which is less than 0.01. Hence, we fail to reject the null hypothesis. It aligns with the nature of the volatility index as it fluctuates around a mean without a clear trend. It is stationary

In the ADF test on simple returns and log returns of FTSE 100, both p-values are 0 which are less than 0.01,
we fail to reject the null hypothesis. Both of them are stationary because they are derived as differences and remove the trends.

In conclusion, except FTSE 100, all series reject the null hypothesis (H0), indicating that they are stationary.

```
Performing ADF Test on FTSE100
ADF Statistic: -1.792290
p-value: 0.384296
Critical Values:
        1%: -3.431094419181841
        5%: -2.861868983995307
        10%: -2.5669451131986794
```

(figure1.8)

```
Performing ADF Test on VIX
ADF Statistic: -6.889152
p-value: 0.000000
Critical Values:
        1%: -3.431093657275364 5
        5%: -2.861868647325225
        10%: -2.566944933987006
```

(figure1.9)

```
Performing ADF Test on simple_returns_FTSE100
ADF Statistic: -23.432864
p-value: 0.000000
Critical Values:
        1%: -3.431094334448478
        5%: -2.861868946553461
        10%: -2.5669450932681404
```

(figure1.10)

```
Performing ADF Test on log_returns_FTSE100
ADF Statistic: -23.245985
p-value: 0.000000
Critical Values:
        1%: -3.431094334448478
        5%: -2.861868946553461
        10%: -2.5669450932681404
```

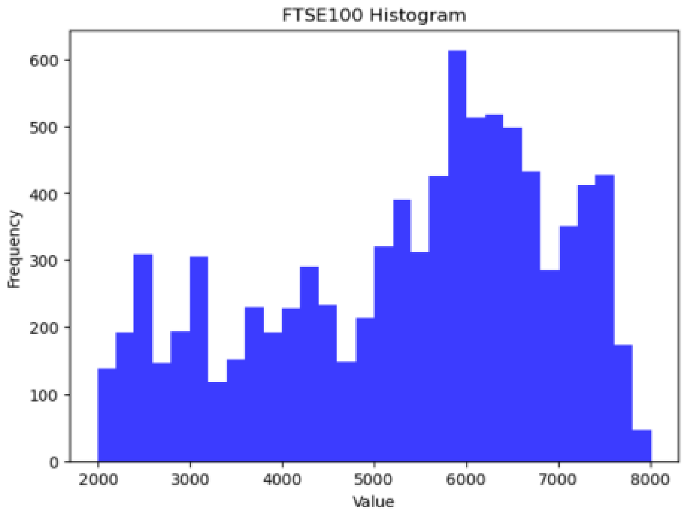
(figure1.11)

3.3 Normality Test

From session 3.1, we conclude that none of them is normally distributed and both log return of FTSE100 and vix series are stationary.

Here we will further test the normality through Jarque Bera Test.
H0: The series is normally distributed (Null hypothesis)
H1: The series is not normally distributed (Alternative hypothesis)

If the JB statistic is high and the p-value < 0.05, we can reject the null hypothesis and determine that the data does not follow a normal distribution.

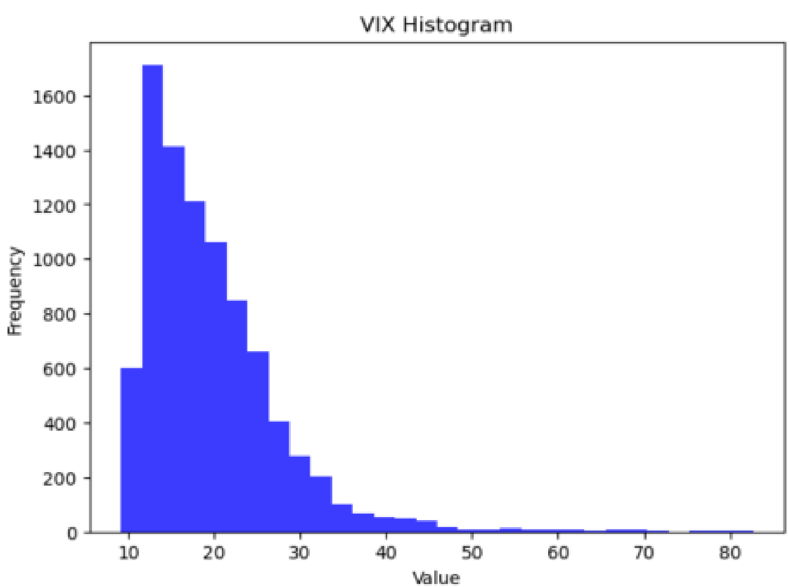


(figure1.12)

From figure1.12, Jarque-Bera normality test for FTSE100: JB stat = 604.844, p-value = 0.000 and therefore it is not normally distributed

Since the JB stat is large and the p-value is less than 0.05, we can reject the null and conclude that the series is not normally distributed. ✓

In figure1.12, the distribution of the FTSE 100 appears slightly right-skewed, with some peaks at around 3000-4000, 5000-6000, 7000-7500

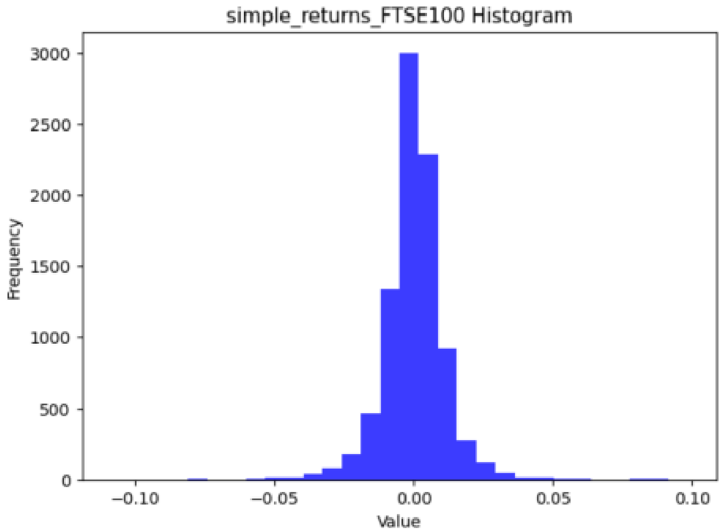


(figure1.13)

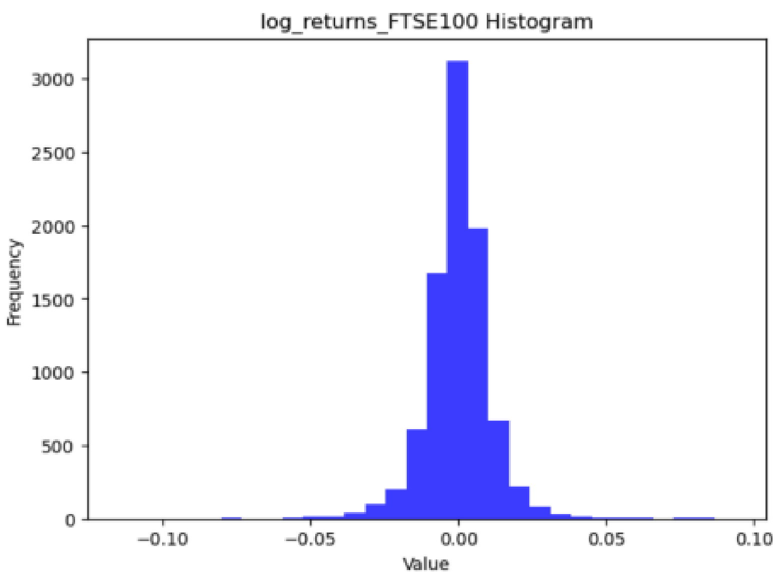
Jarque-Bera normality test for VIX: JB stat = 31187.405, p-value = 0.000

Again, Since the JB stat is large and the p-value is less than 0.05, we can reject the null and conclude that the series is not normally distributed. ✓

In figure 1.13, the distribution of VIX is highly skewed to the right, with a long tail extending toward higher VIX values, indicating occasional extreme volatility in the market.



(figure1.14)



(figure1.15)

Jarque-Bera normality test for log_returns_FTSE100: JB stat = 22904.908, p-value = 0.000

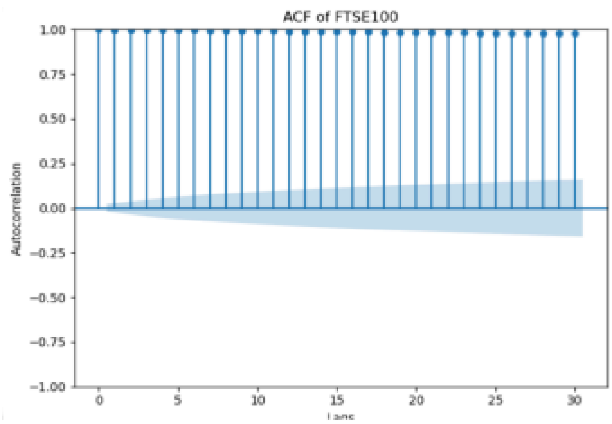
JB stat is huge again and p-value is less than 0.05, hence we reject the null and conclude that the log returns of FTSE100 is not normally distributed.

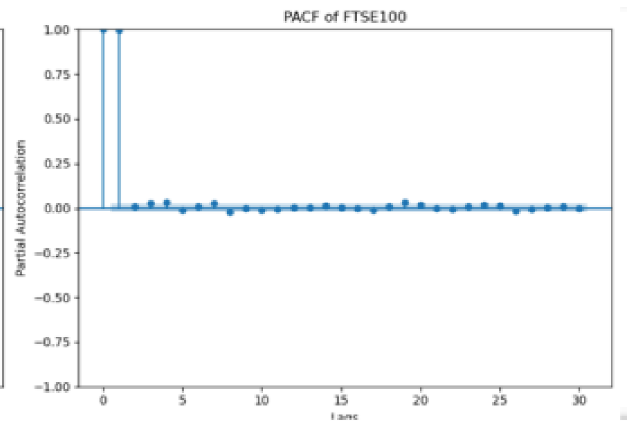
The simple returns (figure1.14)and log returns of FTSE 100 (figure1.15) histogram show that both exhibit a sharp peak at around 0 and fat tails compared to a normal distribution. This highlights the existence of extreme values during volatile periods.

By performing the Jarque-Bera Test to test the series whether they are normally distributed or not. Hence the results suggest that none of the series follow a normally distributed. This conclusion is supported by the kurtosis values, none of the kurtosis values equal 3 (normal distribution kurtosis value is 3) from the table.

3.4 Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF)

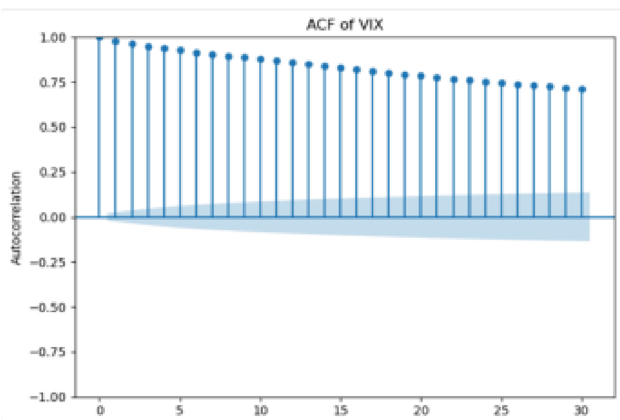
ACF measures the correlation between a time series and its lagged value across various time lags, including direct and indirect correlations. Meanwhile, PACF measures the correlation between a time series and its lagged value after removing the effect of intermediate lags, indicating only a direct correlation. Therefore, it is better to use ACF for the overall correlation structure and PACF for identifying direct relationships.

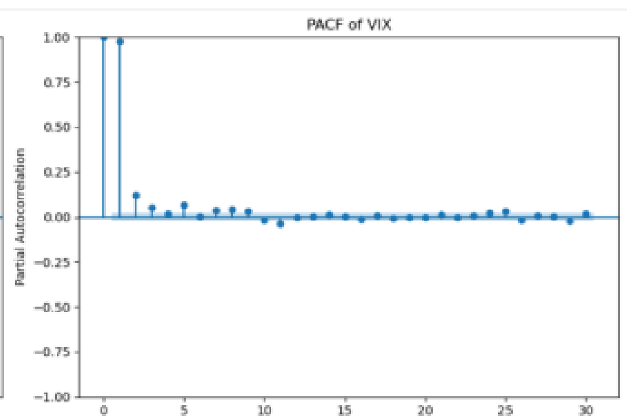




(figure1.16) (figure1.17)

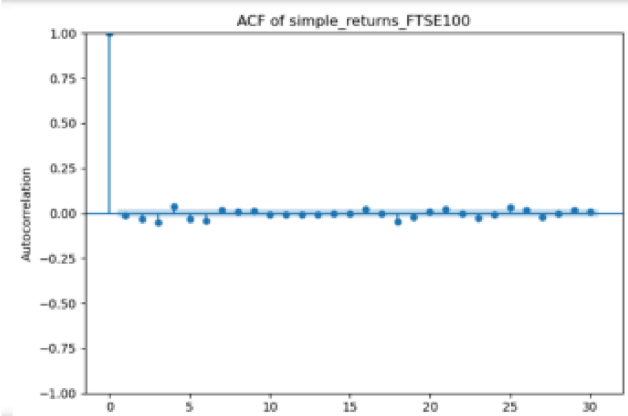
For the FTSE 100, the ACF graph (figure1.16) shows high and persistent correlations over many lags slowly tapering toward zero. This indicates a strong dependence level which is a non-stationary time series. The PACF graph (figure1.17) shows that it spikes at lag 1 only in the first 30 lags. This may show that much of the dependent data in the ACF can be explained by the first lag.



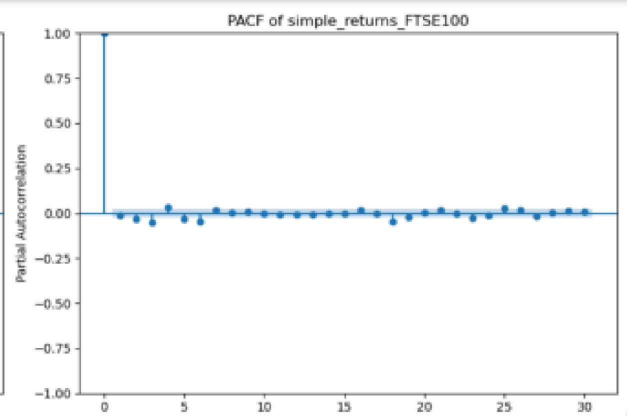


(figure1.18) (figure1.19)

For the VIX, the ACF graph (figure1.18) follows a similar pattern to FTSE 100 level, but slightly faster in tapering toward zero. This indicates a strong dependence level and proves that it is non-stationary. The PACF (figure1.19) also displays significant spikes at lag 1 in the first 30 lags.

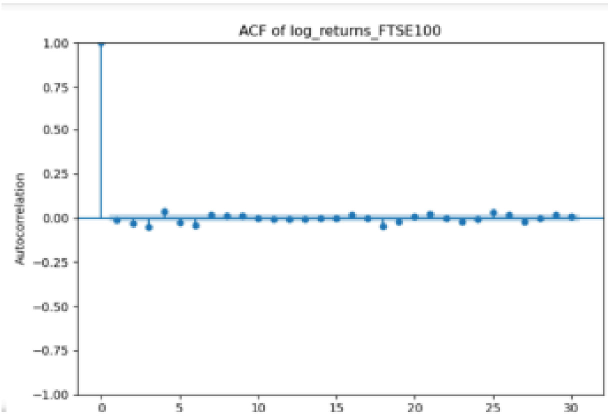


(figure1.20)

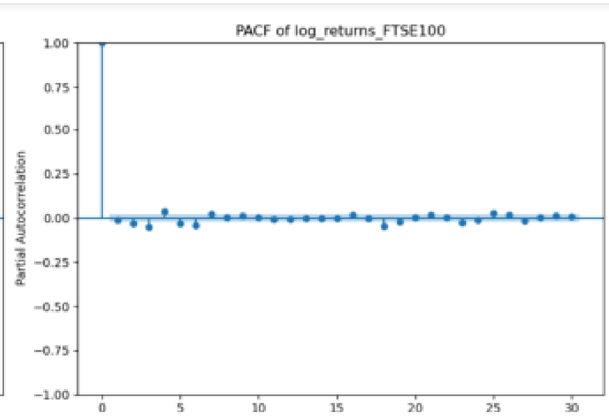


(figure1.21)

For the simple returns of FTSE 100, the ACF (figure1.20) of each time series drops to near-zero after lag 0. The PACF (figure1.21) has no significant spikes except lag 0. This indicates that the series of simple returns of FTSE 100 is weakly dependent or white noise.



(figure1.22)



(figure1.23)

For the log returns of FTSE 100, the ACF graph (figure1.22) is similar to simple returns ACF plots (figure1.20). The ACF for log return FTSE100 (figure1.22) is near zero except lag 0 and PACF(figure1.23) has no significant spikes again beyond lag 0.

For the FTSE 100 and VIX series, AR(1) model could be appropriate as AR(1) model can capture the most important lagged relationship. For the simple returns and log return plots, the graphs show that the series are weakly dependent, it is stationary thus white noise model is enough. In the next session, we will use Ljung-Box Test for further investigation on correlation.

3.5 Test on Correlation (Ljung-Box Test)

Ljung-Box Test

Testing for the weak white noise property of $\hat{\epsilon}_t$ by applying the Ljung-Box test to the residuals.

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_k = 0$$

$$H_A : \rho_j \neq 0 \quad \text{for some} \quad j \in \{1, \ldots, k\}$$

Rejecting H_0 rejects that the residuals are (i.i.d) white noise.

If cannot reject the H_0, further tests are necessary to verify whether the residuals are white noise.

```
Ljung-Box Test Results for FTSE100:
        lb_stat  lb_pvalue
1    8792.76958        0.0
2   17569.59035        0.0
3   26331.32077        0.0
4   35078.91655        0.0
5   43811.92646        0.0
6   52530.53201        0.0
7   61235.59408        0.0
8   69926.54340        0.0
9   78603.21872        0.0
10  87265.25568        0.0
```

(figure1.24)

```
Ljung-Box Test Results for VIX:
        lb_stat  lb_pvalue
1    8445.26864        0.0
2   16624.59130        0.0
3   24570.27422        0.0
4   32294.20064        0.0
5   39841.83714        0.0
6   47206.42816        0.0
7   54413.23686        0.0
8   61484.10982        0.0
9   68429.97446        0.0
10  75235.16976        0.0
```

(figure1.25)

```
Ljung-Box Test Results for simple_returns_FTSE100:
     lb_stat  lb_pvalue
1    1.66946    0.19633
2   11.88637    0.00262
3   36.53279    0.00000
4   47.70693    0.00000
5   55.90672    0.00000
6   72.21160    0.00000
7   75.07786    0.00000
8   75.95157    0.00000
9   76.95780    0.00000
10  77.19594    0.00000
```

(figure1.26)

```
Ljung-Box Test Results for log_returns_FTSE100:
     lb_stat  lb_pvalue
1    1.55917    0.21179
2   11.00996    0.00407
3   33.65636    0.00000
4   45.81164    0.00000
5   53.38927    0.00000
6   69.24522    0.00000
7   72.63179    0.00000
8   73.60806    0.00000
9   74.82104    0.00000
10  74.99765    0.00000
```

(figure1.27)

This table presents the Ljung-Box test results for FTSE 100, VIX, simple returns and log returns of FTSE 100.

The Ljung-Box test examines whether there is significant autocorrelation within the time series at various lags.

From the table results for FTSE 100 (figure1.24) here, we can observe that all the p-values are 0.0. This is indicating that strong evidence of autocorrelation at all lags tested which aligns with the notion that financial time series are generally non-stationary.

Moreover, the results for VIX (figure1.25), it indicates it has similar results to FTSE 100 in which all p-values are 0.0, showing significant autocorrelation at all lags.

Furthermore, the results for simple returns for FTSE 100 (figure1.26) shows that there are no significant spikes against autocorrelation. At lag 1, the p-value is 0.19633 but from lag 2 onward, p-values drop below 0.05. This may present potential predictability in certain market conditions.

In addition, the result for log returns from FTSE 100(figure1.26) is similar to the results for simple returns for FTSE 100. At lag 1, the p-value is 0.21179. From lag 2 onward, significant autocorrelation is observed as all the p-values are 0.0.

	FTSE100	VIX	simple_returns_FTSE100	log_returns_FTSE100
FTSE100	1.000000	-0.071518	0.010815	0.011508
VIX	-0.071518	1.000000	-0.095045	-0.102138
simple_returns_FTSE100	0.010815	-0.095045	1.000000	0.999858
log_returns_FTSE100	0.011508	-0.102138	0.999858	1.000000
std_log_returns_FTSE100	0.011508	-0.102138	0.999858	1.000000
abs_log_returns_FTSE100	-0.052153	0.469986	-0.047202	-0.061197
squared_log_returns_FTSE100	-0.041539	0.425864	-0.068111	-0.084932

(figure1.28)

This table shows the correlation between different series, providing insights into the relationships among them.

Some key observations:

The correlation of FTSE 100 and VIX is -0.071518 which indicates a weakly and negatively correlated. This aligns with the market expectations: VIX (volatility index) often increases during market downturns (FTSE 100 indices)

The correlation of simple returns and log returns for FTSE 100 is 0.999858 which is very close to perfect correlation (perfect = 1). This high correlation reflects a similarity between simple returns and log returns because they are both derived from price changes over time.

How about ARCH effects? -1

4.Conclusion

This report examines the statistical properties of financial time series, focusing on the FTSE 100 and VIX, as well as their simple and log returns. Through various statistical tests and visualizations, we analyzed their stationarity, normality, autocorrelation, and relationships. The results of the Augmented Dickey-Fuller Test (ADF) test confirm that the simple and log returns of the FTSE 100 show characteristics closer to stationarity. Additionally, the Ljung-Box tests indicates that there are no significant autocorrelation in the residuals, supporting that there are no serial correlation. Furthermore, the weak and negative correlation between FTSE 100 and VIX aligns with market expectations during turbulent periods. Overall, this analysis highlights the unique statistical properties of financial time series, including non-stationarity, tail distribution, and deviations from normality.

1b) Q1b) 20/25

1. Introduction

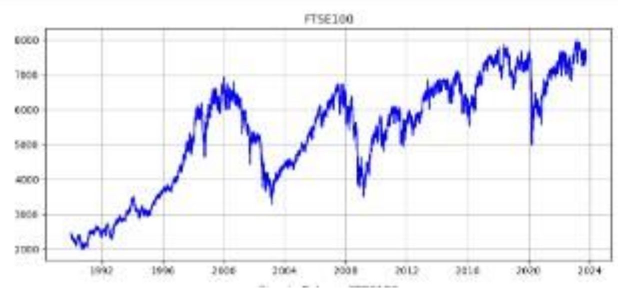
The aim of this report is to find the best model in the class of ARMA (p,q) GARCH (p ∗ , q∗) models (with p, q, p ∗ and q ∗ no greater than 3) for each series based on the appropriate (stationary) transformations of the VIX series and the FTSE100 series data up to 31st December 2021.

In this report, we use of ARMA and GARCH model to find the linear dependencies and the changes in volitility of the data, as the ARMA model only capture linear relationship and autocorrelation of the stationary time series. However, it is important to include volatility clustering as well as it observe investors behaviour in a certain period of time and improve forecasting and portfolio optimization (**Abdollahi, H., Junttila, J.-P. and Lehkonen, H., 2024)**as well as observing investors behaviour (**Suthar, J., 2023)**

2.Data selection

Very good! However, this was already done in Q1a so you could skip this ADF tests here.

Before further analysis, we have to make sure that the time series are stationary, which means the statistical properties (mean, variance, autocorrelation) are constant overtime.



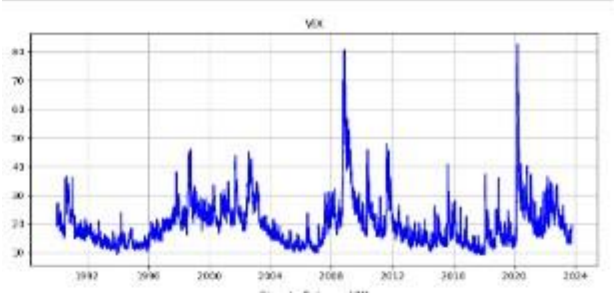
(figure1.2.1)

The plot of FTSE100 has a strong positive trend with strong seasonality. Therefore, it is non-stationary and cannot be used for forecasting.

To solve the problem above, we can transform the FTSE100 data through logarithm transformation.

(figure1.2.2)

As shown in figure 1.2.2, the mean and variance level are now constant overtime without trends and strong seasonality. It is stationary and suitable for forecasting.



(figure1.2.3)

The Vix series is stationary, as the above vix series plot (figure1.2.3) does not have any trends and seasonality, the mean, variance remain constant over time.

To verify, we can conduct the Augmented Dickey-Fuller (ADF) Test. We set the null hypothesis = if failed to be rejected, the time series has a unit root (non-stationary) , alternative hypothesis = when null hypothesis is rejected, the time series does not a unit root (stationary). The rejection criteria is when p value is below the respective critical value, we reject the null hypothesis.

```
Performing ADF Test on FTSE100
ADF Statistic: -1.792290
p-value: 0.384296
Critical Values:
        1%: -3.431094419181841
        5%: -2.861868983995307
        10%: -2.5669451131986794
```

(figure1.2.4)

```
Performing ADF Test on log_returns_FTSE100
ADF Statistic: -23.245985
p-value: 0.000000
Critical Values:
        1%: -3.431094334448478
        5%: -2.861868946553461
        10%: -2.566945093268140
```

(figure1.2.5)

```
Performing ADF Test on VIX
ADF Statistic: -6.889152
p-value: 0.000000
Critical Values:
        1%: -3.4310936572753645
        5%: -2.861868647325225
        10%: -2.566944933987006
```

(figure1.2.6)

As shown above, the p-value of log return of FTSE100 (figure1.2.5)and vix (figure1.2.6) is both 0, lower than the significance level (0.05). Therefore we reject the null hypothesis and conclude they are stationary.
Meanwhile, the p value of FTSE100 is 0.384, higher than the significance level, failed to reject null hypothesis and it is non-stationary.

According to **Arif, A. (2024)**, it is important to forecast using stationary data as statistical data are more meaningful and well- defined when stationary, leading to more accurate and reliable predictions. As explained above, therefore we will use the log return of FTSE100 and Vix series for our modelling below.

-2

3.Model Selection

Before going to perform different model estimation, checks and tests, you should've summarized your modelling selection strategy first - what you were going to do - to give readers an overall picture/roadmap of your analysis! Think of Box-Jenkins analysis for model selection.

Selected model:

FTSE100

ARMA(2,3)-GARCH(1,1)

VIX

How did you come up with these models? Are these best-fitted models and why?

ARMA(3,3)-GARCH(1,1)

3.1 ARMA model

- The ARMA model captures the linear relationship between the time series.

3.1.1 Through ACF and PACF plots

To find the potential model, we can observe the spikes of ACF and PACF plots. ACF plots
In ACF plots, with x axis indicates the number of lags and y axis indicates the autocorrelation. ACF shows the coefficients of correlation between the time series and the lag.

PACF plots

Meanwhile, the PACF plots with x axis labeled with the number of lags and y axis labelled as partial correlation. PACF shows the partial correlation coefficient between the time series and the lag

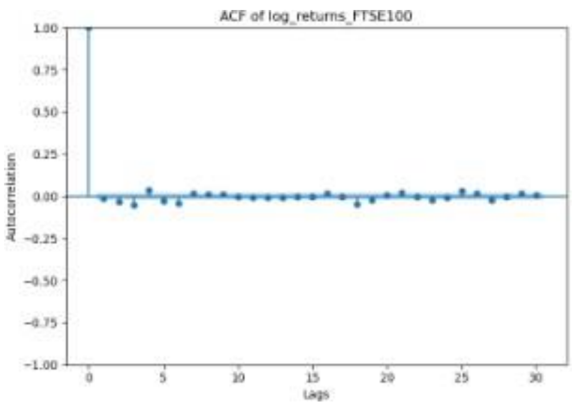
- **FTSE100 log returns result**

In the plot of ACF log returns FTSE100 (figure1.2.7), the spike at lag 0 indicates the autocorrelation =1, it is perfectly correlated

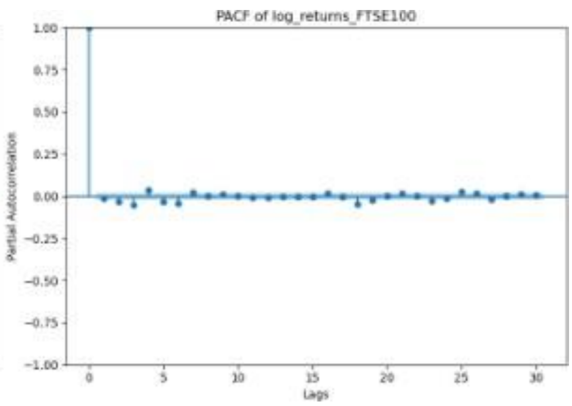
In the PACF of log returns FTSE100 (figure1.2.8) the spike at lag 0 indicates the partial autocorrelation is also 1 and perfectly correlated.

- **VIX series result**

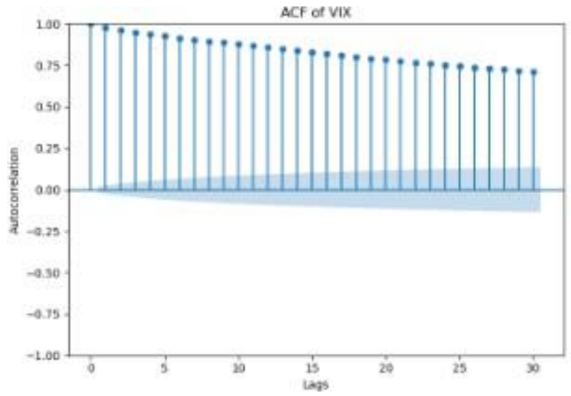
The VIX series, in the PACF vix plot (figure1.2.10) is a sharp cutoff at lag=1 while the ACF vix plot (figure1.2.9)significant spikes at high lags. In figure1.2.9, we can observe the spikes from lags 0 to 30. It shows that it is highly persistent and near to 1, decaying at the same time. Therefore it is stationary.



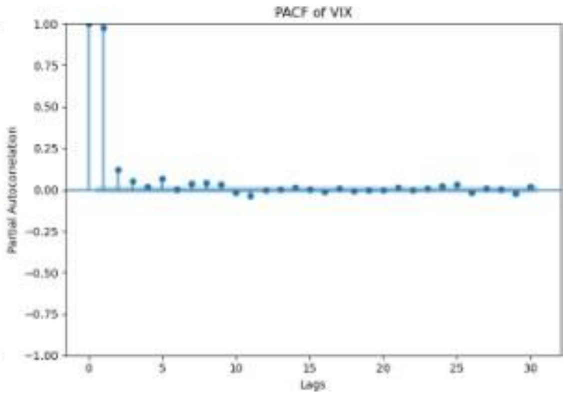
(figure1.2.7)



(figure1.2.8)



(figure1.2.9)



(figure1.2.10)

However, PACF and ACF may not be the best in finding best model in this case. We can find the best model through performing automated model selection technique in Python. Automated model selection provide more information such as the AIC and BIC. ✓

In this report, we will focus on AIC.

- Akaike information criterion (AIC)

$$\mathrm{AIC} = 2k - 2\ln(L)$$

K= the number of estimated parameters in the model ✓

L=maximum likelihood of the model

The lowest AIC find the best model which without overfitting. AIC penalizes the models with more parameters and simpler models that captures the underlying data structure is preferred (the best fit model).

3.1.2 Using Python

Step1: Estimate the mean equation $y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t$ (ARMA component) using the ARIMA() from statsmodels in Python then extract the residuals

Step2: Estimate the variance equation (GARCH component)

$$\text{Variance Equation:} \quad \sigma_t^2 = \omega + \sum_{i=1}^{p^*} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{q^*} \beta_i \sigma_{t-i}^2$$

by fitting the residuals with a zero-mean GARCH model using the arch library

- **ARMA model in log return of FTSE100**

  First we run the ARMA model for log return of FTSE100, (with p and q no greater than 3), and the lowest aic = best fit model (figure1.2.11)

```
from statsmodels.tsa.arima.model import ARIMA
import warnings
from statsmodels.tools.sm_exceptions import ValueWarning

#For log return of FTSE100:
lowest_aic = float('inf')
best_ARMA_FTSE = None

for p in range(1,4):
    for q in range(1,4):
        try:
            warnings.simplefilter('ignore', ValueWarning)
            model = ARIMA(data['log_returns_FTSE100'] * 100, order=(p, 0, q))
            fit_model = model.fit()
            if fit_model.aic < lowest_aic:
                lowest_aic = fit_model.aic
                best_ARMA_FTSE = fit_model
        except:
            continue

print(best_ARMA_FTSE.summary())
```

(figure1.2.11)

```
                               SARIMAX Results
==============================================================================
Dep. Variable:     log_returns_FTSE100   No. Observations:                 8347
Model:                  ARIMA(2, 0, 3)   Log Likelihood              -12530.219
Date:                 Wed, 27 Nov 2024   AIC                          25074.438
Time:                         17:56:59   BIC                          25123.646
Sample:                     01-04-1990   HQIC                         25091.245
                          - 12-31-2021
Covariance Type:                   opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0132      0.012      1.139      0.255      -0.009       0.036
ar.L1         -0.6755      0.055    -12.346      0.000      -0.783      -0.568
ar.L2         -0.4978      0.057     -8.779      0.000      -0.609      -0.387
ma.L1          0.6632      0.055     12.010      0.000       0.555       0.771
ma.L2          0.4588      0.057      7.994      0.000       0.346       0.571
ma.L3         -0.0799      0.007    -12.272      0.000      -0.093      -0.067
sigma2         1.1787      0.009    136.500      0.000       1.162       1.196
===================================================================================
Ljung-Box (L1) (Q):                   0.00   Jarque-Bera (JB):             20974.06
Prob(Q):                              0.99   Prob(JB):                         0.00
Heteroskedasticity (H):               1.21   Skew:                            -0.34
Prob(H) (two-sided):                  0.00   Kurtosis:                        10.74
===================================================================================
```
(figure1.2.12)

The result shows that ARMA (2,3) has the lowest AIC (25074.438) and therefore it is the best model.

Let' s further anaylse the log return of FTSE100 SARIMAX results (figure1.2.12). It has a Ljung-Box Test (L1) (Q) = 0.00 with prob(Q) =0.99, the high p value indicates that there is no significant autcorrelation in the residuals at lag 1. In terms of normality, it has a JB statistic of 20974.06 with Prob (JB)=0.00. The low p value indicates that the residuals are not normally distributed. It has some skewness (-0.42) which is insignificant and a high kurtosis (10.74) indicates the residuals' distribution is leptokurtic (kurtosis >3.0), fat tails with outliers and the heteroskedasticity (1.21) indicates that the variance of the error terms are not constant in the observation.

```
ARCH Test for VIX: LM stat = 1632.902, p-value = 0.000
ARCH Test for log_returns_FTSE100: LM stat = 1572.399, p-value = 0.000
```

(figure1.2.13)

By conducting the ARCH test for FTSE100 log return (figure1.2.13),(H0=no autoregressive conditional heteroskedasticity, H1=there is autoregressive conditional heteroskedasticity). FTSE 100 log return p value =0 less than significance level (0.05), reject null hypothesis. And the high LM stat (1572.399) we can confirm that there is a large arch effect remain in FTSE100 log return and therefore GARCH test is used to consider the peak values and nullifies the effect of kurtosis.

- GARCH model in log return of FTSE100

By running the code,

Here, you kept the mean equation as ZeroMean. It is better to use the best fitted ARMA model found above.

```
best_aic_garch = float('inf')
best_garch_model_ftse = None

for p_star in range(1, 4):
    for q_star in range(1, 4):
        try:
            garch_model = arch_model(data['log_returns_FTSE100'] * 100, mean="Ze
            garch_result = garch_model.fit(disp="off")
            if garch_result.aic < best_aic_garch:
                best_aic_garch = garch_result.aic
                best_garch_model_ftse = garch_result
                best_order = (p_star,q_star)
        except:
            continue

print(best_garch_model_ftse.summary())
print('Best GARCH order for Log return FTSE 100:',best_order)
```

(figure1.2.14)

We can get the following result

```
                        Zero Mean - GARCH Model Results
==============================================================================
Dep. Variable:      log_returns_FTSE100   R-squared:                       0.000
Mean Model:                   Zero Mean   Adj. R-squared:                  0.000
Vol Model:                        GARCH   Log-Likelihood:               -11084.1
Distribution:                    Normal   AIC:                           22174.2
Method:              Maximum Likelihood   BIC:                           22195.3
                                          No. Observations:                 8347
Date:                  Wed, Nov 27 2024   Df Residuals:                     8347
Time:                          18:17:54   Df Model:                            0
                              Volatility Model
===========================================================================
                 coef    std err          t      P>|t|       95.0% Conf. Int.
---------------------------------------------------------------------------
omega          0.0173  3.865e-03      4.472  7.744e-06 [9.710e-03,2.486e-02]
alpha[1]       0.0909  1.118e-02      8.128  4.351e-16   [6.897e-02,  0.113]
beta[1]        0.8933  1.299e-02     68.775      0.000     [  0.868,  0.919]
===========================================================================

Covariance estimator: robust
Best GARCH order for Log return FTSE 100: (1, 1)
```

(figure1.2.15)

Let’ s further analyse the result, by running the code with the lowest AIC as the best GARCH model, we conclude that the Best GARCH order for log return FTSE100 is GARCH(1,1).

Combining ARMA and GARCH best model we get ARMA(2,3)-GARCH(1,1)

How did you interpret the estimated output of the best ARMA-GARCH model? e.g. the mean equation dynamics, and the persistence in the variance equation

- ARMA model in VIX

Repeat the same procedure for vix series

-1

```
lowest_aic = float('inf')
best_ARMA_VIX = None

for p in range(1,4):
    for q in range(1,4):
        try:
            warnings.simplefilter('ignore', ValueWarning)
            model = ARIMA(data['VIX'] * 100, order=(p, 0, q))
            fit_model = model.fit()
            if fit_model.aic < lowest_aic:
                lowest_aic = fit_model.aic
                best_ARMA_VIX = fit_model
        except:
            continue

print(best_ARMA_VIX.summary())
```

(figure1.2.16)

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                    VIX   No. Observations:                 8347
Model:                 ARIMA(3, 0, 3)   Log Likelihood              -54196.036
Date:                Wed, 27 Nov 2024   AIC                         108408.071
Time:                        17:57:24   BIC                         108464.308
Sample:                    01-04-1990   HQIC                        108427.279
                         - 12-31-2021
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const       1945.9869    160.367     12.135      0.000    1631.674    2260.300
ar.L1          0.8229      0.075     10.903      0.000       0.675       0.971
ar.L2          0.8006      0.079     10.074      0.000       0.645       0.956
ar.L3         -0.6290      0.037    -16.823      0.000      -0.702      -0.556
ma.L1          0.0281      0.075      0.373      0.709      -0.120       0.176
ma.L2         -0.7176      0.037    -19.561      0.000      -0.790      -0.646
ma.L3          0.0607      0.009      6.457      0.000       0.042       0.079
sigma2      2.554e+04    121.885    209.506      0.000    2.53e+04    2.58e+04
===================================================================================
Ljung-Box (L1) (Q):                   0.01   Jarque-Bera (JB):            325603.28
Prob(Q):                              0.91   Prob(JB):                         0.00
Heteroskedasticity (H):               2.44   Skew:                             2.43
Prob(H) (two-sided):                  0.00   Kurtosis:                        33.21
-----------------------------------------------------------------------------------
```

(figure1.2.17)

Let’ s further anaylse the vix SARIMAX results (figure1.2.16). It has a Ljung-Box Test (L1) (Q) = 0.01 with prob(Q) =0.91, the high p value indicates that there is no significant aut0correlation. In terms of normality, it has a JB statistic of 325603.28 with Prob (JB)=0.00. The low p value indicates that the residuals are not normally distributed. It has some skewness (2.43), positively skewed and a high kurtosis (33.21) indicates the residuals' distribution is leptokurtic (kurtosis >3.0), fat tails with outliers and the heteroskedasticity (2.44) indicates that the variance of the error terms are not constant in the observation.

By conducting the ARCH test for vix (figure1.2.13),(H0=no autoregressive conditional heteroskedasticity, H1=there is autoregressive conditional heteroskedasticity). p value =0 less than significance level (0.05), reject null hypothesis. And the high LM stat (1632.902) we can confirm that there is a large arch effect remain in vix and therefore GARCH test is used to consider the peak values and nullifies the effect of kurtosis.

```
                       Zero Mean - GARCH Model Results
==============================================================================
Dep. Variable:      log_returns_FTSE100   R-squared:                       0.000
Mean Model:                   Zero Mean   Adj. R-squared:                  0.000
Vol Model:                        GARCH   Log-Likelihood:                -11084.1
Distribution:                    Normal   AIC:                           22174.2
Method:              Maximum Likelihood   BIC:                           22195.3
                                          No. Observations:                 8347
Date:                  Wed, Nov 27 2024   Df Residuals:                     8347
Time:                          18:17:54   Df Model:                            0
                              Volatility Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
omega          0.0173  3.865e-03      4.472  7.744e-06 [9.710e-03,2.486e-02]
alpha[1]       0.0909  1.118e-02      8.128  4.351e-16   [6.897e-02,  0.113]
beta[1]        0.8933  1.299e-02     68.775      0.000     [  0.868,  0.919]
============================================================================

Covariance estimator: robust
Best GARCH order for Log return FTSE 100: (1, 1)
```

Should be the equation for VIX here!

(figure1.2.18)

Repeat the same step as FTSE100 log return in figure (1.2.14) by running the code with the lowest AIC as the best GARCH model, the Best GARCH order for vix is GARCH(1,1).
Combine the ARMA and GARCH model, the best model for vix is ARMA(3,3)-GARCH(1,1)

We then perform a post-GARCH ARCH test to check the remaining heteroskedasticity in the residuals of the model and get the following result.

```
Post-GARCH ARCH Test for FTSE100: LM stat = 42.436, p-value = 0.000
Post-GARCH ARCH Test for VIX: LM stat = 10.768, p-value = 0.376
```

The post-GARCH ARCH test is performed by regressing the squared residuals from the FTSE100 model against the lagged values of the residuals and check if there is autocorrelation in the squared errors and any remaining arch effect. The aim of this test is to find the conditional volatility model has capture volatility clustering. If there is a significant ARCH effect remain, further modelling is needed.

In FTSE100, the LM Statistic is 42.236 and a low p value (p-value=0) , which indicates that after fitting the GARCH , there is a large arch effect remain in the residual. It suggests that the current GARCH test has not capture all the volatility clustering in log return FTSE100 and further modeling such as EGARCH/ TGARCH is required. Meanwhile, the vix series has a high p-value (0.376) which indicates that it has no large arch effect in the residuals and all the volatility clustering is captured. The current GARCH model is good enough for vix.

3.2Residual check

3.2.1 Jarque Bera Test

Last but not least, let' s perform a Jarque-Bera Test to check the normality of residuals.

Jarque-Bera Test

- Formula

$$J\,B = \frac{n}{6}\left(S^2 + \frac{(\,K - 3\,)^2}{4} \right)$$

- hypothesis
    - H0=The data is normally distributed
    - H1 =The data does not follow a normal distribution

Compute the result

```
Jarque-Bera Test for FTSE100:
JB Statistic: 113800.344; p-value: 0.000
Jarque-Bera Test for VIX:
JB Statistic: 778.638; p-value: 0.000
```

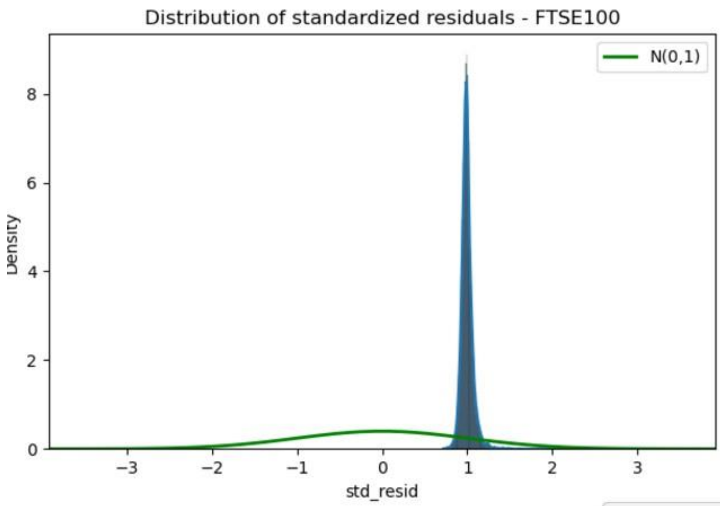
(figure1.2.19)

For both FTSE100 and vix series, the p value is 0, less than significance level 0.05. Therefore reject the null hypothesis and therefore the residuals of FTSE100 and vix are both not normally distributed.

The high JB statistic, FTSE100 113600.344 and VIX 778.638 indicates that the deviation in the normal distribution is high with skewness.

What are your recommendations given this result? -1

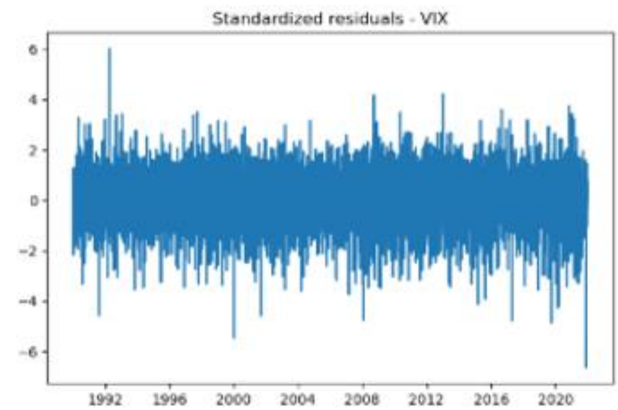
3.2.2 Standardized residuals



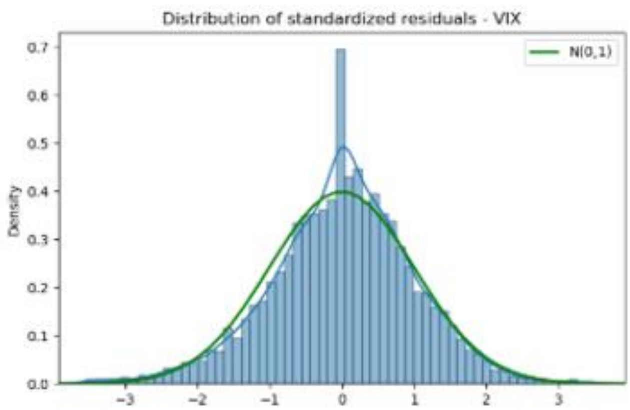
Something went wrong with the standardized residuals for FTSE100 returns here. The histogram centers at 1, not 0 so what you were plotting here didn't seem to be the standardized residuals from ARMA-GARCH which should have a zero mean by construction.

-1

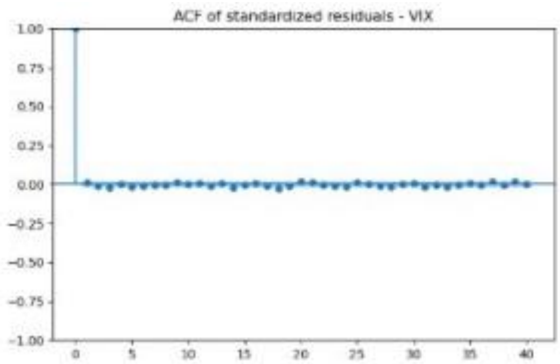
(figure1.2.20)



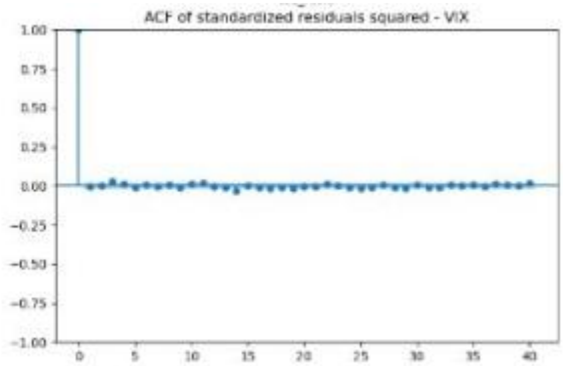
(figure1.2.21)



(figure1.2.22)



(figure1.2.23)



(figure1.2.24)

In figure1.2.20, the ftse100 distribution in standard residual is slightly negatively skewed which means there is some arch effect left, some residuals are still correlated.
In figure1.2.21, standardized residuals for vix, the plot is plotted time against volatility. there is no trend, seasonarity nor patterns. As there is no pattern shown and lags of acf and pacf of standard residuals in vix =0 (figure1.2.23 and 1.2.24) indicates there is no autocorrelation remain in vix series. We can conclude that all volatility clustering has been captured and therefore ARMA(3,3)-GARCH(1,1) is the best model for vix

Granger-causality andpulse Response Analyses?

4. Conclusion

Based on the above modeling selection strategy, we concluded that in FTSE100, with the transformation of log return FTSE100 log return, ARMA(2,3)-GARCH(1,1) at the moment and further modeling is required. For vix, we concluded that ARMA(3,3)-GARCH(1,1) has captured all the volatility clustering and there is no autocorrelation between residuals. Therefore, it is the best fit model.

1c) Q1c.i) 8/10

AcF 324_Question1_Part C i

From question b we know that the best ARMA(p, q)-GARCH(p*, q*) model for log return of FTSE is ARMA(3, 3)-GARCH(1, 1), which for VIX is ARMA(2, 3)-GARCH(1, 1). For simplicity, we create a competing model ARMA(1,1)-GARCH(1,1) for both series.

You switched the best models between FTSE100 and VIX here. Will take results as given for subsequent analysis! -1

We denote the ARMA(3, 3)-GARCH(1, 1) model as FTSE_Model_1, ARMA(2, 3)-GARCH(1, 1) as VIX_Model_1, the ARMA(1,1)-GARCH(1,1) models as FTSE_Model_2 and VIX_Model_2 respectedly.

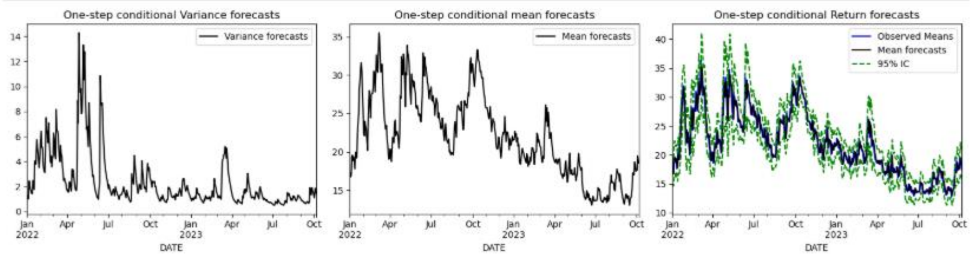
<u>Forecast For VIX under VIX_Model_1</u>

- The one step forecast is as follows:

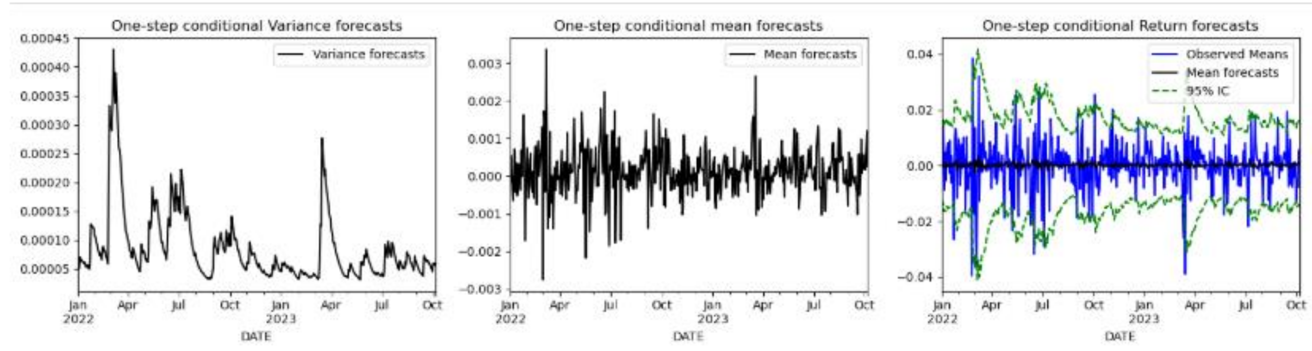
1]:

DATE	true_value	mean_forecast	var_forecast
2022-01-03	16.600000	17.309878	1.360314
2022-01-04	16.910000	16.808006	1.253979
2022-01-05	19.730000	16.979126	1.074127
2022-01-06	19.610001	19.400322	2.396719
2022-01-07	18.760000	19.441405	1.985265
...	...	...	...
2023-10-02	17.610001	17.499009	1.282227
2023-10-03	19.780001	17.586466	1.096861
2023-10-04	18.580000	19.451544	1.880498
2023-10-05	18.490000	18.563765	1.715336
2023-10-06	17.450001	18.512487	1.438391

460 rows × 3 columns



Forecast For VIX under VIX_Model_2



This didn't look like the forecasts for VIX but for FTSE100 returns!

-1

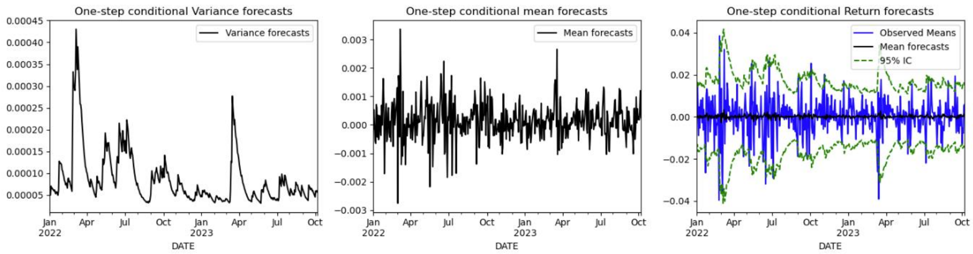
DATE	true_value	mean_forecast	var_forecast
2022-01-03	0.000000	0.000128	0.000055
2022-01-04	0.016213	0.000128	0.000051
2022-01-05	0.001558	0.000128	0.000073
2022-01-06	-0.008886	0.000128	0.000067
2022-01-07	0.004673	0.000128	0.000069
...	...	...	...
2023-10-02	-0.012885	0.000128	0.000046
2023-10-03	-0.005407	0.000128	0.000060
2023-10-04	-0.007754	0.000128	0.000058
2023-10-05	0.005248	0.000128	0.000060
2023-10-06	0.005767	0.000128	0.000058

```
For model 1:
 rmspe: 1.4512752710496108 mapr: 1.035343547917168 mdape: 0.7845532673732487
For model 2:
 rmspe: 1.4562066619501404 mapr: 1.0371200659970314 mdape: 0.7788967803281786
```

From the plot above, we can see that under both models, most of the predictions fall within the 95% conditional interval, thus the mean forecast is relatively reliable. However, both models have shown a prediction that is very similar to the true value, suggesting a possibility of overfitting, thus reacting too much to the noise.

From the evaluation, we can see that the ARMA(2,3))-GARCH(1, 1) model has slightly better performance in Mean Absolute Percentage Error, suggesting it may provide more reliable forecasts in the average case. However, the ARMA(1,1)-GARCH(1,1) outperforms in the other two evaluations, indicating that it can better handle extreme values and outliers. Overall, the two models have relatively small differences in all criteria, thus having relatively similar performance.

Forecast For Log Return of FTSE100 under FTSE_Model_1



[71]:

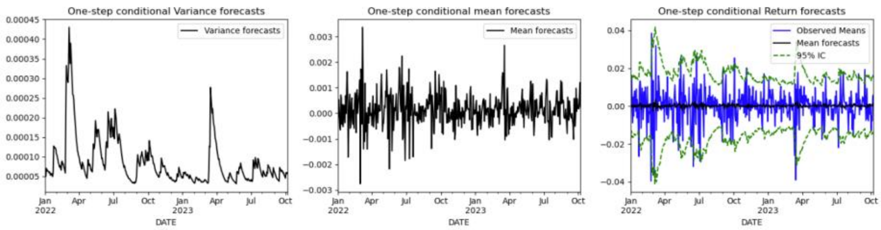
DATE	true_value	mean_forecast	var_forecast
2022-01-03	0.000000	-0.000145	0.000054
2022-01-04	0.016213	0.000532	0.000049
2022-01-05	0.001558	0.000009	0.000070
2022-01-06	-0.008886	-0.000523	0.000064
2022-01-07	0.004673	-0.000654	0.000066
...	...	...	...
2023-10-02	-0.012885	0.000277	0.000045
2023-10-03	-0.005407	0.000171	0.000059
2023-10-04	-0.007754	0.000576	0.000057
2023-10-05	0.005248	0.001197	0.000060
2023-10-06	0.005767	0.000252	0.000056

460 rows × 3 columns

Forecast For Log Return of FTSE100 under FTSE_Model_2

DATE	true_value	mean_forecast	var_forecast
2022-01-03	0.000000	0.000128	0.000055
2022-01-04	0.016213	0.000128	0.000051
2022-01-05	0.001558	0.000128	0.000073
2022-01-06	-0.008886	0.000128	0.000067
2022-01-07	0.004673	0.000128	0.000069
...	...	...	...
2023-10-02	-0.012885	0.000128	0.000046
2023-10-03	-0.005407	0.000128	0.000060
2023-10-04	-0.007754	0.000128	0.000058
2023-10-05	0.005248	0.000128	0.000060
2023-10-06	0.005767	0.000128	0.000058

460 rows × 3 columns




```
For FTSE model 1:
 rmspe: 0.009046879703738353 mapr: 0.00621673172611388 mdape: 0.004203905799376285
For FTSE model 2:
 rmspe: 0.009045849401274828 mapr: 0.006243526943390957 mdape: 0.004101848610735665
```

From the plot above, we can see that under both models, most of the predictions fall within the 95% conditional interval, thus the mean forecast is relatively reliable. Both the conditional mean forecast and conditional variance forecast have obtained extreme values around April in both 2022 and 2023. This might suggest that the market is more volatile during the period.

From the evaluation, we can see that the two models have very similar performance regarding RMSPE, while ARMA(3, 3)-GARCH(1, 1) outperforms in MAPR and the ARMA(1,1)-GARCH(1,1) outperforms in MDAPE. This also indicates that the ARMA(3, 3)-GARCH(1, 1) has better reliable forecast performance and ARMA(1,1)-GARCH(1,1) better handles extreme values. Overall, the two models have relatively small differences in all criteria and, thus, have relatively similar performance, similar to what was discovered regarding VIX.

Q1c.ii) 7/10

c)ii) When comparing the best model ARMA(3,3)-GARCH(1,1) for FTSE and ARMA(2,3)-GARCH(1,1) for VIX to the competing model for both ARMA(1,1)-GARCH(1,1) in terms of predictive ability we first calculate the forecasts for the out of sample period 01/01/2022 – 06/10/2023 ✓

DATE	true_value	mean_forecast	var_forecast
1990-01-04	19.220000	NaN	NaN
1990-01-05	20.110000	NaN	NaN
1990-01-08	20.260000	NaN	NaN
1990-01-09	22.200000	NaN	NaN
1990-01-10	22.440000	NaN	NaN
...	...	...	...
2023-10-02	17.610001	17.499009	1.282227
2023-10-03	19.780001	17.586466	1.096861
2023-10-04	18.580000	19.451544	1.880498
2023-10-05	18.490000	18.563765	1.715336
2023-10-06	17.450001	18.512487	1.438391

ARMA (3,3)

Should be ARMA-GARCH here

DATE	true_value	mean_forecast	var_forecast
1990-01-04	-0.004923	NaN	NaN
1990-01-05	-0.002900	NaN	NaN
1990-01-08	-0.005415	NaN	NaN
1990-01-09	0.002054	NaN	NaN
1990-01-10	-0.009775	NaN	NaN
...	...	...	...
2023-10-02	-0.012885	0.000277	0.000045
2023-10-03	-0.005407	0.000171	0.000059
2023-10-04	-0.007754	0.000576	0.000057
2023-10-05	0.005248	0.001197	0.000060
2023-10-06	0.005767	0.000252	0.000056

✓

ARMA(2,3)

After getting this data we then calculate the error for each model.

Loss differential = error term1 – error term2

✓

We create a null and alternative hypothesis

H0 – if the loss differential = 0 both nodels have equal predictive accuracy

H1 – loss differential doesn't = 0

Using this we can complete a DM test

	Test	Statistic	P-value
0	Sign Test	0.839254	0.401327
1	DM Test	0.910206	0.362714

	Test	Statistic	P-value
0	Sign Test	-0.466252	0.641035
1	DM Test	0.341198	0.732954

Which table is for FTSE100, which for VIX?

For comparing to a t table I am using a significance level of 0.05 and because this is a two tailed test it is halved

As the p values for both exceeds 0.05 we reject the null hypothesis that the models have the same predictive ability. As the null hypothesis was rejected this suggests that the models do not significantly in their ability to predict conditional variances. ARMA(1,1)-GARCH(1,1) is the better better model at conditional variances and the extreme values

Here you performed DM tests for conditional mean forecasts? How about conditional variance forecasts?

Can you perform DM tests for conditional variance forecasts? Why?

-3

1d) Q1d) 10/15

```
                                SARIMAX Results
==============================================================================
Dep. Variable:     log_returns_FTSE100   No. Observations:                 8347
Model:                  ARIMA(2, 0, 3)   Log Likelihood               25904.254
Date:                 Sun, 01 Dec 2024   AIC                         -51794.509
Time:                         15:45:30   BIC                         -51745.301
Sample:                     01-04-1990   HQIC                        -51777.701
                          - 12-31-2021
Covariance Type:                   opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0001      0.000      1.121      0.262   -9.63e-05       0.000
ar.L1         -0.3783      0.084     -4.500      0.000      -0.543      -0.214
ar.L2         -0.2565      0.087     -2.935      0.003      -0.428      -0.085
ma.L1          0.3649      0.084      4.362      0.000       0.201       0.529
ma.L2          0.2206      0.087      2.522      0.012       0.049       0.392
ma.L3         -0.0712      0.006    -11.837      0.000      -0.083      -0.059
sigma2         0.0001   8.53e-07    138.218      0.000       0.000       0.000
===================================================================================
Ljung-Box (L1) (Q):                   0.00   Jarque-Bera (JB):             21808.64
Prob(Q):                              0.96   Prob(JB):                         0.00
Heteroskedasticity (H):               1.21   Skew:                            -0.36
Prob(H) (two-sided):                  0.00   Kurtosis:                        10.89
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
                       Zero Mean - GARCH Model Results
==============================================================================
Dep. Variable:              Residuals   R-squared:                       0.000
Mean Model:                 Zero Mean   Adj. R-squared:                  0.000
Vol Model:                      GARCH   Log-Likelihood:                27351.9
Distribution:                  Normal   AIC:                          -54697.9
Method:            Maximum Likelihood   BIC:                          -54676.8
                                        No. Observations:                 8347
Date:                Sun, Dec 01 2024   Df Residuals:                     8347
Time:                        15:45:30   Df Model:                            0
                              Volatility Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
omega      2.3531e-06  8.331e-12  2.824e+05      0.000 [2.353e-06,2.353e-06]
alpha[1]       0.1000  1.881e-04    531.562      0.000   [9.962e-02,  0.100]
beta[1]        0.8776  2.374e-03    369.651      0.000     [  0.873,  0.882]
============================================================================

Covariance estimator: robust
```

Table 4.1

```
                               SARIMAX Results
==============================================================================
Dep. Variable:     log_returns_FTSE100   No. Observations:                 8347
Model:                ARIMA(1, 0, 1)   Log Likelihood               25878.447
Date:               Mon, 02 Dec 2024   AIC                         -51748.894
Time:                       01:56:35   BIC                         -51720.776
Sample:                   01-04-1990   HQIC                        -51739.290
                        - 12-31-2021
Covariance Type:                 opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0001      0.000      1.054      0.292      -0.000       0.000
ar.L1      -1.645e-06      0.003     -0.001      1.000      -0.006       0.006
ma.L1      -1.646e-06      0.003     -0.001      1.000      -0.006       0.006
sigma2         0.0001   8.29e-07    143.064      0.000       0.000       0.000
===================================================================================
Ljung-Box (L1) (Q):                   1.53   Jarque-Bera (JB):             22074.61
Prob(Q):                              0.22   Prob(JB):                         0.00
Heteroskedasticity (H):               1.21   Skew:                            -0.28
Prob(H) (two-sided):                  0.00   Kurtosis:                        10.95
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
[2] Covariance matrix is singular or near-singular, with condition number 2.48e+17. Standard errors may be unstable.
                       Zero Mean - GARCH Model Results
==============================================================================
Dep. Variable:              Residuals   R-squared:                       0.000
Mean Model:                 Zero Mean   Adj. R-squared:                  0.000
Vol Model:                      GARCH   Log-Likelihood:                27356.1
Distribution:                  Normal   AIC:                          -54706.2
Method:            Maximum Likelihood   BIC:                          -54685.1
                                        No. Observations:                 8347
Date:                Mon, Dec 02 2024   Df Residuals:                     8347
Time:                        01:56:35   Df Model:                            0
                              Volatility Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
omega      2.3749e-06  1.938e-13  1.225e+07      0.000 [2.375e-06,2.375e-06]
alpha[1]       0.1000  8.272e-06  1.209e+04      0.000   [9.998e-02,  0.100]
beta[1]        0.8800  2.146e-03    409.992      0.000     [  0.876,  0.884]
============================================================================

Covariance estimator: robust
```

Table 4.2

In this section, we are going to compare the ARMA (3,3) and ARMA (1,1) models. Firstly, we are going to compare by analysing its AIC. For ARMA (3,3) we have an AIC value of –54697.9 whilst ARMA (1,1) with –54706.2. The lower the AIC value, the better the model is, which indicates that ARMA (1,1) is much preferred in this experiment. That being said, we cannot just depend and make a decision on which model is better based on the AIC model solely. Hence, more investigation will be examined below to compare and determine the best ARMA model here.

To perform the trading strategy, we first merged the FTSE 100 price and the forecasted return generated by the two different models. Subsequently, by looping through every row, we base our buying or selling on the forecasted return on the forecast for the next day and integrate the outcomes to form a dataset. In comparison, we assume there are no transaction costs, and the transactions have no market impact. The datasets are as follows.

DATE	FTSE100_price	forecasted_return	Cash Balance	Share Holdings	Holding Value	Portfolio Value
2022-01-03	7384.500000	-0.000145	4001.500000	13.0	95998.500000	100000.000000
2022-01-04	7505.200195	0.000532	4001.500000	13.0	97567.602535	101569.102535
2022-01-05	7516.899902	0.000009	101721.198726	0.0	0.000000	101721.198726
2022-01-06	7450.399902	-0.000523	101721.198726	0.0	0.000000	101721.198726
2022-01-07	7485.299805	-0.000654	4412.301261	13.0	97308.897465	101721.198726
...	...	...	...	...	...	...
2023-10-02	7510.700195	0.000277	2391.528248	13.0	97639.102535	100030.630783
2023-10-03	7470.200195	0.000171	2391.528248	13.0	97112.602535	99504.130783
2023-10-04	7412.500000	0.000576	2391.528248	13.0	96362.500000	98754.028248
2023-10-05	7451.500000	0.001197	2391.528248	13.0	96869.500000	99261.028248
2023-10-06	7494.600098	0.000252	2391.528248	13.0	97429.801274	99821.329522

460 rows × 6 columns

Table 4.3 for ARMA (3,3)

Based on table 4.3 above, we can identify the forecasted return for certain days in January 2022 and October 2023 based on the ARMA (3,3) model. In January 2022, we can imply that there was a fluctuation trend of forecasted returns. On the 3rd of January, it can be seen that the value was –0.000145. The asset return might experience a minor decrease in the amount, which suggests that it should be sold. This could prevent them from bearing a potential loss.

Moving on to the next two days, we can examine the forecasted return portrays a positive value of 0.000532 and 0.000009 which indicates that there was an increase in the asset value, potentially because of positive news of the assets in the market. Therefore, it was a good time to purchase and invest in the assets. On the 4th of October, we were still not able to buy the assets since our cash balance was lesser than the holding value of the assets. Meanwhile, on the 5th, the positive value was extremely small. Even so, we managed to purchase the assets due to the fact that our cash balance is more than the holding value.

Furthermore, on the 6th and 7th of January, it can be analysed that the values were back to negatives, of –0.000523 and –0.000654 respectively. However, as the holding value on the 6th is 0, we were not able to sell the assets, but it was possible for us to sell the assets on the 7th since the holding value is more than the cash balance.

On the other hand, in October 2023, the forecasted return values were all positive, with a daily fluctuation in the magnitudes. Firstly, the return for 2nd October, with the amount 0.000277 indicates that the assets should be stable and buying the assets was a good option. Following that, the value on 3rd October declined to 0.000171, which is still a good deal to buy, but needs to be cautious with the slight difference of magnitude.

In contrast, the value increased on the 4th, with the amount of 0.000576, implying a buying signal and a good deal for investment. Even better, the value on 5th October rose to 0.001197, which shows a significant climb from the day before. This shows a solid and favorable event impacting the asset on the day, which suggests that the assets should definitely be purchased. Following that, the value on the 6th of October diminished to 0.000252, inclining that it was still a buy signal, but should be considered of certain factors, such as news impact, diversification and market trends. Nonetheless, we can conclude that for October values, we were unqualified to purchase the assets, as we did not have sufficient cash balance as compared to the holding values of the assets consecutively.

DATE	FTSE100_price	forecasted_return	Cash Balance	Share Holdings	Holding Value	Portfolio Value
2022-01-03	7384.500000	0.000128	4001.5	13.0	95998.500000	100000.000000
2022-01-04	7505.200195	0.000128	4001.5	13.0	97567.602535	101569.102535
2022-01-05	7516.899902	0.000128	4001.5	13.0	97719.698726	101721.198726
2022-01-06	7450.399902	0.000128	4001.5	13.0	96855.198726	100856.698726
2022-01-07	7485.299805	0.000128	4001.5	13.0	97308.897465	101310.397465
...	...	...	...	...	...	...
2023-10-02	7510.700195	0.000128	4001.5	13.0	97639.102535	101640.602535
2023-10-03	7470.200195	0.000128	4001.5	13.0	97112.602535	101114.102535
2023-10-04	7412.500000	0.000128	4001.5	13.0	96362.500000	100364.000000
2023-10-05	7451.500000	0.000128	4001.5	13.0	96869.500000	100871.000000
2023-10-06	7494.600098	0.000128	4001.5	13.0	97429.801274	101431.301274

460 rows × 6 columns

Something doesn't look right here. With ARMA(1,1) mean equation, forecasted returns should not stay constant like this. This seems to be forecasts from a Constant Mean-GARCH model!

-1

Table 4.4 for ARMA (1,1)

In accordance with Table 4.4, we can analyse the forecasted return for ARMA (1,1) model. The same dates and years are being implied in this model, as compared to Table 4.1 above, to ensure fairness in comparing both of the models. In ARMA (1,1) model, it can be seen that forecasted return values remain similar for all the dates in January 2022 and October 2023, in contrast to ARMA (3,3) model in Table 4.1. The value of 0.000128 for this model implies a positive value, where it could be suggested that the market was experiencing low volatility and stable condition. The asset should be bought, nevertheless, if the goal is to earn short term profit, it is best not to purchase first due to limited chance of price appreciation and low forecasted returns. ✓

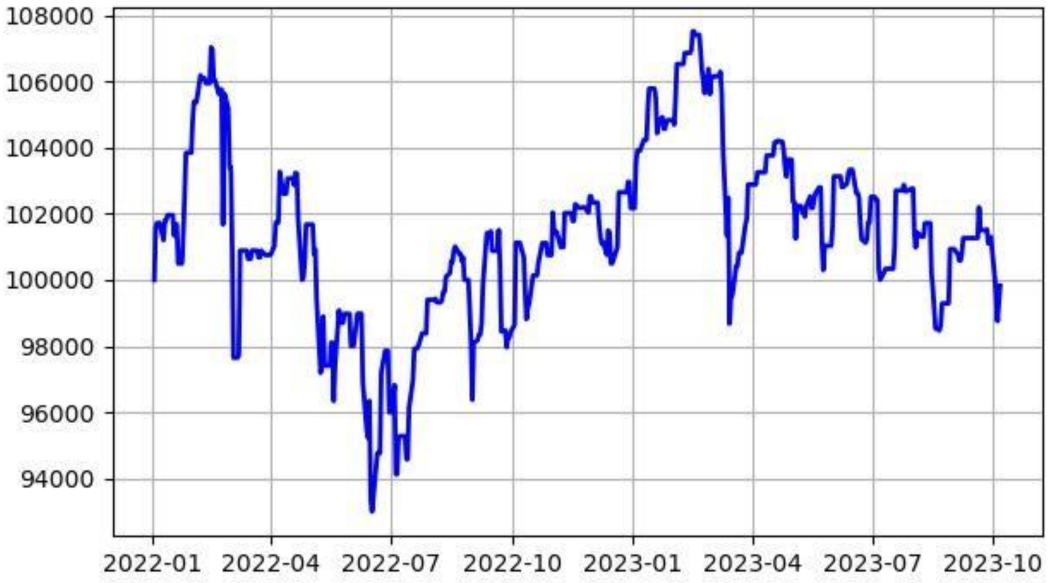
Conclusion for 1(d)

Therefore, from both of the tables, it is obvious that the portfolio value at the end of the out-sample period is larger under the forecast generated by the ARMA (1,1) model, which is 101431, than 99821 under the ARMA (3,3) model. This is mainly caused by the selling and buying behaviour difference under the two-forecasting model. ✓

You could also tabulate Profit/Loss table for different strategies to make the comparison easier to follow.

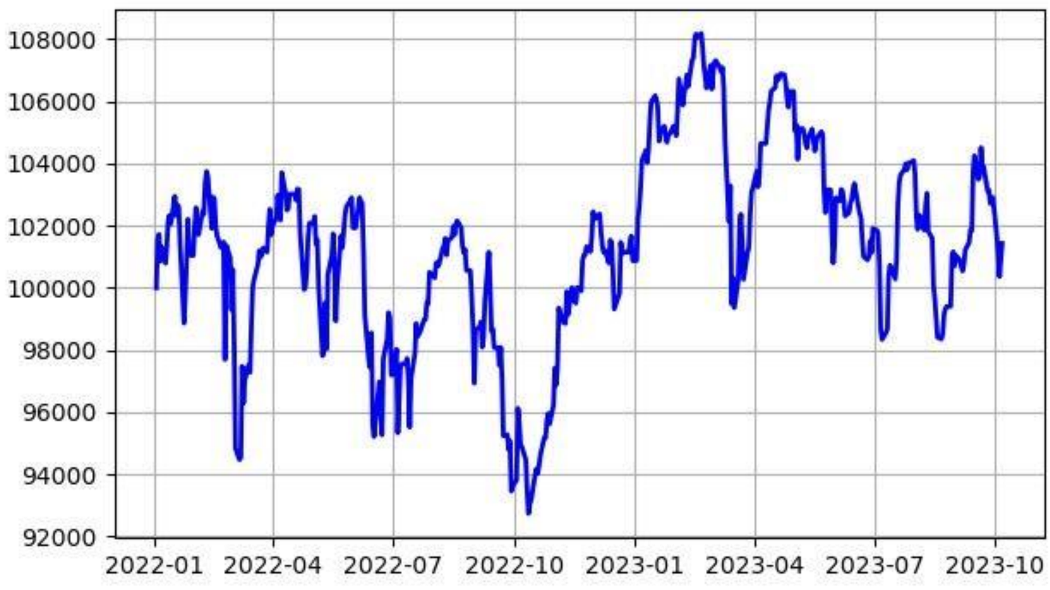
What are the assumptions made for implementing trading strategies based on ARMA-GARCH? e.g. transactions costs, interest paid on cash balances. How would these factors affect the performance of the ARMA-GARCH-based trading strategies?

-4



Graph 4.1

Under ARMA (3,3), the forecasted return fluctuates more because it has a higher ARMA order thus closer to the true return, which leads to more frequent operating even when the fluctuation is only short-term. This could lead to the possibility of missing gains from holding during positive trends, as a temporary negative return for a day during a positive long-term trend would lead to unnecessary selling behaviour , Consequently, when the price resumes its upward movement, the investor remains in cash and misses the benefit of the continued price increase. As can be seen from Graph 4.1, the different stages period and fluctuations indicate that it is not capturing the upward trend. Indirectly, this implies that we do not have sufficient money if we buy or sell the asset too early.



Graph 4.2

On the other hand, the mean forecast is almost constant under the ARMA (1,1) model due to a lower sensitivity to noise or small fluctuations, which avoided unnecessary transactions and guided the investor to buy at the beginning and hold it to the end. From graph 4.2, the trend seems smoother, which captures most of the trends in the market. This can make good use of the long-term forecast and capture the positive trend, thus making a better profit.

Question 2 Q2) 15/20

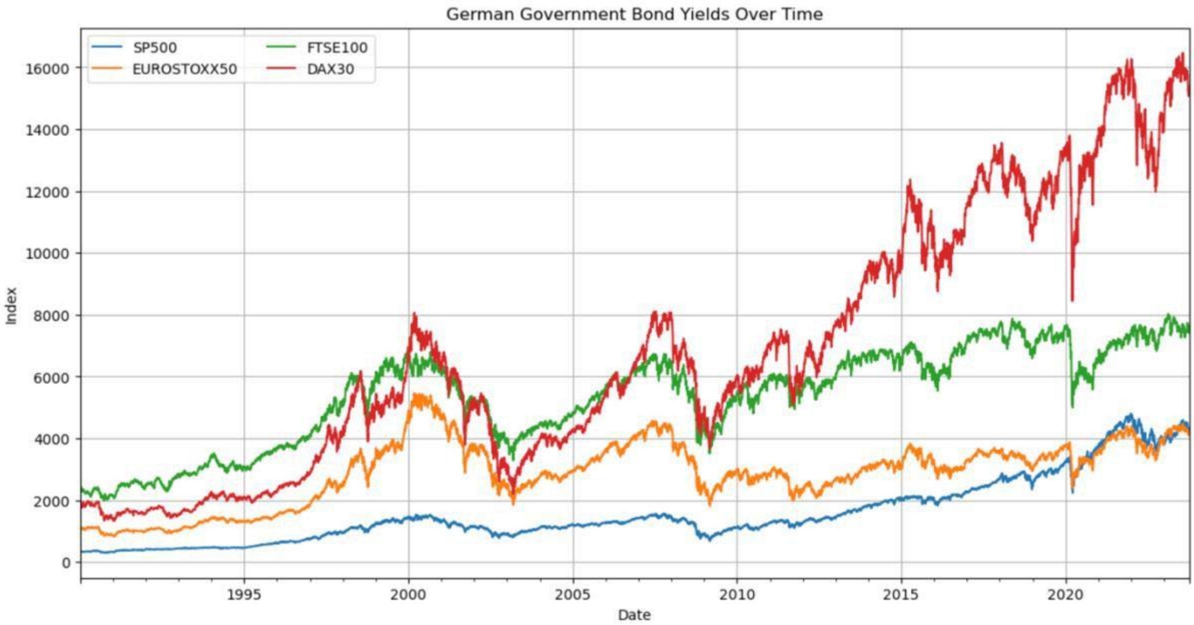
From the StockIndexFXDATA.csv file, choose the SP500, EUROSTOXX50, FTSE100, and DAX30 price series and consider them as a 4 dimensional system. Find an appropriate multivariate model (VAR or VEC), with no more than 5 lags, to explain these 4 stock indices. Write a report outlining your model selection strategy and interpret the estimated results of your selected model.

Introduction

In this assignment, we aim to investigate the relationships between four stock price series (SP500, EUROSTOXX50, FTSE100, and DAX30) using visual interpretation and statistical methods. The analysis will begin by determining whether the four price series follow a normal distribution and non-stationary process by using appropriate statistical tests. After the determination, we can select the most suitable model (VAR/VEC Model) based on the abovementioned properties.

To enhance the explanatory power of the model, we can then further improve the model by fitting the appropriate value of lags based on different test criteria. Finally, we will analyse the result generated by the selected multivariate model and focus on key elements (including cointegrating equation, α Coefficients, β Coefficients) and overall model dynamics. This comprehensive approach will allow us to understand the interrelationships among the stock indices and interpret the economic implications of the findings.

Visual Interpretation



From the graph, we can observe that the four price series selected all illustrated upward trends with more fluctuations over time, indicating no constant mean with increasing variance.

A stationary process, by definition, has a constant mean and variance over time, which cannot show here. Hence, the graph can suggest that the processes of SP500, EUROSTOXX50, FTSE100, and DAX30 price series follow non-stationary processes.

Statistical Methods

	YYYY	MM	DD	DAX30	DJ30	NIKKEI225	FTSE100	CAC40	EUROSTOXX50
count	8809.000000	8809.000000	8809.000000	8809.000000	8809.000000	8809.000000	8809.000000	8809.000000	8809.000000
mean	2006.386423	6.492564	15.724713	6807.835666	13432.568837	17370.128170	5369.328667	4051.364560	2879.153863
std	9.747823	3.439000	8.802296	4109.328310	8761.745403	6067.708625	1567.905437	1491.895004	1052.663038
min	1990.000000	1.000000	1.000000	1322.680000	2365.100000	7054.980000	1990.200000	1441.170000	818.500000
25%	1998.000000	4.000000	8.000000	3547.840000	8017.590000	11959.330000	4197.500000	2955.110000	2284.669000
50%	2006.000000	7.000000	16.000000	5952.920000	10812.870000	17031.630000	5755.300000	4087.490000	3007.510000
75%	2015.000000	9.000000	23.000000	9915.560000	17425.030000	20858.300000	6572.330000	5145.160000	3593.762000
max	2023.000000	12.000000	31.000000	16469.750000	36799.648440	38915.870000	8014.299805	7577.000000	5464.430000
Median	2006.000000	7.000000	16.000000	5952.920000	10812.870000	17031.630000	5755.300000	4087.490000	3007.510000
Skewness	0.001554	0.000328	0.007723	0.584777	1.013811	0.534290	-0.481288	0.092517	-0.239832
Kurtosis	1.799962	1.799138	1.805586	2.271992	3.184889	2.909998	2.150211	2.183085	2.415055
Jarque-Bera	412981.888042	412981.888042	412981.888042	412981.888042	412981.888042	412981.888042	412981.888042	412981.888042	412981.888042
JB pvalue	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

SP500	SMI	VIX	USD_EUR	GBP_EUR	CHF_USD	USD_GBP
8809.000000	8809.000000	8809.000000	8809.000000	8809.000000	8809.000000	8809.000000
1565.374711	6651.601997	19.593476	1.206559	0.774572	1.196087	1.569601
1066.325764	2788.438090	7.900345	0.147594	0.090388	0.247156	0.202189
295.460000	1287.600000	9.140000	0.828700	0.570700	0.722800	1.072754
893.580000	5044.800000	13.880000	1.107450	0.689550	0.969100	1.435250
1265.420000	6871.300000	17.800000	1.209250	0.790380	1.179400	1.570700
2033.110000	8659.690000	23.010000	1.311000	0.854510	1.412300	1.670900
4796.560059	12970.530270	82.690002	1.597850	0.980300	1.820100	2.108200
1265.420000	6871.300000	17.800000	1.209250	0.790380	1.179400	1.570700
1.249449	-0.198918	2.131786	-0.159141	-0.258229	0.439082	0.212196
3.864857	2.322014	11.182563	2.794049	1.783062	2.027769	2.640033
412981.888042	412981.888042	412981.888042	412981.888042	412981.888042	412981.888042	412981.888042
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

Jarque-Bera Test:

- Test whether a given time series or dataset follows a normal distribution.

H_0: The series is normally distributed (Null hypothesis)
H_1: The series is not normally distributed (Alternative hypothesis)

Test statistic:

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

where T is the sample size, S is the skewness and K is the kurtosis.

$$JB \sim \chi^2_{(2)}$$

The test statistic follows a chi-squared distribution with 2 degree of freedom.

Rejection criteria: Reject the null hypothesis (H_0) if p-value < significance level (e.g. 0.05)

From the statistics above, we can observe that the JB test statistics of the four price series are huge, and the JB p-values are all less than 0.05. Hence, we can reject the null and conclude that the SP500, EUROSTOXX50, FTSE100, and DAX30 price series are not normally distributed.

```
Performing ADF Test on SP500
ADF Statistic: 0.750935
p-value: 0.990801
Critical Values:
	1%: -3.431093911070841
	5%: -2.861868759471996
	10%: -2.5669449936834594

Performing ADF Test on EUROSTOXX50
ADF Statistic: -1.813576
p-value: 0.373743
Critical Values:
	1%: -3.4310931502035724
	5%: -2.8618684232610603
	10%: -2.5669448147162024

Performing ADF Test on FTSE100
ADF Statistic: -1.794399
p-value: 0.383246
Critical Values:
	1%: -3.4310942497344024
	5%: -2.8618689091201364
	10%: -2.566945073342138

Performing ADF Test on DAX30
ADF Statistic: -0.496478
p-value: 0.892702
Critical Values:
	1%: -3.4310934036530005
	5%: -2.861868535254939
	10%: -2.5669448743312677
```

Augmented DF-Test:
- Test whether the time series is stationary or non-stationary

H_0: The time series is non-stationary (i.e. has a unit root) (Null hypothesis)
H_1: The time series stationary (Alternative hypothesis)

$$\begin{aligned} Y_t &= \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t, \\ &= \rho Y_{t-1} + \zeta_1 \Delta Y_{t-1} + \zeta_2 \Delta Y_{t-2} + \ldots + \zeta_{p-1} \Delta Y_{t-p+1} + \varepsilon_t, \end{aligned}$$

$H_0: \phi = 1$
$H_1: \phi < 1$

$$\Delta Y_t = \alpha Y_{t-1} + \zeta_1 \Delta Y_{t-1} + \zeta_2 \Delta Y_{t-2} + \ldots + \zeta_{p-1} \Delta Y_{t-p+1} + \varepsilon_t,$$

$H_0: \alpha = 0$
$H_1: \alpha < 0$

where $\alpha = \rho - 1 = \phi_1 + \phi_2 + \ldots + \phi_p - 1.$
where $\rho = \phi_1 + \phi_2 + \cdots + \phi_p$, and $\zeta_j \equiv -[\phi_{j+1} + \phi_{j+2} + \cdots + \phi_p]$ for $j = 1, 2, \cdots, p-1$.

Rejection criteria: Reject the null hypothesis (H_0) if p-value < significance level (e.g. 0.05)

From the statistics above, we can observe that the ADF test static of the four price series is higher than the critical value of 5%. Meanwhile, the p-values of the four price series are all greater than 0.05. Hence, we fail to reject the null for all four series and conclude that the SP500, EUROSTOXX50, FTSE100, and DAX30 price series are non-stationary processes (i.e. have a unit root).

Johansen's Cointegration Trace test

No. of CE(s)		Eigenvalue	Trace Stat	0.05 Crit Value	
0	None	0.004389	55.247097	47.8545	Reject the H_0 → No pair of relationship
1	At most 1	0.001287	16.509898	29.7961	Do not reject the H_0 → One cointegrating relationship
2	At most 2	0.000573	5.165572	15.4943	↑
3	At most 3	0.000014	0.120760	3.8415	Stop here!

Johansen's Cointegration Max-Eigen test

No. of CE(s)		Eigenvalue	Max-Eigen Stat	0.05 Crit Value	
0	None	0.004389	38.737199	27.5858	Reject the H_0 → No pair of relationship
1	At most 1	0.001287	11.344326	21.1314	Do not reject the H_0 → One cointegrating relationship
2	At most 2	0.000573	5.044811	14.2639	↑
3	At most 3	0.000014	0.120760	3.8415	Stop here!

Test for Co-Integration:

- Test to determine whether a group of time series are cointegrated, meaning that they share a long-run equilibrium relationship despite being non-stationary individually.

(a) Johansen's Trace Test

H_0: There are h cointegrating relationships (h = 0, 1, 2, …, n-1) (Null hypothesis)
H_1: There are n cointegrating relationships (i.e. System is stationary) (Alternative hypothesis)

Testing whether the remaining eigenvalues from h+1 to n equal to 0 or not.

Under H_0: $\ln\mathcal{L}_0 = -\frac{nT}{2}\ln(2\pi) - \frac{nT}{2} - \frac{T}{2}\ln(\det(\hat{\Sigma}_{uu})) - \frac{T}{2}\sum_{i=1}^{h}\ln(1-\hat{\lambda}_i)$

Under H_A: $\ln\mathcal{L}_A = \text{————————} - \frac{T}{2}\sum_{i=1}^{n}\ln(1-\hat{\lambda}_i)$

Test statistic:

$$LR = 2(\ln\mathcal{L}_A - \ln\mathcal{L}_0) = -T\sum_{i=h+1}^{n}\ln(1-\hat{\lambda}_i) \overset{a}{\sim} \text{some non-standard asymptotic distribution}$$

(b) Johansen's Maximum Eigenvalue Test

H_0: There are h cointegrating relationships (h = 0, 1, 2, …, n-1) (Null hypothesis)
H_1: There are h+1 cointegrating relationships (i.e. System is stationary) (Alternative hypothesis)

Testing whether the remaining eigenvalues of h+1 equal to 0 or not.

Test statistic:

$$LR = 2(\ln\mathcal{L}_A - \ln\mathcal{L}_0) = -T\ln(1-\hat{\lambda}_{h+1}) \overset{a}{\sim} \text{some non-standard asymptotic distr.}$$

Rejection criteria: Reject the null hypothesis (H_0) if test statistic > critical value (e.g. 0.05 critical value)

To determine the number of cointegrating relations, we proceed sequentially from h=0 to h=n−1 until we fail to reject the null hypothesis. From the statistic above, both Johansen's Trace Test and Johansen's Maximum Eigenvalue Test fail to reject the null when h=1 since the test statistic is smaller than the critical value. Therefore, we can conclude that there is one cointegrating relationship in these four price series.

Because the series are non-stationary and cointegration exist, we select the VEC model.

VECM Order Selection (* highlights the minimums)

	AIC	BIC	FPE	HQIC
0	27.73	27.75	1.101e+12	27.73
1	27.61	27.65*	9.823e+11	27.62
2	27.60	27.65	9.702e+11	27.62
3	27.59	27.65	9.619e+11	27.61*
4	27.59	27.66	9.601e+11	27.61
5	27.59*	27.67	9.597e+11*	27.62

Model Selection Criteria:

- Choose the best model by using model selection criteria such as AIC and BIC, the model with the least value of the criterion is typically chosen.
- Information criteria penalize the model likelihood that should be maximized, by a penalty function which increases with the number of model parameters k.

- Akaike Information Criterion (AIC):

Minimise! → $$AIC(k) = -\frac{2\ln\mathcal{L}}{T} + \frac{2k}{T} \overset{\varepsilon_t \sim N(0,\sigma^2)}{=} \ln\hat{\sigma}^2 + \frac{2k}{T}$$

- Bayes Information Criterion (BIC) or Schwarz (SIC) Information Criterion:

Minimise! → $$BIC(k) = -\frac{2\ln\mathcal{L}}{T} + \frac{k\ln T}{T} \overset{\varepsilon_t \sim N(0,\sigma^2)}{=} \ln\hat{\sigma}^2 + \frac{k\ln T}{T}$$

- Hannan-Quinn Information Criterion (HQ):

Minimise! $$HQ(k) = -\frac{2\ln\mathcal{L}}{T} + \frac{2k\ln(\ln T)}{T} \overset{\varepsilon_t \sim N(0,\sigma^2)}{=} \ln\hat{\sigma}^2 + \frac{2k\ln(\ln T)}{T}$$

where $\hat{\sigma}^2 = 1/T\sum_{t=1}^{T}\hat{\varepsilon}_t^2$, $\hat{\varepsilon}_t$ denotes the residuals from the model, k the number of estimated parameters, T the number of observations and $\mathcal{L}$ the likelihood function of the model.

The **FPE** is calculated as follows:

$$FPE(p) = \frac{(n+p)}{(n-p)} \cdot \hat{\sigma}^2$$

Where:

- n is the number of observations (data points),
- p is the number of estimated parameters (model order),
- $\hat{\sigma}^2$ is the estimate of the error variance (often the mean squared error of the residuals after fitting the model).

We can choose p with the help of the model selection criteria by fitting in a VEC(p) model to select the value of p, which minimizes the value of each selection criterion. The statistics above show that AIC, BIC, FPE and HQIC suggest the value of p to be 5, 1, 5 and 3, respectively. However, since we would like to balance the model's complexity and the model fitting the data reasonably well, we chose lag 3, suggested by HQIC.

VEC Model Summary

VECM Summary:

Det. terms outside the coint. relation & lagged endog. parameters for equation SP500

	coef	std err	z	P>|z|	[0.025	0.975]
L1.SP500	-0.0926	0.013	-7.142	0.000	-0.118	-0.067
L1.EUROSTOXX50	0.0214	0.014	1.556	0.120	-0.006	0.048
L1.FTSE100	-0.0145	0.007	-2.036	0.042	-0.028	-0.001
L1.DAX30	0.0012	0.005	0.234	0.815	-0.009	0.011
L2.SP500	-0.0200	0.013	-1.504	0.133	-0.046	0.006
L2.EUROSTOXX50	-0.0636	0.014	-4.625	0.000	-0.091	-0.037
L2.FTSE100	0.0040	0.007	0.559	0.576	-0.010	0.018
L2.DAX30	0.0383	0.005	7.643	0.000	0.028	0.048

Det. terms outside the coint. relation & lagged endog. parameters for equation EUROSTOXX50

	coef	std err	z	P>|z|	[0.025	0.975]
L1.SP500	0.4805	0.024	20.073	0.000	0.434	0.527
L1.EUROSTOXX50	-0.0775	0.025	-3.052	0.002	-0.127	-0.028
L1.FTSE100	-0.0127	0.013	-0.970	0.332	-0.038	0.013
L1.DAX30	-0.0260	0.009	-2.807	0.005	-0.044	-0.008
L2.SP500	0.0868	0.025	3.542	0.000	0.039	0.135
L2.EUROSTOXX50	-0.0346	0.025	-1.363	0.173	-0.084	0.015
L2.FTSE100	-0.0154	0.013	-1.179	0.238	-0.041	0.010
L2.DAX30	0.0107	0.009	1.154	0.248	-0.007	0.029

Det. terms outside the coint. relation & lagged endog. parameters for equation FTSE100

	coef	std err	z	P>|z|	[0.025	0.975]
L1.SP500	0.7305	0.035	21.003	0.000	0.662	0.799
L1.EUROSTOXX50	-0.0651	0.037	-1.765	0.078	-0.137	0.007
L1.FTSE100	-0.0413	0.019	-2.169	0.030	-0.079	-0.004
L1.DAX30	-0.0492	0.013	-3.648	0.000	-0.076	-0.023
L2.SP500	0.1633	0.036	4.585	0.000	0.093	0.233
L2.EUROSTOXX50	-0.0511	0.037	-1.385	0.166	-0.123	0.021
L2.FTSE100	-0.0414	0.019	-2.178	0.029	-0.079	-0.004
L2.DAX30	0.0162	0.013	1.202	0.229	-0.010	0.043

Det. terms outside the coint. relation & lagged endog. parameters for equation DAX30

	coef	std err	z	P>|z|	[0.025	0.975]
L1.SP500	1.1452	0.059	19.520	0.000	1.030	1.260
L1.EUROSTOXX50	0.0053	0.062	0.085	0.932	-0.117	0.127
L1.FTSE100	-0.0326	0.032	-1.015	0.310	-0.096	0.030
L1.DAX30	-0.1395	0.023	-6.135	0.000	-0.184	-0.095
L2.SP500	0.2087	0.060	3.474	0.001	0.091	0.326
L2.EUROSTOXX50	-0.0404	0.062	-0.650	0.516	-0.162	0.082
L2.FTSE100	-0.0467	0.032	-1.460	0.144	-0.110	0.016
L2.DAX30	0.0313	0.023	1.381	0.167	-0.013	0.076

Loading coefficients (alpha) for equation SP500

	coef	std err	z	P>|z|	[0.025	0.975]
ec1	-0.0010	0.001	-1.474	0.140	-0.002	0.000

Loading coefficients (alpha) for equation EUROSTOXX50

	coef	std err	z	P>|z|	[0.025	0.975]
ec1	0.0052	0.001	4.250	0.000	0.003	0.008

Loading coefficients (alpha) for equation FTSE100

	coef	std err	z	P>|z|	[0.025	0.975]
ec1	0.0039	0.002	2.198	0.028	0.000	0.007

Loading coefficients (alpha) for equation DAX30

	coef	std err	z	P>|z|	[0.025	0.975]
ec1	0.0089	0.003	2.947	0.003	0.003	0.015

Cointegration relations for loading-coefficients-column 1

	coef	std err	z	P>|z|	[0.025	0.975]
beta.1	1.0000	0	0	0.000	1.000	1.000
beta.2	-0.3599	0.106	-3.403	0.001	-0.567	-0.153
beta.3	0.6475	0.115	5.631	0.000	0.422	0.873
beta.4	-0.4303	0.031	-13.791	0.000	-0.491	-0.369
const	-1140.8335	269.880	-4.227	0.000	-1669.790	-611.878

Results

General equation:

$$\Delta y_t = a Z_{t-1} + \theta_1 \Delta y_{t-1} + \theta_2 \Delta y_{t-2} + \varepsilon$$
$$Z_{t-1} = B_1 Y_{1,t-1} + B_2 Y_{2,t-1} + B_3 Y_{3,t-1} + B_4 Y_{4,t-1} + C$$

Cointegrating equation:

1 cointegrating relationship ($Z_{t-1} = \Gamma' Y_{t-1}$)

$$Z_{t-1} = Y_{1,t-1} - 0.3599 Y_{2,t-1} + 0.6475 Y_{3,t-1} - 0.4303 Y_{4,t-1} - 1140.8335$$

Z_t = Long non-relationship between the integrated variables Y_t

---

Interpretation

β Coefficients:

- Represents the cointegration relations between the variables in the system.

- **β1 = 1.0000 (Coefficient for SP500):** SP500 is the base or reference variable in this cointegrating equation. (p = 0.000 ➔ Significant!)
- **β2 = -0.3599 (Coefficient for EUROSTOXX50):** A 1-unit increase in EUROSTOXX50 decreases SP500 by 0.3599 in the long run. (p = 0.001 ➔ Significant!)
- **β3 = 0.6475 (Coefficient for FTSE100):** A 1-unit increase in FTSE100 increases SP500 by 0.6475 in the long run. (p = 0.000 ➔ Significant!)
- **β4 = -0.4303 (Coefficient for DAX30):** A 1-unit increase in DAX30 decreases SP500 by 0.3599 in the long run. (p = 0.000 ➔ Significant!)
- **Constant (Intercept in the cointegration equation):** SP500 will be -1140.8335 in the long run if all of the beta are zero. (p = 0.000 ➔ Significant!)

The cointegration equation is statistically significant, indicating the presence of one long-term equilibrium relationship (p=0.000).

SP500 Equation:

$$\Delta_{Y_1,t} = -0.0010\,Z_{t-1} - 0.0926\,\Delta_{Y_1,t-1} + 0.0214\,\Delta_{Y_2,t-1} - 0.0145\,\Delta_{Y_3,t-1} + 0.0012\,\Delta_{Y_4,t-1}$$
$$- 0.0200\,\Delta_{Y_1,t-2} - 0.0636\,\Delta_{Y_2,t-2} + 0.0040\,\Delta_{Y_3,t-2} + 0.0383\,\Delta_{Y_4,t-2} + \varepsilon_t$$

EUROSTOXX50 Equation:

$$\Delta_{Y_2,t} = 0.0052\,Z_{t-1} + 0.4805\,\Delta_{Y_1,t-1} - 0.0775\,\Delta_{Y_2,t-1} - 0.0127\,\Delta_{Y_3,t-1} - 0.0260\,\Delta_{Y_4,t-1}$$
$$+ 0.0868\,\Delta_{Y_1,t-2} - 0.0346\,\Delta_{Y_2,t-2} - 0.0154\,\Delta_{Y_3,t-2} + 0.0107\,\Delta_{Y_4,t-2} + \varepsilon_t$$

FTSE100 Equation:

$$\Delta_{Y_3,t} = 0.0039\,Z_{t-1} + 0.7305\,\Delta_{Y_1,t-1} - 0.0651\,\Delta_{Y_2,t-1} - 0.0413\,\Delta_{Y_3,t-1} - 0.0492\,\Delta_{Y_4,t-1}$$
$$+ 0.1633\,\Delta_{Y_1,t-2} - 0.0511\,\Delta_{Y_2,t-2} - 0.0444\,\Delta_{Y_3,t-2} + 0.0162\,\Delta_{Y_4,t-2} + \varepsilon_t$$

DAX30:

$$\Delta_{Y_4,t} = 0.0089\,Z_{t-1} + 1.1452\,\Delta_{Y_1,t-1} + 0.0053\,\Delta_{Y_2,t-1} - 0.0326\,\Delta_{Y_3,t-1} - 0.1395\,\Delta_{Y_4,t-1}$$
$$+ 0.2087\,\Delta_{Y_1,t-2} - 0.0404\,\Delta_{Y_2,t-2} - 0.0467\,\Delta_{Y_3,t-2} + 0.0313\,\Delta_{Y_4,t-2} + \varepsilon_t$$

α Coefficients:
- Reflect the speed at which a variable adjusts to deviations from the long-run equilibrium.

- **SP500 (α1 = -0.0010):** The small negative coefficient indicates that SP500 will adjust downwards slowly to return to equilibrium. (p = 0.140 ➔ Not Significant!)
- **EUROSTOXX50 (α2 = 0.0052)**: The small positive coefficient indicates that EUROSTOXX50 will adjust upwards slowly to return to equilibrium. (p = 0.000 ➔ Significant!)
- **FTSE100 (α3 = 0.0039)**: The small positive coefficient indicates that FTSE100 will adjust upwards slowly to return to equilibrium. (p = 0.028 ➔ Significant!)
- **DAX30 (α4 = 0.0089)**: The small positive coefficient indicates that DAX30 will adjust upwards slowly to return to equilibrium. (p = 0.003 ➔ Significant!)

DAX30 has the largest alpha coefficient, suggesting that DAX30 has the fastest adjustment speed among the variables in the system, meaning that it will correct more quickly if it deviates from the long-run equilibrium compared to the others.

There is no strong evidence showing that SP500 can be significantly adjusted to correct deviations from the long-run equilibrium, meaning that α1 is not a significant factor in this model (p=0.140).

The interpretations of individual coefficients here are not that meaningful because you need to "keep other variables fixed". Should perform Granger causaliy and/or Impulse Response Analysis instead.

Overall model dynamics

Example: SP500 Equation

- **L1.SP500**: The coefficient of −0.0926 indicates that a one-unit increase in SP500 at lag 1 is associated with a decrease of 0.0926 in SP500 in the current period. (p = 0.000 ➔ Significant!)
- **L1.EUROSTOXX50**: The coefficient of 0.0214 indicates that a one-unit increase in EUROSTOXX50 at lag 1 is associated with an increase of 0.0214 in SP500 in the current period.
 (p = 0.120 ➔ Not Significant!)
- **L1.FTSE100**: The coefficient of −0.0145 indicates that a one-unit increase in FTSE100 at lag 1 is associated with a decrease of 0.0145 in SP500 in the current period. (p = 0.042 ➔ Significant!)
- **L1.DAX30**: The coefficient of 0.0012 indicates that a one-unit increase in DAX30 at lag 1 is associated with an increase of 0.0012 in SP500 in the current period.
 (p = 0.815 ➔ Not Significant!) ✓
- **L2.SP500**: The coefficient of −0.0200 indicates that a one-unit increase in SP500 at lag 2 is associated with a decrease of 0.0200 in SP500 in the current period. (p = 0.133 ➔ Not Significant!)
- **L2.EUROSTOXX50**: The coefficient of −0.0636 indicates that a one-unit increase in EUROSTOXX50 at lag 2 is associated with a decrease of 0.0636 in SP500 in the current period. (p = 0.000 ➔ Significant!)
- **L2.FTSE100**: The coefficient of 0.0040 indicates that a one-unit increase in FTSE100 at lag 2 is associated with an increase of 0.0040 in SP500 in the current period. (p = 0.576 ➔ Not Significant!)
- **L2.DAX30**: The coefficient of 0.0383 indicates that a one-unit increase in DAX30 at lag 2 is associated with an increase of 0.0383 in SP500 in the current period. (p = 0.000 ➔ Significant!)

SP500 (lag 1), FTSE100 (lag 1), EUROSTOXX50 (lag 2), and DAX30 (lag 2) show a significant relationship between SP500 (p-value < 0.05).

Some variables, such as EUROSTOXX50 (lag 1), DAX30 (lag 1), and FTSE100 (lag 2), are not statistically significant at the 5% level, implying their impact on SP500 in the current period is not strong enough to be considered reliable (p-value > 0.05).

	lb_stat	lb_pvalue
1	0.105950	0.744803
2	0.112622	0.945245
3	0.195824	0.978261
4	1.407046	0.842969
5	1.660927	0.893785
6	1.662509	0.947979
7	1.662556	0.976143
8	3.534578	0.896490
9	4.404779	0.882812
10	4.459710	0.924235

Ljung-Box Test:
- Testing whether the residuals follow the weak white noise property.

$H_0: \rho_1 = \rho_2 = \ldots = \rho_K = 0$ → Implies that the residual are uncorrelated (white noise)

$H_1: \rho_1 \neq 0$ for some $j \in \{1, \ldots, k\}$ → Implies that the residual are serial correlated

Test statistic:

$$Q_{LB}(k) = T(T+2) \sum_{j=1}^{k} \frac{\hat{\rho}_j(\hat{\varepsilon})^2}{T-j} \overset{a}{\sim} \chi^2_{(k-p-q)},$$

← Chi-square distribution

Rejection criteria: Reject the null hypothesis (H_0) if p-value < significance level (e.g. 0.05)

From the statistics above, we can observe that the p-values of all lags are higher than the significant level of 0.05. Indicating that we failed to reject the null and that no significant serial correlation exists in the residuals. Therefore, we can conclude that all lags of the VEC model are well-specified since it can capture the temporal dynamics of the system.

Conclusion

In this question, we successfully analyse the relationship among the four price series (SP500, EUROSTOXX50, FTSE100, and DAX30) within different time series using visual interpretation and statistical methods. The series is then determined to be non-normal and non-stationary, with one cointegrating relationship identified in the Johansen tests. Based on these findings, we can confidently select the Vector Error Correction (VEC) Model to explain the stock relationships further.

The lag value of 3 was determined by using the model section criteria (HQIC) since we would like to balance the model complexity and fit. Key results from the VEC model revealed significant cointegration coefficients and adjustment speeds, providing insights into the stock indices' long-term equilibrium and dynamic interdependencies.
The Ljung-Box Test further ensures the model's adequacy by checking there is no significant serial correlation among the residuals. Overall, the study provides a solid foundation for analysing the interrelated dynamics of financial markets and their broader economic significance.

Granger-causality and Impulse Response Analyses? -5

Reference:

Abdollahi, H., Junttila, J.-P. and Lehkonen, H. (2024) 'Clustering Asset Markets based on volatility connectedness to political news', *Journal of International Financial Markets, Institutions and Money*, pp. 1–25. doi:10.1016/j.intfin.2024.102004.

Arif, A. (2024) *Stationary time-series analysis*, *Timescale*. Available at: https://www.timescale.com/learn/stationary-time-series-analysis (Accessed: 05 December 2024).

CFI(2024)VIX - Volatility Index. Available at: https://corporatefinanceinstitute.com/resources/career-map/sell-side/capital-markets/vix-volatility-index (Accessed: 05 December 2024).

Forsdick, S. (2024) *FTSE 100: 40 years of ups and downs at the top of British business*, *Raconteur*. Available at: https://www.raconteur.net/finance/ftse-100-40-years (Accessed: 05 December 2024).

Suthar, J. (2023) *Volatility clustering – what is it, examples, importance, causes*, *WallStreetMojo*. Available at: https://www.wallstreetmojo.com/volatility-clustering/ (Accessed: 05 December 2024).

AI Appendix

Tool	Purpose	Prompt	Outcomes for this assignment
Thetawise	Creating the structure of question 1a	Structuring for descriptive data	Headings and subheadings
Thetawise	creating the structure of 1b	what is the step- by-step guide in a report for ARMA best model	Headings and subheadings
Perplexity	Finding available resources for definition of vix and ftse100 definitiobn	Definition of VIX and FTSE100	CFI definition for vix Definition of volatility clustering FTSE 100 is a benchmark of the UK Stock market
Perplexity	Finding references for FTSE100 index stock price changes	FTSE 100 index price changes	timeline for FTSE100 index price in 40 years