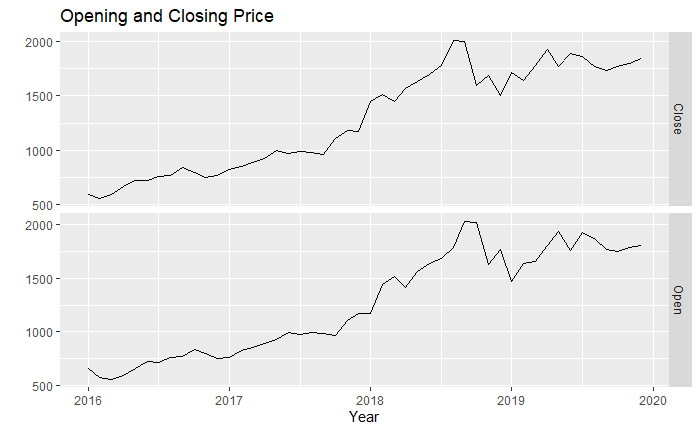
# Time Series Regression

There might be cases that one time series is a combination of the AR and MA components and at the same time depends on one or more Time series. In such cases we can use the ARIMA model along with the xreg input which models the time series as with the predictor variables.

In our case, we have tried to use this concept to forecast the Closing price of the stock given the Opening pric. This kind of model can help us predict the Closing value of the stock at the end of a month. Another underlying fact here is that there is a strong correlation between the Opening and Closing Price of the stock.

We start with checking if the linear regression model fits the data well. Below plot suggests that there might be association in these two variables.



Below is the summary of the linear model:

## Call:  
## lm(formula = Close ~ Open, data = amazon.vol.close.ts)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -418.83 -49.82 -2.97 51.92 252.69   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 69.42273 47.65039 1.457 0.152   
## Open 0.96312 0.03532 27.272 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 115.6 on 46 degrees of freedom  
## Multiple R-squared: 0.9418, Adjusted R-squared: 0.9405   
## F-statistic: 743.7 on 1 and 46 DF, p-value: < 2.2e-16

The p-value and R-squared suggests that the model holds good and we can use this in the xreg part in ARIMA.

## Fitting ARIMA model:

(fit <- Arima(amazon.vol.close.ts[,"Close"],

order=c(1,1,0), season=list(order=c(2,1,0)),

xreg=amazon.vol.close.ts[,"Open"]))

## Series: amazon.vol.close.ts[, "Close"]   
## Regression with ARIMA(1,1,0)(2,1,0)[12] errors   
##   
## Coefficients:  
## ar1 sar1 sar2 xreg  
## -0.3130 -1.1043 -0.7208 0.0530  
## s.e. 0.2653 0.1361 0.1613 0.2292  
##   
## sigma^2 estimated as 8678: log likelihood=-218.27  
## AIC=446.54 AICc=448.61 BIC=454.32

The AIC value here is lower than the auto.arima function. Hence we go ahead with this model. The below residual plot suggest that there is still one spike in ACF plot and we can investigate further. For the scope of this project, we will use this model.

