F70TS – Time Series and Machine Learning

Student ID: H00381869

<u>Question 1</u>

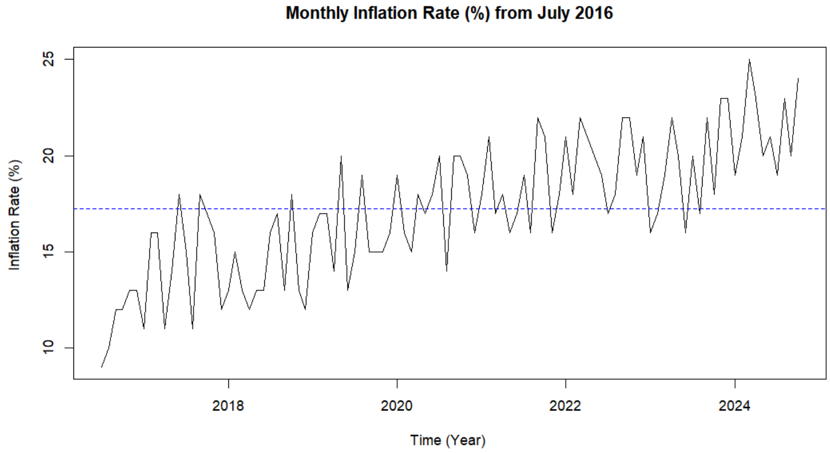

Figure 1: Time series plot of original data.

According to Figure 1, the original data shows a clear upward trend over time. As seen in Figure 1, the blue dashed lines indicated the mean of the inflation rate (%) where the mean appears to be at 17.24% inflation, possibly suggest that the average inflation rate is increasing. Hence, the data appears to have a strong non-stationary behaviour.

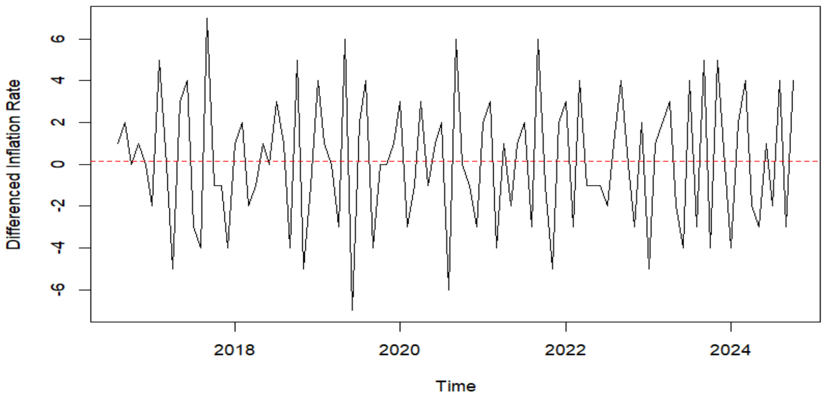

Figure 2: Time series plot the first difference of data.

Based on Figure 2, the first-differenced data appears to have achieved stationarity, which is critical for ARIMA modelling. After differencing, the trend has been removed, and the data oscillates around a mean close to zero, with constant variance.

Question 2

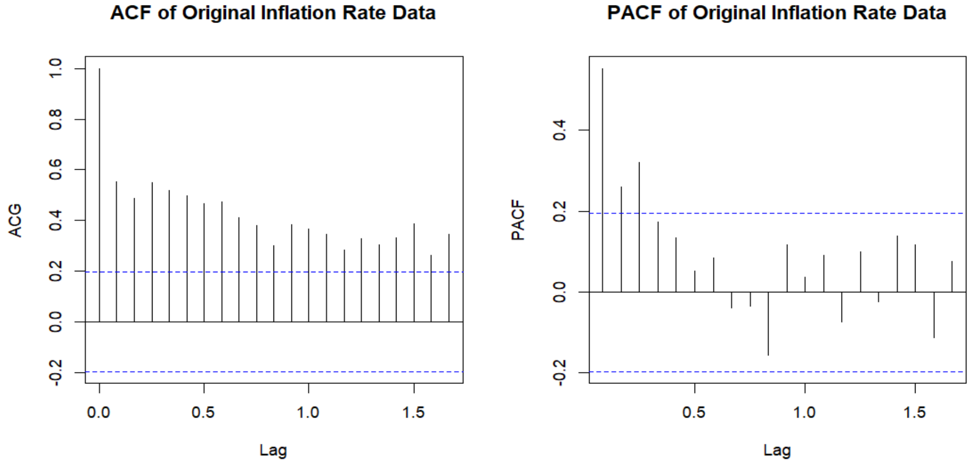

Figure 3: ACF (left-hand plot) and PACF (right-hand plot) of the original data.

Based on the left plot in Figure 3, the ACF of the original data shows significant autocorrelations that remain above the error bounds for multiple lags. This behaviour confirms that the data is non-stationary and requires differencing to achieve stationarity.

In the right plot of Figure 3, the PACF shows a significant spike at lag 1, followed by smaller spikes that gradually decay, with most autocorrelation values falling within the error bounds. This pattern suggests an AR (1) model, where the current value is primarily influenced by its immediate past value, with diminishing influence from earlier lags.

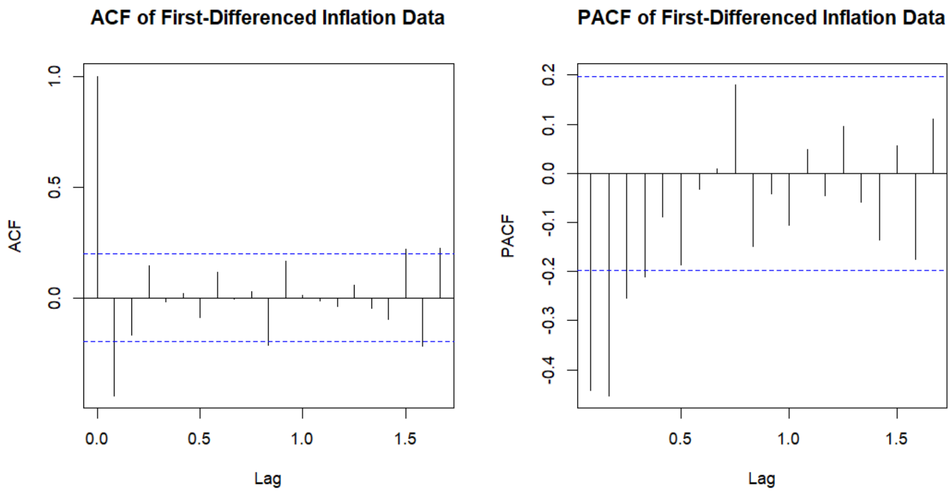

Figure 4: ACF (left-hand plot) and PACF (right-hand plot) of first-differenced data

The ACF plot on the left in Figure 4, shows that there is a significant positive correlation at lag 1, indicating a strong autocorrelation between the current value and the previous value of the first differenced series and hence suggest that the data follows a MA (1) model. The ACF values at lag 10 and lags 18 to 20 fall outside the confidence interval, but it is expected as approximately 5% of values will lie outside the bounds due to random variation in a 95% confidence interval. The PACF plot in Figure 4 exhibits an exponentially decaying pattern after the first lag, where the partial autocorrelations diminish rapidly and remain within the confidence interval for higher lags.

This characteristic is typical of MA processes and further supports the identification of an MA (1) model, as the influence of past values decays quickly and does not extend beyond the first lag. The absence of significant partial autocorrelations beyond lag 1 confirms that there is no need for an autoregressive (AR) component in this model.

Question 3

Based on the observations from ACF and PACF of the first differenced data, the suitable ARIMA (p,1, q) model is ARIMA (0,1,1).

Question 4

A drift is included in the model because the original data no longer has a trend after differencing but there was an upward trend before. Hence, by fitting the ARIMA (0,1,1) model with drift, information that are extracted out from the summary are as below:

- MA (1) coefficient (θ_1) $= -1.0000$
- Drift (μ) = 0.0938
- Variance (σ^2) = 4.807

By substituting the parameters values for d = 1 and t =1, we achieve the equation of the fitted model:

$$(1-B)^d(X_t - \mu t) = (1+\theta_t B)\varepsilon_t$$

$$(1-B)(X_t - 0.0938) = (1-B)\varepsilon_t$$

Where $\varepsilon_t \sim WN(0{,}4.807)$ is the white noise process.

The AIC value for ARIMA (0,1,1) model is 444.97. The inclusion of drift has lowered the AIC, suggesting that it improves the model, However, the MA (1) coefficient being exactly -1.0000 might indicate that the data behaves somewhat like a random walk with drift, implying that future values are largely determined by the drift component rather than past observations.

Question 5

Model	AIC	BIC	MAE	RMSE
ARIMA (0,1,1) with Drift	444.97	452.76	1.872102	2.159346
ARIMA (2,1,5)	461.48	482.24	1.899294	2.267672

Table 1: Table of Comparison for AIC and BIC of ARIMA (0,1,1) and ARIMA (2,1,5).

According to Table 1, the AIC and BIC values for the ARIMA (0,1,1) with drift model are 444.97 and 452.76, respectively, which are significantly lower than those of the ARIMA (2,1,5) model. Lower AIC and BIC values indicate that the ARIMA (0,1,1) with drift provides a better fit to the data while using fewer parameters. This suggests that the simpler model captures the essential patterns in the data more efficiently, without introducing unnecessary complexity.

The ARIMA (2,1,5) model includes multiple AR (autoregressive) and MA (moving average) terms, increasing the model's complexity and the potential for overfitting. Overfitting occurs when a model captures noise or minor fluctuations in the data, which can impair its predictive performance on new data.

In terms of forecast accuracy, the Mean Absolute Error (MAE) for ARIMA (0,1,1) with drift is lower than that of the ARIMA (2,1,5) model, indicating that, on average, the forecasts from the simpler model are closer to the actual values. Additionally, the Root Mean Squared Error (RMSE) for the

ARIMA (0,1,1) with drift is also lower, suggesting that this model does a better job of minimizing larger forecast errors, making it more reliable for predictions.

Therefore, the ARIMA (0,1,1) with drift is the more appropriate model as it strikes a balance between simplicity and goodness of fit. With fewer parameters, it reduces the risk of overfitting while still maintaining strong predictive power.

Question 6

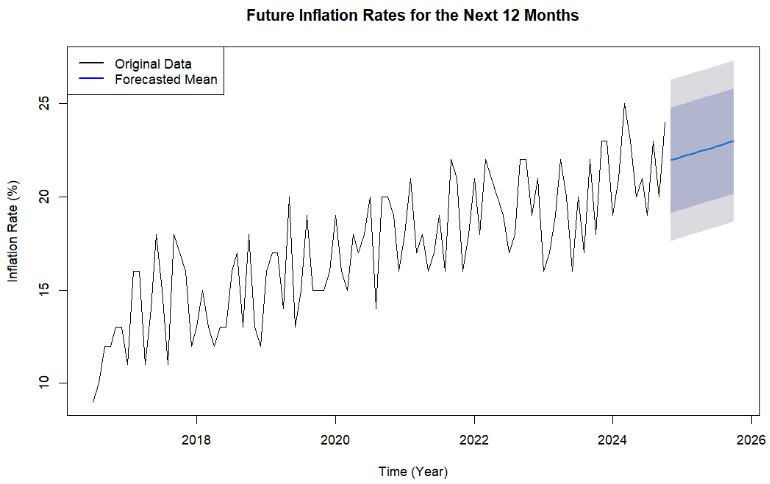

Figure 5: Plot of future inflation rates for the next 12 months.

Based on Figure 5, the black line represents the original data, where a clear upward trend is visible, along with noticeable fluctuations indicating that the inflation rate has been increasing over time. This trend suggests persistent inflationary pressures in the economy, possibly driven by factors such as supply constraints, increased demand, or monetary policy influences.

The blue line represents the forecasted mean inflation rate for the next 12 months. According to the ARIMA (0,1,1) model with drift, the inflation rate is expected to continue rising but at a slightly slower pace compared to the historical data, reflecting a gradual growth trend in the forecast. This indicates a potential moderation in inflation, suggesting that while prices are still expected to rise, the rate of increase may stabilize or decelerate somewhat, possibly due to anticipated economic interventions or changing market conditions.

In the plot, the shaded region around the forecast depicts the 95% confidence intervals, indicating that the actual future inflation rate is expected to fall within this range with 95% confidence. As observed, the width of the interval increases as the forecast horizon extends, highlighting the growing uncertainty in predictions over time. This widening of the confidence intervals reflects the model's increasing difficulty in making accurate long-term forecasts due to the potential for unexpected economic events or structural changes.