IE360 – ASSIGNMENT 3 - FATMANUR YAMAN – 2019402204 – 18.05.2023

1. **For an AR(2) model: Yt = ϕ1Yt−1 + ϕ2Yt−2 + Et, find the autocorrelations for the first 2 lags.**

1) AR(2)

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t \qquad Var(Y_t) = \gamma_0$$

Let's obtain γ_k's for $k \geq 1$.
Multiply both sides by Y_{t-k}.

$$Y_{t-k} Y_t = \phi_1 Y_{t-1} Y_{t-k} + \phi_2 Y_{t-2} Y_{t-k} + \overbrace{\varepsilon_t Y_{t-k}}^{\text{independent}}$$

$$E[Y_{t-k} Y_t] = \phi_1 E[Y_{t-1} Y_{t-k}] + \phi_2 E[Y_{t-2} Y_{t-k}] + \underbrace{E[\varepsilon_t Y_{t-k}]}_{0}$$

$$Cov(Y_{t-k} Y_t) = E[Y_{t-k} Y_t] - \underbrace{E[Y_{t-k}]E[Y_t]}_{0} = \gamma_k$$

$$Cov(Y_{t-k}, Y_t) = \phi_1 Cov(Y_{t-1}, Y_{t-k}) + \phi_2 Cov(Y_{t-2}, Y_{t-k}) + \overbrace{Cov(\varepsilon_t, Y_{t-k})}^{0} \quad k=1,2,\ldots$$

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} \qquad \gamma_0 = var(Y_t) \text{ (by stationary)}$$

autocovariance coefficient $\rho_k = \dfrac{\gamma_k}{\gamma_0} = \phi_1 \dfrac{\gamma_{k-1}}{\gamma_0} + \phi_2 \dfrac{\gamma_{k-2}}{\gamma_0}$

$$\boxed{\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}}$$

$$\rho_1 = \phi_1 \overset{1}{\rho_0} + \phi_2 \overset{\rho_1}{\rho_{-1}}$$

$$\rho_1 = \phi_1 + \phi_2 \rho_1$$

$$\Rightarrow \boxed{\rho_1 = \frac{\phi_1}{1-\phi_2}}$$ 1 lag autocorrelation.

$$\rho_2 = \phi_1 \rho_1 + \phi_2 \rho_0$$

$$= \phi_1 \cdot \frac{\phi_1}{1-\phi_2} + \phi_2 \cdot 1$$

$$\Rightarrow \boxed{\rho_2 = \frac{\phi_1^2 + \phi_2 - \phi_2^2}{1-\phi_2}}$$ 2 lag autocorrelation

2. **Consider the data provided in the file "IBMStocks.xls", listing weekly stock prices for IBM.**

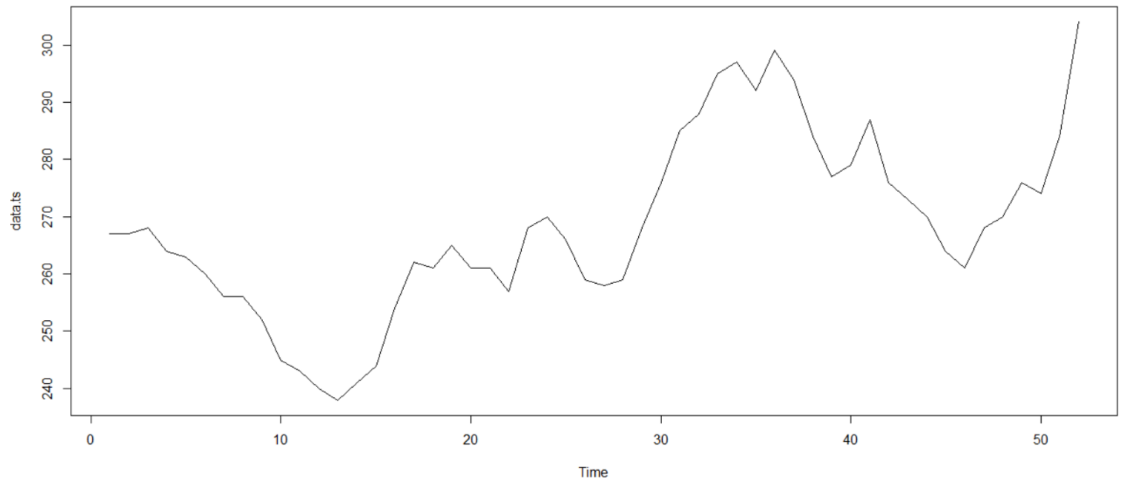
a. **Using R obtain a time series plot of the data, plots of the sample autocorrelations and the sample partial autocorrelations. What are your main observations?**

The Code:

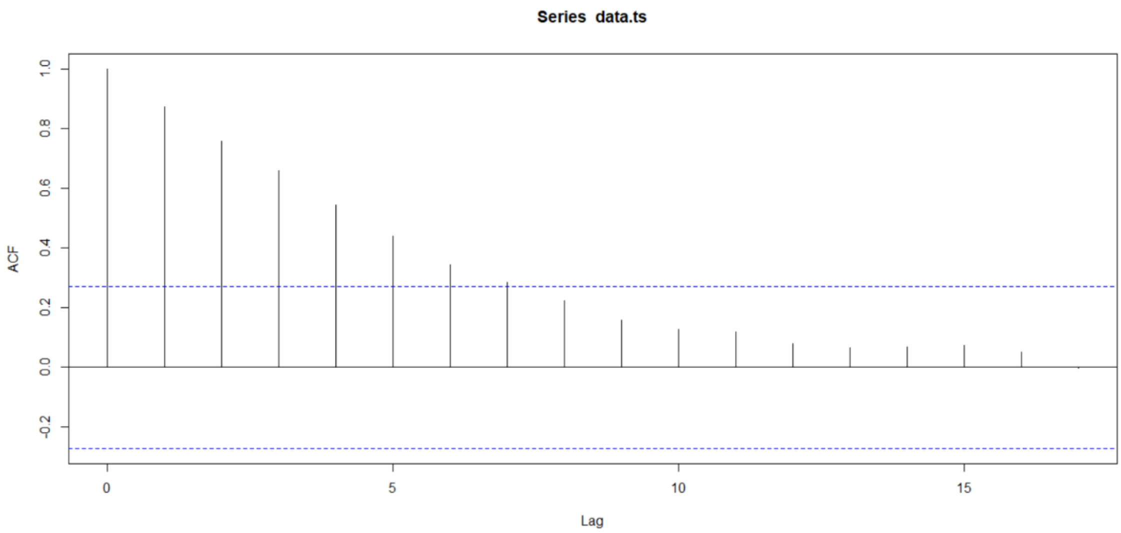
```
data.ts<-ts(data$'IBM Stock Price')
data
#Plot
plot(data.ts)
#Autocorrelation Function
acf(data.ts)
#Partial Autocorrelation
pacf(data.ts)
#Test
adf.test(data.ts)
```

The Output:

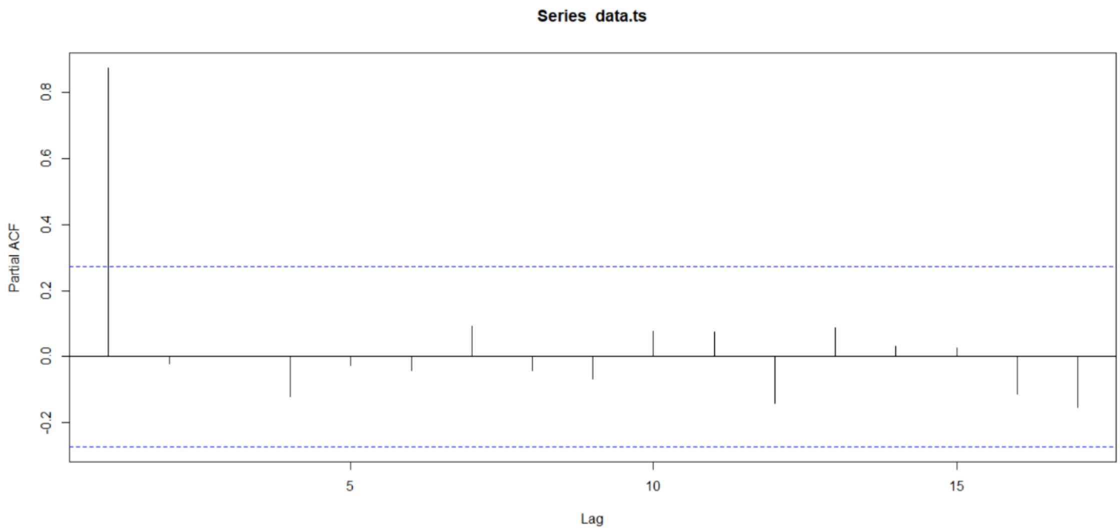
1. **Time Series Plot**



2. **Autocorrelation Function Plot**



3. **Partial Autocorrelation Function Plot**



4. **Augmented Dickey-Fuller Test Result**

```
        Augmented Dickey-Fuller Test

data:  data.ts
Dickey-Fuller = -2.5817, Lag order = 3, p-value = 0.3404
alternative hypothesis: stationary
```

The Observations:

When the time series plot is examined, it can be seen that there is an increasing linear trend in the data. So, by just looking at the time series plot, it can be said that the data is not stationary.

The autocorrelation function is persistent. It is not cutting off or trailing of near zero after a few lags. Slow decay in the autocorrelation function is a clear indication of non-stationary time series which means that the data includes trend. The observation from the time series plot is proved.

The partial autocorrelation function cutts of at 1 and it is an indication of AR(1) model where the data is dependent to the previous observation.

In addition to these, the p-value obtained from Augmented Dickey-Fuller Test is 0.3404. It is a small value, thus the null hypothesis can be rejected. The test claims that the data is non-stationary.

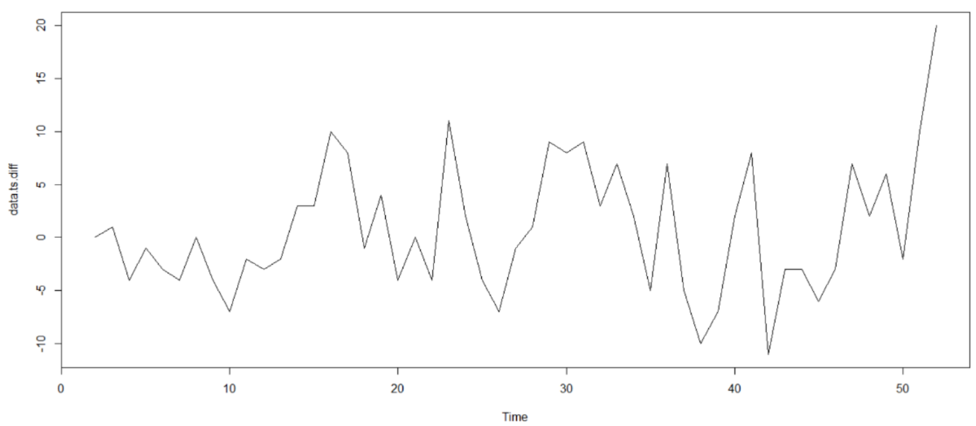
b. **Is the series look stationary? If not, what correction would you recommend to remedy this situation?**

The Code

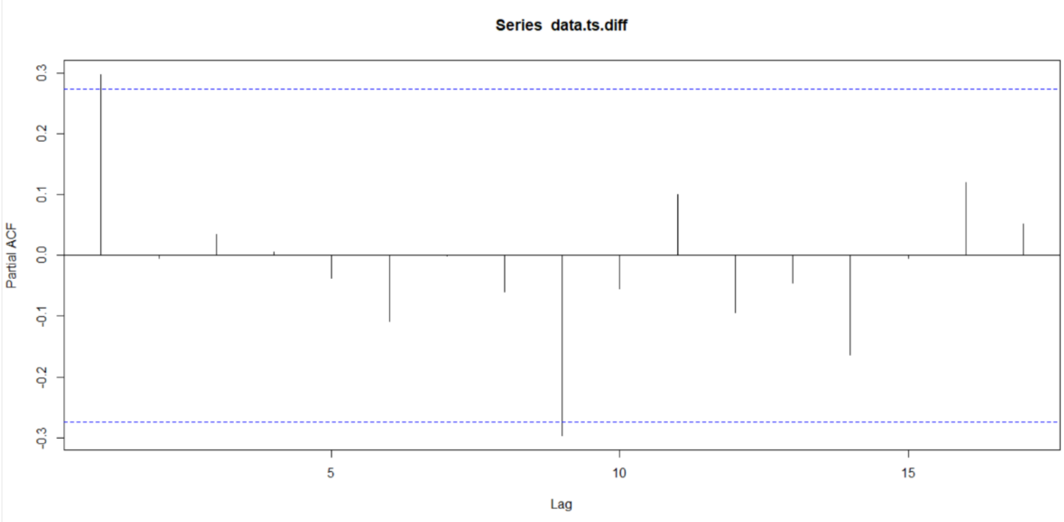
```
#Taking the first difference.
data.ts.diff<-diff(data.ts)
plot(data.ts.diff)
acf(data.ts.diff)
pacf(data.ts.diff)
adf.test(data.ts.diff)

#Taking the second difference.
data.ts.diff.diff<-diff(data.ts.diff)
plot(data.ts.diff.diff)
acf(data.ts.diff.diff)
pacf(data.ts.diff.diff)
adf.test(data.ts.diff.diff)
```

First Difference Output

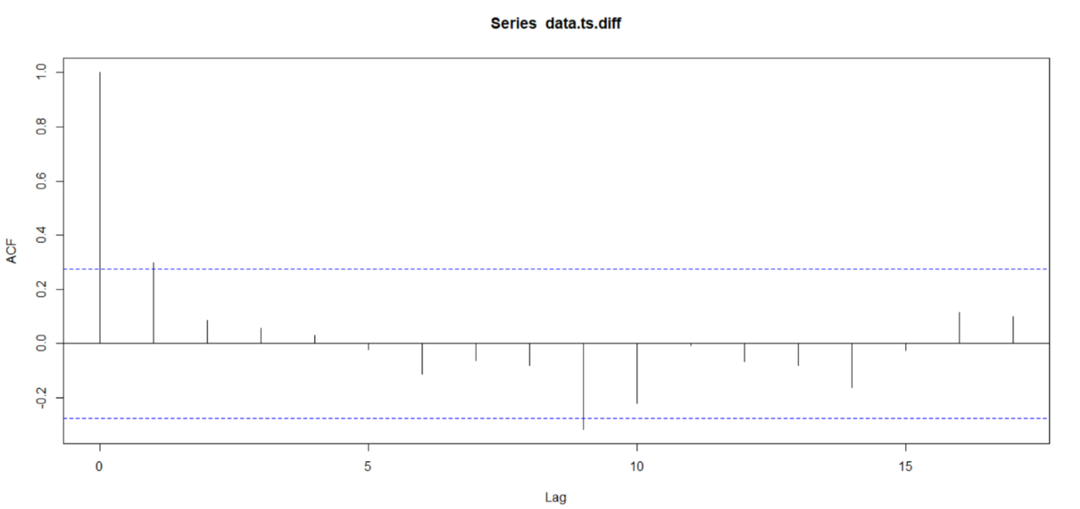
1. **Time Series Plot**



2. Autocorrelation Function Plot



3. Partial Autocorrelation Function Plot



4. Augmented Dickey Fuller Test

```
        Augmented Dickey-Fuller Test

data:  data.ts.diff
Dickey-Fuller = -2.1918, Lag order = 3, p-value = 0.4973
alternative hypothesis: stationary
```

The Observations & Comments

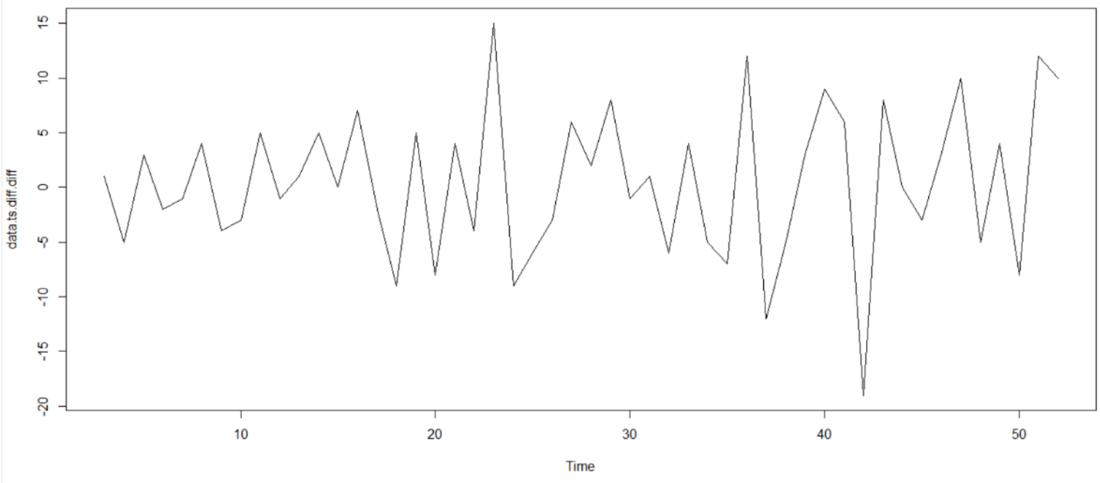
Since the data is showin AR(1) type of dependancy, taking the first difference will make the data stationary. As can be seen from the time series plot of the new data that is obtained by taking the first difference, the mean of the data is near to 0 and the values seems like randomly distributed around the mean 0. In addition to this, the ACF and PACF plots do not contain slow decay and indicate the trend in the new data. But, the plots are not enough to make a conclusion. The Augmented Dickey-Fuller Test has a p-value of 0.4973 which claims that the data is non-stationary. Taking the first difference is not enough. Thus, the second difference should be taken.

Second Difference Output

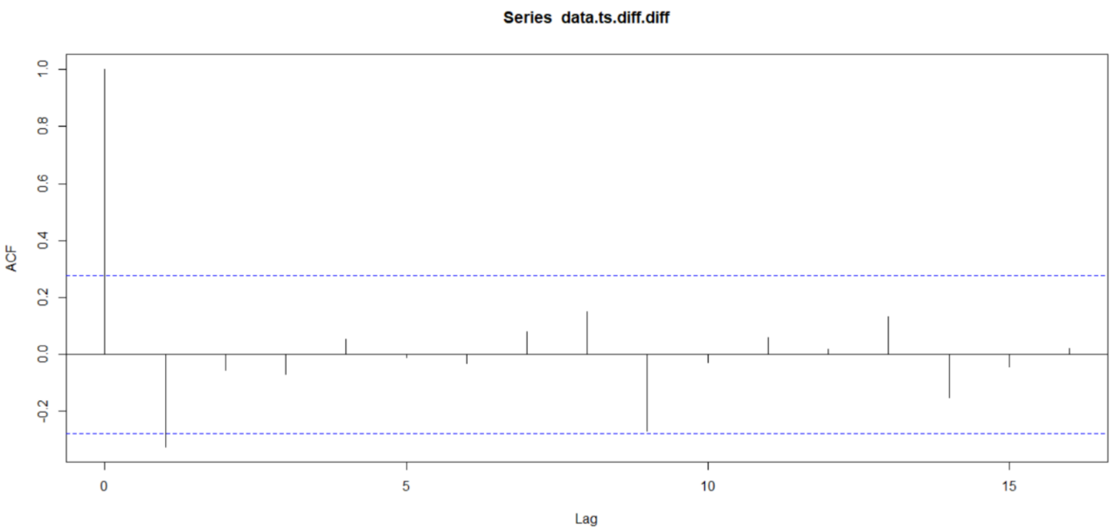
```
        Augmented Dickey-Fuller Test

data:  data.ts.diff.diff
Dickey-Fuller = -4.3792, Lag order = 3, p-value = 0.01
alternative hypothesis: stationary
```

The second difference has a p-value of 0.01 which shows that the data is stationary.

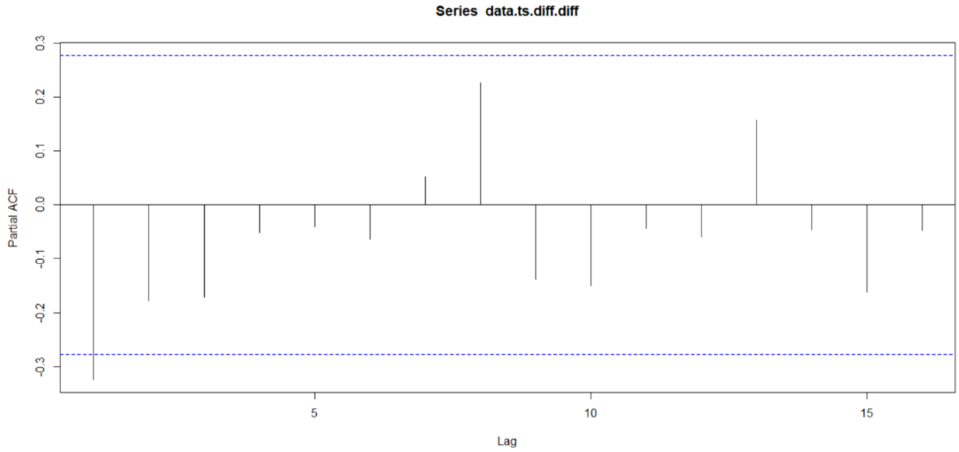
1. Time Series Plot



2. Autocorrelation Plot



3. Partial Autocorrelation Plot



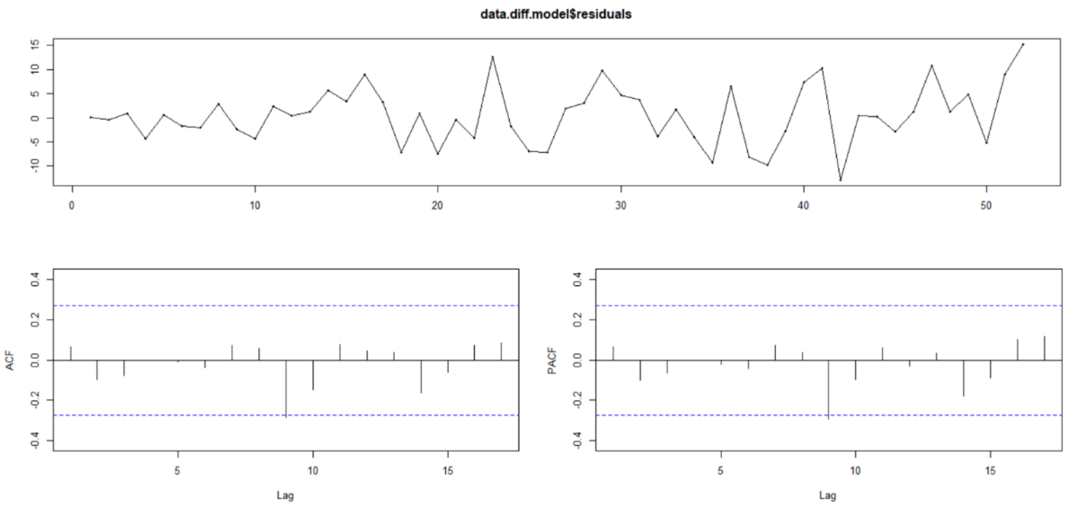
c. Fit an ARIMA model to the IBM Stock Price data. Check the adequacy of the fit.

The ACF cuts off after lag 1. But in PACF plot, there is not any pattern. The ARIMA(0,2,1) model can capture the data patterns and make more accurare forecasts. Thus, I will build ARIMA(0,2,1) model.

The Code

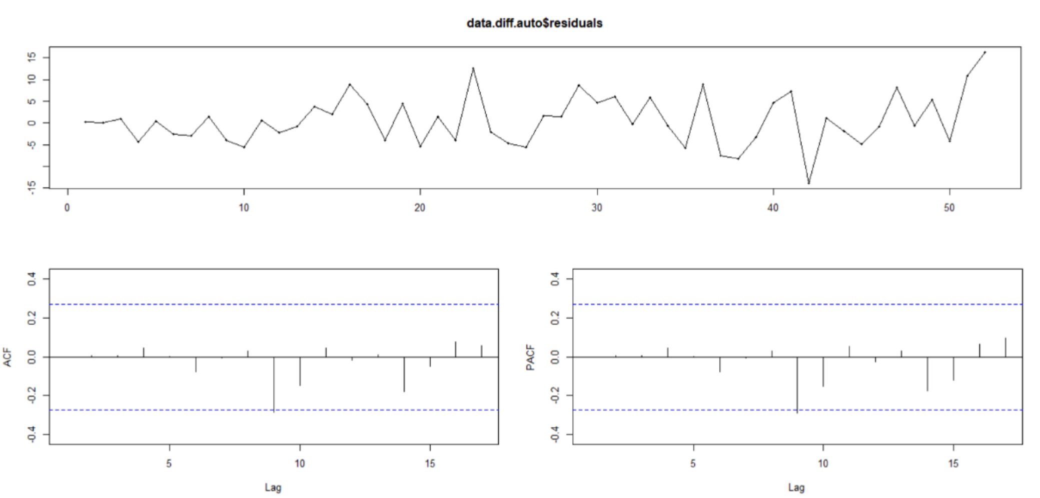
```
#Fit ARIMA(0,2,1) model.
data.diff.model<-Arima(data.ts, order=c(0,2,1))
data.diff.model
tsdisplay(data.diff.model$residuals)
#Fit auto.arima model.
data.diff.auto<-auto.arima(data.ts)
data.diff.auto
tsdisplay(data.diff.auto$residuals)
```

ARIMA(0,2,1) MODEL - OUTPUT




```
Series: data.ts
ARIMA(0,2,1)

Coefficients:
          ma1
      -0.5869
s.e.   0.1640

sigma^2 = 38.46:  log likelihood = -161.89
AIC=327.78   AICc=328.03   BIC=331.6
```

THE AUTO-ARIMA MODEL - OUTPUT




```
Series: data.ts
ARIMA(1,1,0)

Coefficients:
         ar1
      0.3687
s.e.  0.1457

sigma^2 = 34.21:  log likelihood = -162.02
AIC=328.03   AICc=328.28   BIC=331.89
```

The Results

The AIC, AICc, and BIC models are compared, the ARIMA(0,2,1) model is very slightly better than the second model. I think it can be said that the model is adequate.

d. Forecast the IBM stock price for the first week of January of the next year.

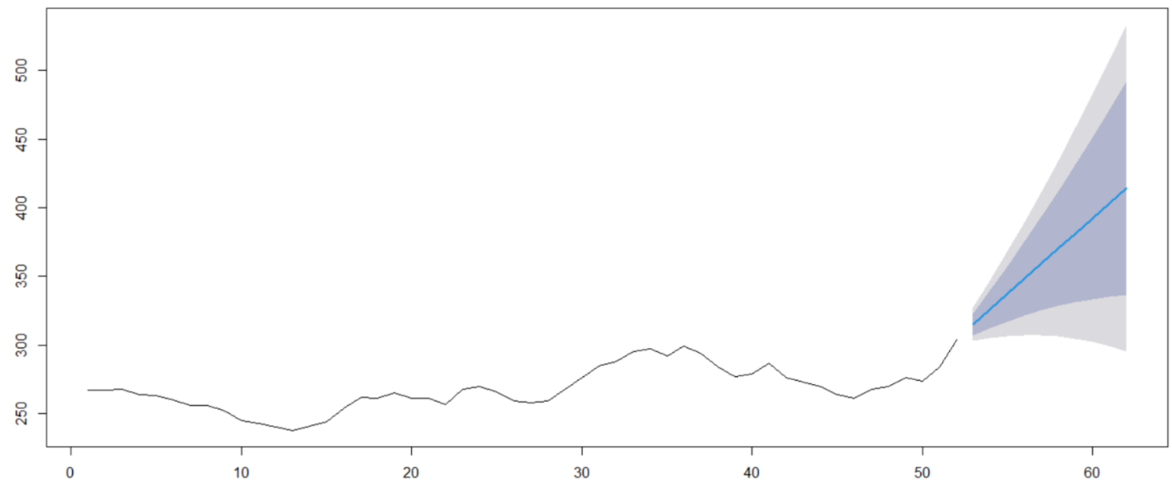
The Code

```
#Forecast.
data.ts.forecast<-forecast(data.diff.model)
data.ts.forecast
plot(data.ts.forecast)
```

The Output

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
53	315.0489	307.1015	322.9962	302.8945	327.2033
54	326.0978	312.3395	339.8560	305.0563	347.1392
55	337.1466	317.1479	357.1454	306.5613	367.7320
56	348.1955	321.4243	374.9668	307.2525	389.1386
57	359.2444	325.1695	393.3193	307.1313	411.3575
58	370.2933	328.4040	412.1826	306.2291	434.3575
59	381.3422	331.1515	431.5329	304.5821	458.1022
60	392.3911	333.4349	451.3472	302.2254	482.5568
61	403.4399	335.2751	471.6048	299.1908	507.6891
62	414.4888	336.6908	492.2868	295.5070	533.4706

Forecasts from ARIMA(0,2,1)



The Result

The forecast for the first week of January of the next year is 315.0489 (point forecast).

3. For the ARMA(1,2) model: Yt = 0.8Yt−1 + Et + 0.7Et−1 + 0.6Et−2 , show that ρk = 0.8ρk−1 for k ≥ 3 ρ2 = 0.8ρ1 + 0.6σ2 /γ0, where, as usual, "t ∼ iid(0, σ2) Also, find γ0 = V ar(Yt).

ARMA(1,2)

③ $Y_t = \phi_1 Y_{t-1} + w_2 \varepsilon_{t-2} + w_1 \varepsilon_{t-1} + \varepsilon_t$

$$E[Y_t Y_{t-k}] = E[\phi_1 Y_{t-1} Y_{t-k}] + \underbrace{E[w_2 \varepsilon_{t-2} Y_{t-k}]}_{0} + \underbrace{E[w_1 \varepsilon_{t-1} Y_{t-k}]}_{0} + \underbrace{E[\varepsilon_t Y_{t-k}]}_{0}$$

$$\left.\begin{array}{l} E[\varepsilon_{t-2} Y_{t-k}] = 0 \\ E[\varepsilon_{t-1} Y_{t-k}] = 0 \\ E[\varepsilon_t Y_{t-k}] = 0 \end{array}\right\}$$ while $k \geq 3$ because Y_{t-k} depends on $\varepsilon_{t-3}, \varepsilon_{t-4}, \varepsilon_{t-5}, \ldots$

Thus; $\gamma_k = \phi_1 \gamma_{k-1}$

$$\frac{\gamma_k}{\gamma_0} = \phi_1 \frac{\gamma_{k-1}}{\gamma_0} \Rightarrow \boxed{\rho_k = \phi_1 \rho_{k-1}} \rightarrow \rho_k = 0.8 \rho_{k-1} \text{ for } k \geq 3.$$

$\gamma_2 = \phi_1 \gamma_1 + w_2 E[Y_{t-2}\varepsilon_{t-2}] \rightarrow Y_{t-2}$ contains ε_{t-2}

$\gamma_1 = \phi_1 \gamma_0 + w_2 E[Y_{t-1}\varepsilon_{t-2}] + w_1 E[Y_{t-1}\varepsilon_{t-1}] \rightarrow Y_{t-1}$ contains ε_{t-1} & ε_{t-2}

$\gamma_0 = \phi_1 \gamma_{-1} + w_2 E[Y_t \varepsilon_{t-2}] + w_1 E[Y_t \varepsilon_{t-1}] + E[Y_t \cdot \varepsilon_t] \rightarrow Y_t$ has all of them.

$Y_t = \phi_1 Y_{t-1} + w_2 \varepsilon_{t-2} + w_1 \varepsilon_{t-1} + \varepsilon_t$

Y_{t-1} does not include ε_t. It depends on $\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots$

$$E[Y_t \varepsilon_t] = \phi_1 \underbrace{E[\varepsilon_t Y_{t-1}]}_{} + w_2 \underbrace{E[\varepsilon_{t-2}\varepsilon_t]}_{} + w_1 \underbrace{E[\varepsilon_{t-1}\varepsilon_t]}_{} + \underbrace{E[\varepsilon_t \cdot \varepsilon_t]}_{}$$

$$= 0 + 0 + 0 + \sigma^2 = \sigma^2$$

$$E[Y_t \varepsilon_{t-1}] = \phi_1 \underbrace{E[\varepsilon_{t-1} Y_{t-1}]}_{} + w_2 \underbrace{E[\varepsilon_{t-2}\varepsilon_{t-1}]}_{} + w_1 \underbrace{E[\varepsilon_{t-1}\varepsilon_{t-1}]}_{} + \underbrace{E[\varepsilon_t \varepsilon_{t-1}]}_{}$$

$$= \phi_1 \cdot \sigma^2 + 0 + w_1 \cdot \sigma^2 + 0$$

$$= \sigma^2(\phi_1 + w_1)$$

$$E[Y_t \varepsilon_{t-2}] = \phi_1 \underbrace{E[Y_{t-1}\varepsilon_{t-2}]}_{} + w_2 \underbrace{E[\varepsilon_{t-2}\varepsilon_{t-2}]}_{} + w_1 \underbrace{E[\varepsilon_{t-1}\varepsilon_{t-2}]}_{} + \underbrace{E[\varepsilon_t \varepsilon_{t-2}]}_{}$$

$$= \phi_1 [\sigma^2(\phi_1 + w_1)] + w_2\sigma^2 + 0 + 0$$

$$= \phi_1^2 \sigma^2 + \phi_1 w_1 \sigma^2 + w_2 \sigma^2$$

$$= \sigma^2(\phi_1^2 + \phi_1 w_1 + w_2)$$

$$\gamma_0 = \phi_1 \gamma_{-1} + w_2 \cdot [\sigma^2(\phi_1^2 + \phi_1 w_1 + w_2)] + w_1 \sigma^2(\phi_1 + w_1) + \sigma^2$$

$\gamma_{-1} \rightarrow \gamma_1$

$$\gamma_1 = \phi_1 \gamma_0 + w_2 \sigma^2(\phi_1 + w_1) + w_1 \sigma^2$$

$$= 0.8\gamma_0 + 0.6\sigma^2(0.8 + 0.7) + 0.7\sigma^2$$

$$= 0.8\gamma_0 + 1.6\sigma^2$$

$$\gamma_0 = 0.8\gamma_1 + 0.6[\sigma^2(0.8^2 + 0.8 \cdot 0.7 + 0.6)] + 0.7\sigma^2(0.8 + 0.7) + \sigma^2$$

$$= 0.8\gamma_1 + 1.08\sigma^2 + 1.05\sigma^2 + \sigma^2$$

$$= 0.8\gamma_1 + 3.13\sigma^2$$

$$= 0.8(0.8\gamma_0 + 1.6\sigma^2) + 3.13\sigma^2$$

$$\gamma_0 = 0.64\gamma_0 + 4{,}41\sigma^2$$

$$\boxed{\gamma_0 = 12{,}25\sigma^2}$$

$$\boxed{\begin{aligned}\rho_2 &= \phi_1 \frac{\gamma_1}{\gamma_0} + \frac{w_2 \cdot \sigma^2}{\gamma_0} \\ &= 0.8 \cdot \rho_1 + \frac{0.6\sigma^2}{\gamma_0}\end{aligned}}$$