**FORECASTING TIME SERIES**

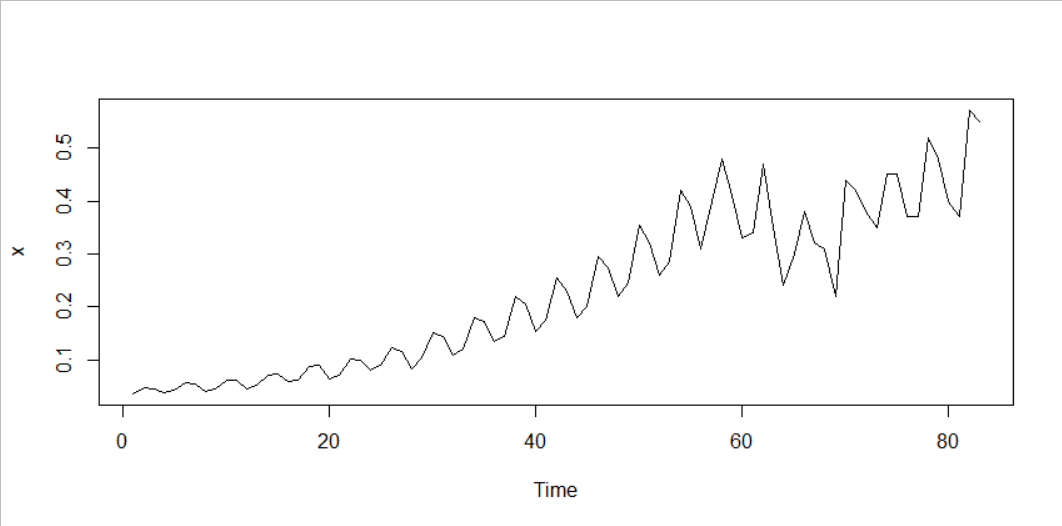
**Group Assignment -2**

**QUESTION 1**

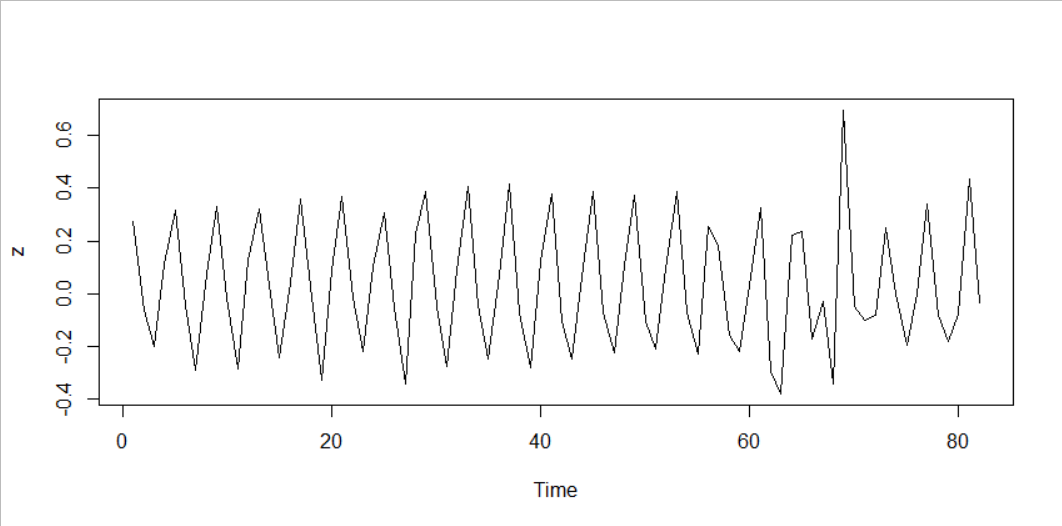
Find at least two linear time series models, using the Box-Jenkins methodology, for the quarterly earnings per share of Coca-Cola Company from the first quarter of 1983 to the third quarter of 2009. Identify your models using the entire available sample (coca\_cola\_earnings.csv)

**ANSWER 1**

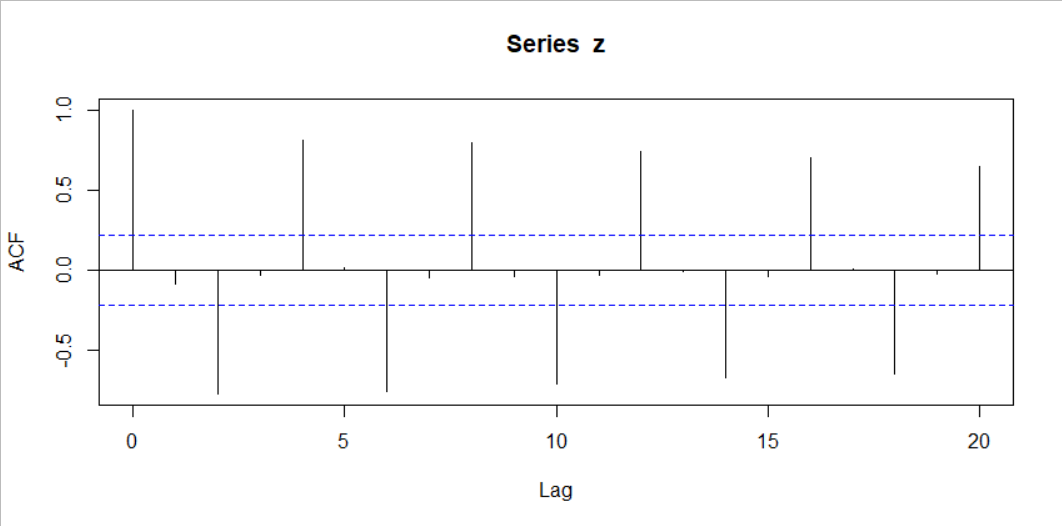
Looking at the data sample, we can clearly see that the data has a trend, hence is seasonal and that it is not stationary in the mean or variance.



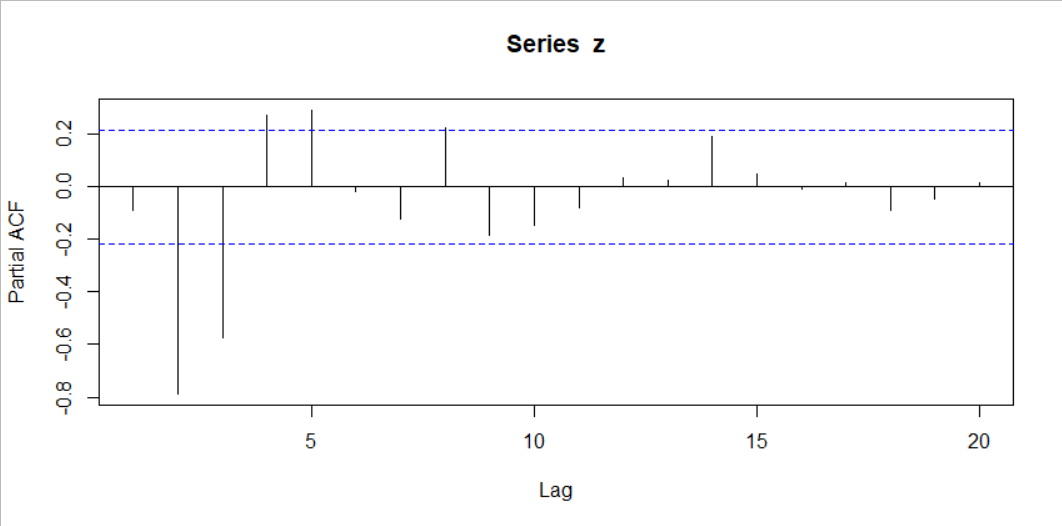
We establish that the after taking a log difference, and running the ADF(1) and OCSB(1) formal tests we need to take 1 regular and 1 seasonal difference each to achieve stationarity in the mean. After taking the log and 1 regular difference we can see that the data is stationary in both mean and variance.



After plotting the the ACF and PACF, we can see that ACF is clearly seasonal with lags out of bound at a frequency of 2 (2, 4, 6, 8, 10..etc.).



The PACF decays towards zero slowly, with lags 2, 3, 4, 5 and 8 out of bound.



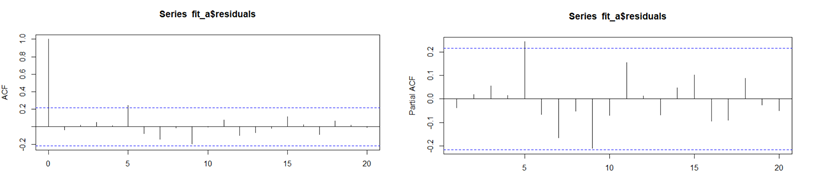
Our seasonal parameter is s=4, as we have quarterly data.

**Our two models are:**

**Model -1: MA(1) & S-MA(1)**

After fitting the model to (0, 1, 1), (0, 1, 1) with period = 4, and looking at the coefficients we can conclude that MA(1) and SMA(1) are significant with 0 outside the values.

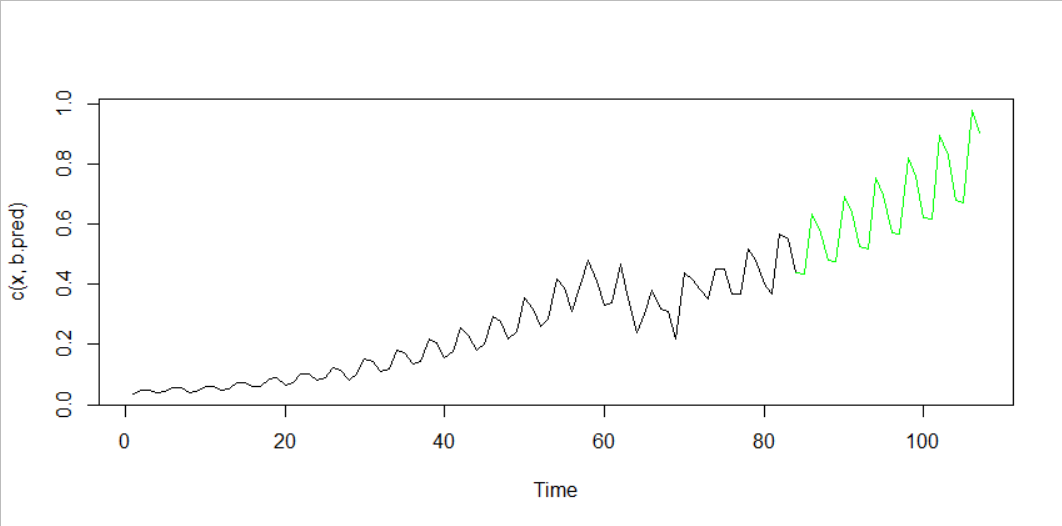
We look at the ACF & PACF of the residuals.



With only 1 lag out of bound at 5, for both ACF, PACF, we can say it doesn’t look like WN.

However with Box test, P-value on lag 5 (0.3801) is higher than 0.05, we can consider the lag out of bounds as WN and proceed with this model.

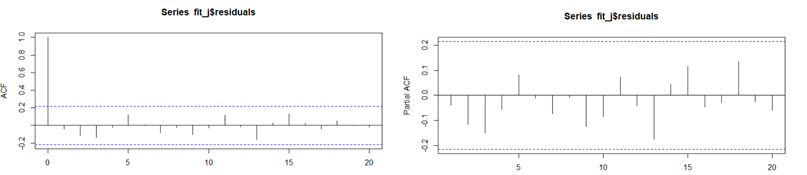
We then calculate the point prediction for 24 quarters (6 years) ahead, however before doing so, we transform the data back to its original state to undo the logarithmic conversion. We then plot the real data with the point predictions to conclude the model successfully can forecast the quarterly earnings per share for coca-cola company, 6 years into the future.



**Model -2: AR(2), MA(2) & S-MA(2)**

After fitting the model to (2, 1, 2), (0, 1, 2) with period = 4 and looking at the coefficients we can conclude that AR(2), MA(2) and SMA(2) are significant with 0 outside the values.

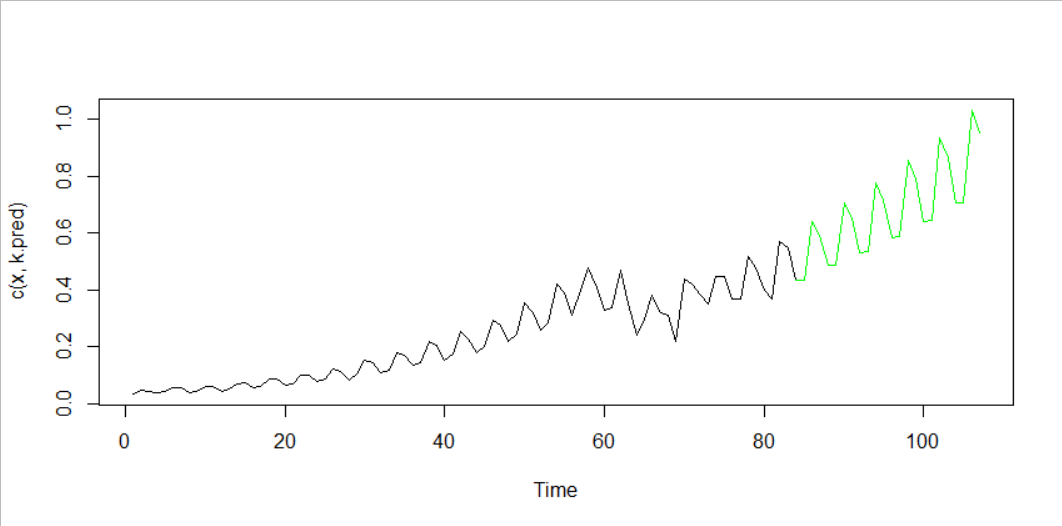
We look at the ACF & PACF of the residuals.



With no lag out of bound we can say it is WN.

With Box test, P-value on lag 5 (0.7256) is higher than 0.05, we confirm it as WN and proceed with this model.

We then calculate the point prediction for 24 quarters (6 years) ahead, however before doing so, we transform the data back to its original state to undo the logarithmic conversion. We then plot the real data with the point predictions to conclude the model successfully can forecast the quarterly earnings per share for coca-cola company, 6 years into the future.



**QUESTION 2**

For the models identified in the previous step, leave for example the last 24 real values to compare all the models in terms of forecasting (out of sample forecasting exercise). What is the best model and why is this your choice?

**ANSWER 2**

Comparing our two models (0,1,1)(0,1,1) and (2,1,2)(0,1,2) with the MAPE and MSFE values, we can conclude that (0,1,1)(0,1,1) is the best model as it has the lowest MAPE and MSFE values for 24 step-ahead forecast.

