pstat274_lab06_aoxu

AO XU

2022-11-04

1.

(a) Summarize how you would carry out the following steps in time series analysis. You can describe briefly by words, or write down R commands you would use to implement these steps.

Assume we have a dataset abcd.

Step 1 Data processing

op <- par(mfrow=c(1,2))

acf(abcd)

pacf(abcd)

par(op)

abcd.diff <- diff(abcd,1)

ts.plot(abcd.diff, main = “De-trended data”)

Step 2 Model identification

op <- par(mfrow=c(1,2))

acf(abcd.diff)

pacf(abcd.diff)

par(op)

Step 3 Model estimation

(fit.ar <- ar(abcd.diff, method=“yw”))

95% CI for phi1

ar1.se <- sqrt(fit.ar$asy.var.coef)

c(fit.ar$ar - 1.96 * ar1.se, fit.ar$ar + 1.96*1.96*ar1.se)

Step 4 Model selection

library(qpcR)

Calculate AICc for ARMA models with p and q running from 0 to 5

```
aiccs <- matrix(NA, nr = 6, nc = 6)
dimnames(aiccs) = list(p=0:5, q=0:5)
for(p in 0:5)
{
for(q in 0:5)
{
aiccs[p+1,q+1] = AICc(arima(abcd.diff, order = c(p,0,q), method="ML"))
}
}
aiccs
(aiccs==min(aiccs))
```

Step 5 Model diagnostics

Pick AR(1) and perform residual analysis:

```
fit = arima(abcd, order=c(1,1,0), method="ML")
```

Test for independence of residuals

```
Box.test(residuals(fit), type="Ljung")
```

Test for normality of residuals

```
shapiro.test(residuals(fit))
ts.plot(residuals(fit),main = "Fitted Residuals")
par(mfrow=c(1,2),oma=c(0,0,2,0))
```

Plot diagnostics of residuals

```
op <- par(mfrow=c(2,2))
```

acf

```
acf(residuals(fit),main = "Autocorrelation")
```

pacf

```
pacf(residuals(fit),main = "Partial Autocorrelation")
```

Histogram

```
hist(residuals(fit),main = "Histogram")
```

q-q plot

```
qqnorm(residuals(fit))
qqline(residuals(fit),col ="blue")
```

Add overall title

title(“Fitted Residuals Diagnostics”, outer=TRUE)

par(op)

Step 6 Data forecast

Predict 10 future observations and plot

mypred <- predict(fit, n.ahead=10)

ts.plot(abcd, xlim=c(0,89))

points(79:88,mypred$pred)

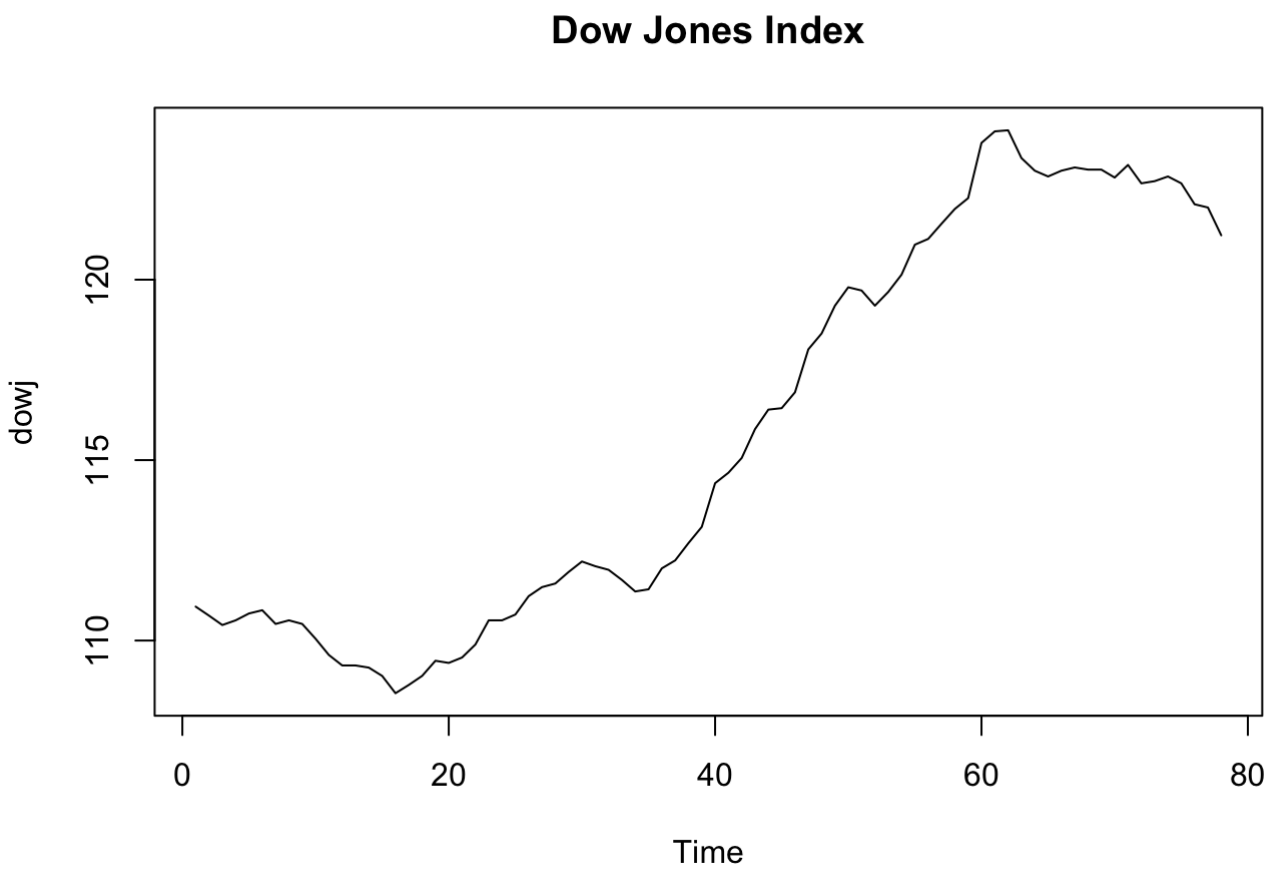
lines(79:88,mypred$pred + 1.96 * mypred$se,lty=2)

lines(79:88,mypred$pred - 1.96 * mypred$se,lty=2)

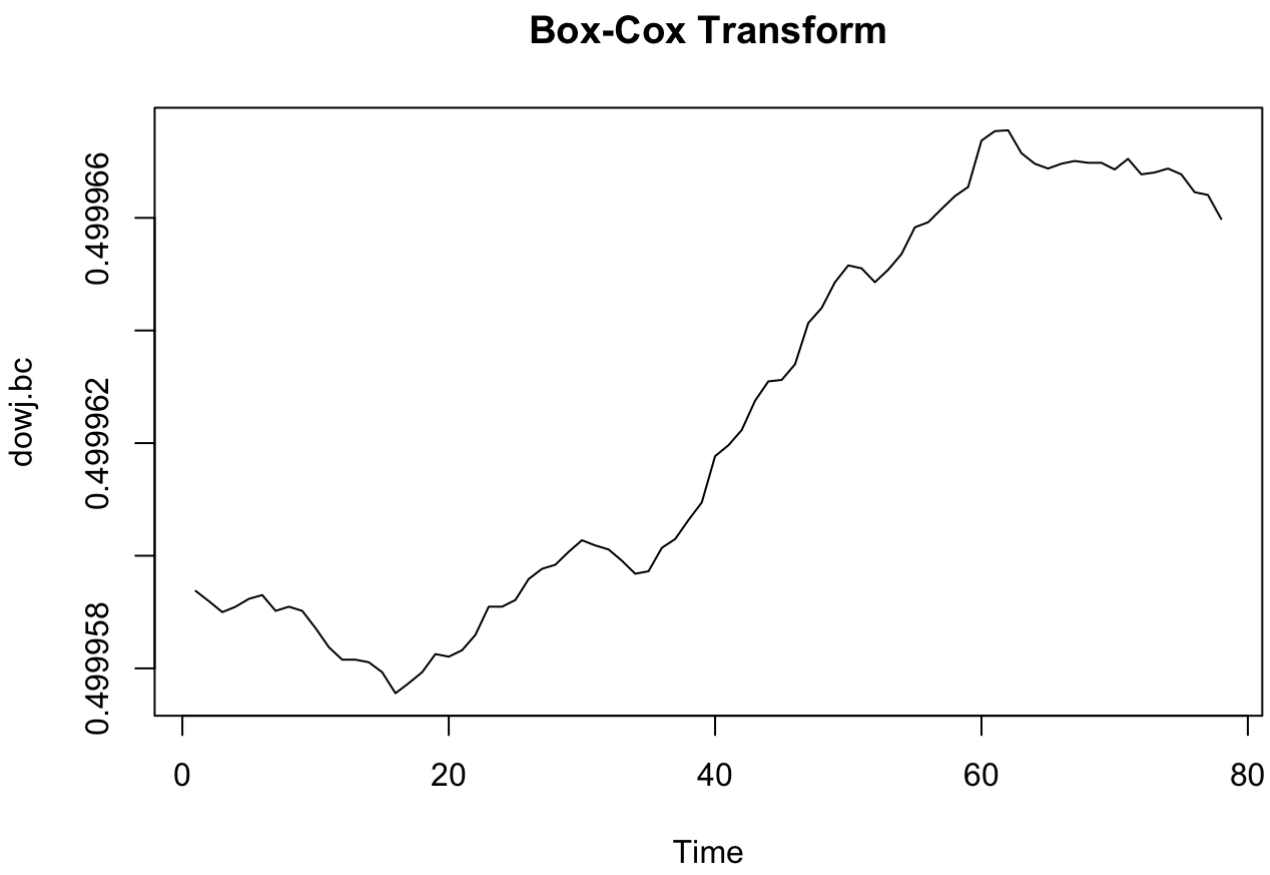
(b) Review this week’s lab material, Dow Jones Index question part 3) (‘Make the data stationary’). Is dierencing once at lag 1 sucient to make the data stationary? If yes, justify it. If no, try to make it stationary. Please write related R Codes.

```
library(MASS)
# Load data
dowj_data <- scan("dowj.txt")
```

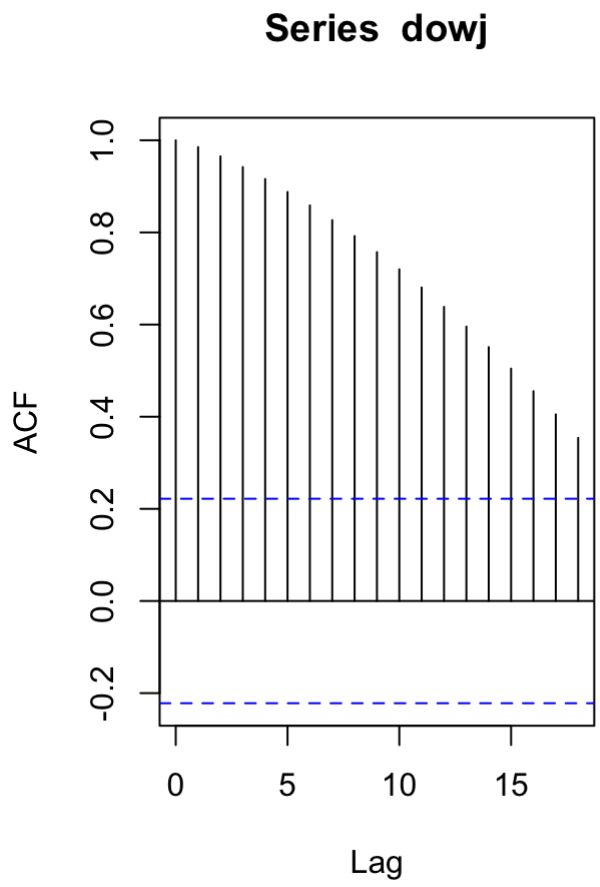
```
dowj <- ts(dowj_data)
# Plot data
ts.plot(dowj,main = "Dow Jones Index")
```

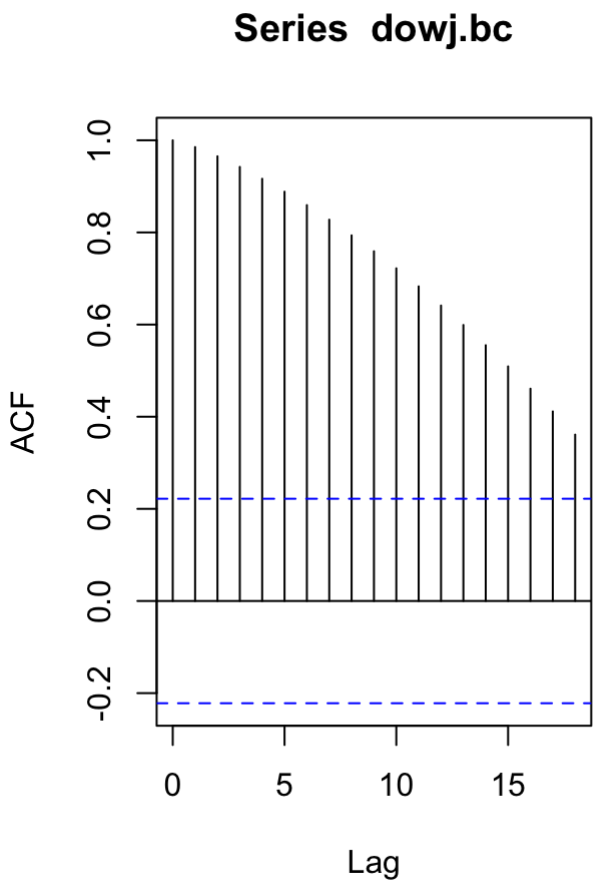



```
t = 1:length(dowj)
bcTranform = boxcox(dowj~t, plotit=FALSE)
lambda = bcTranform$x[which(bcTranform$y==max(bcTranform$y))]
dowj.bc = (1/lambda)*(dowj^lambda-1)
ts.plot(dowj.bc,main="Box-Cox Transform")
```

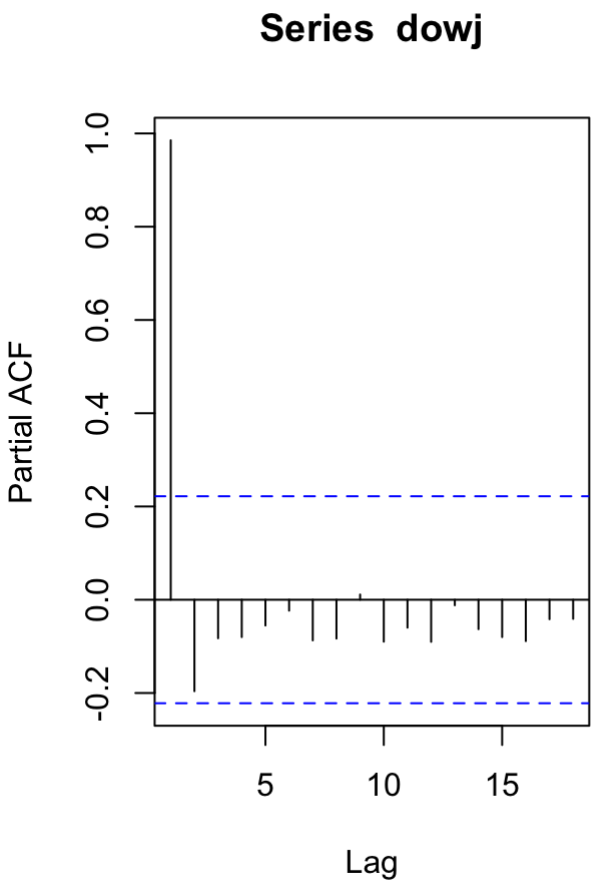


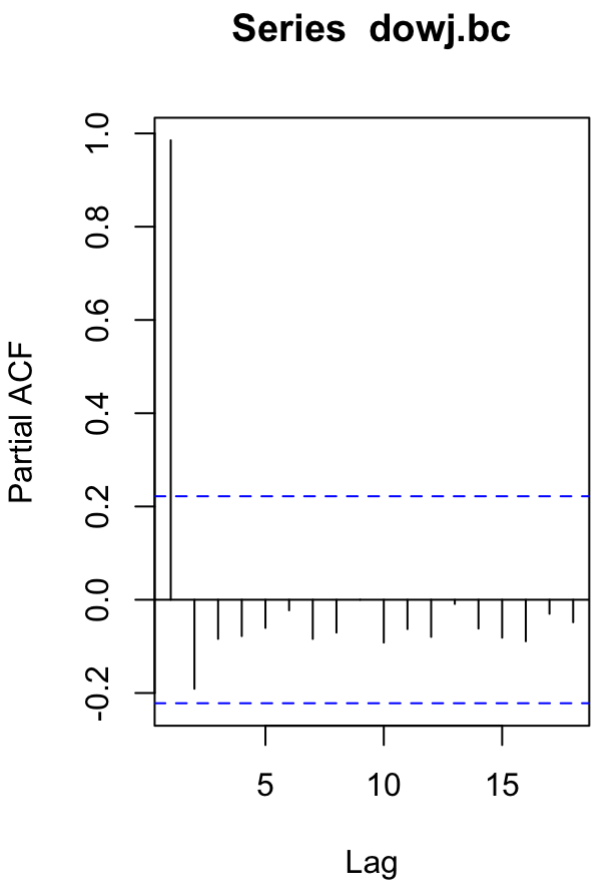

```
op <- par(mfrow=c(1,2))
acf(dowj)
acf(dowj.bc)
```



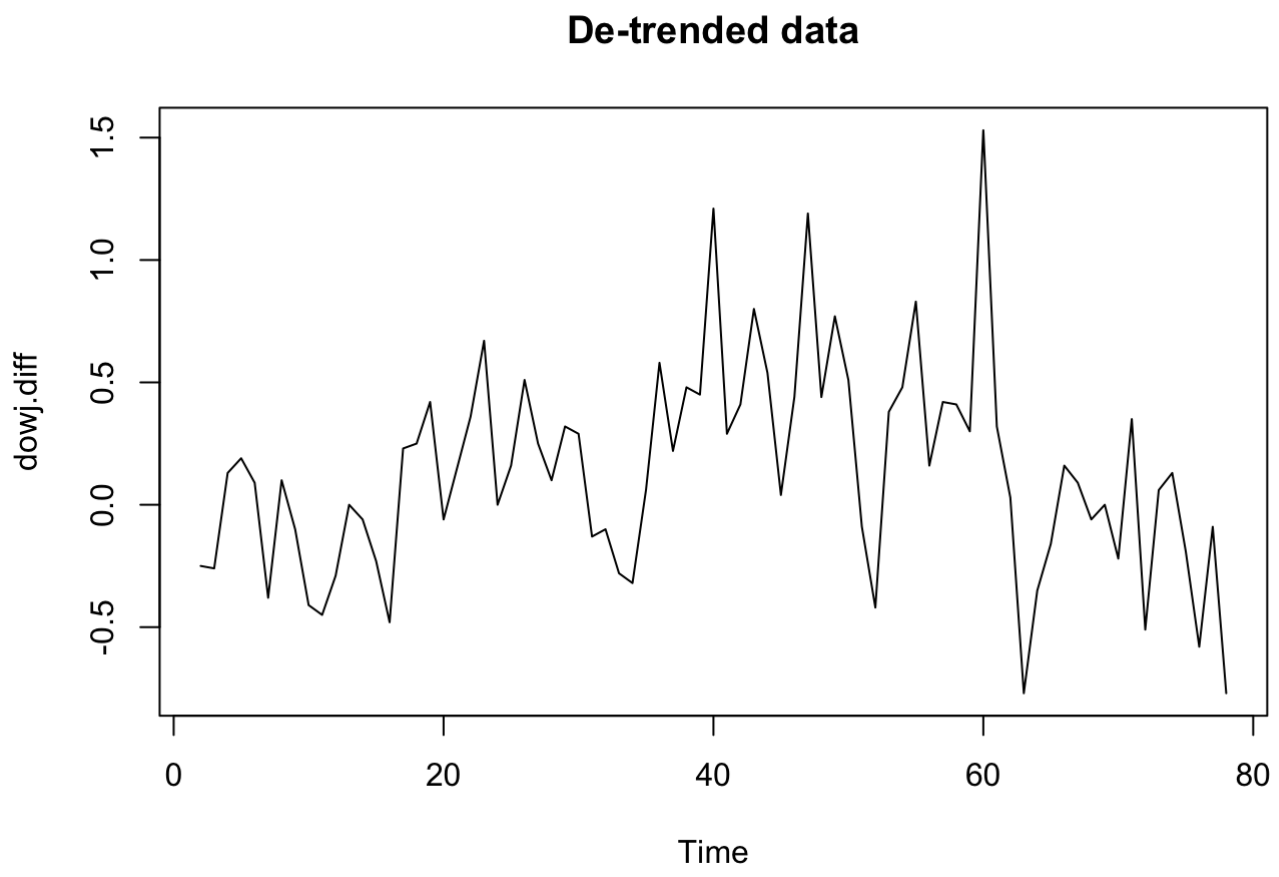



```
pacf(dowj)
pacf(dowj.bc)
```

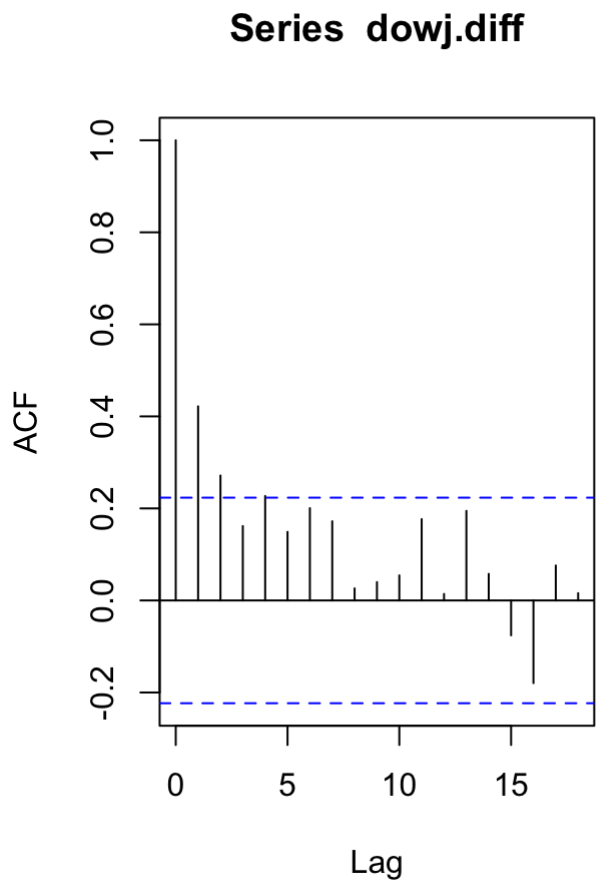


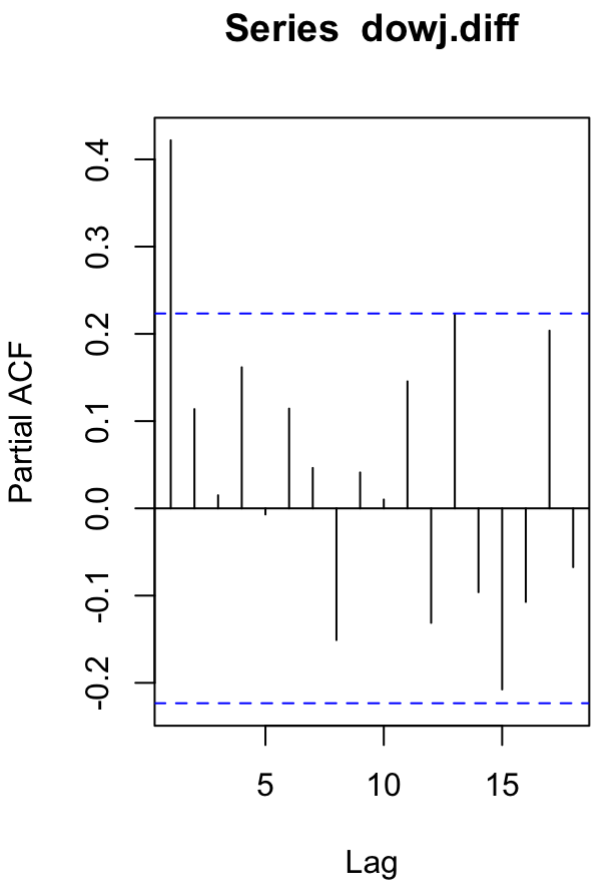



```
par(op)
dowj.diff <- diff(dowj,1)
ts.plot(dowj.diff, main = "De-trended data")
```




```
op <- par(mfrow=c(1,2))
acf(dowj.diff)
pacf(dowj.diff)
```

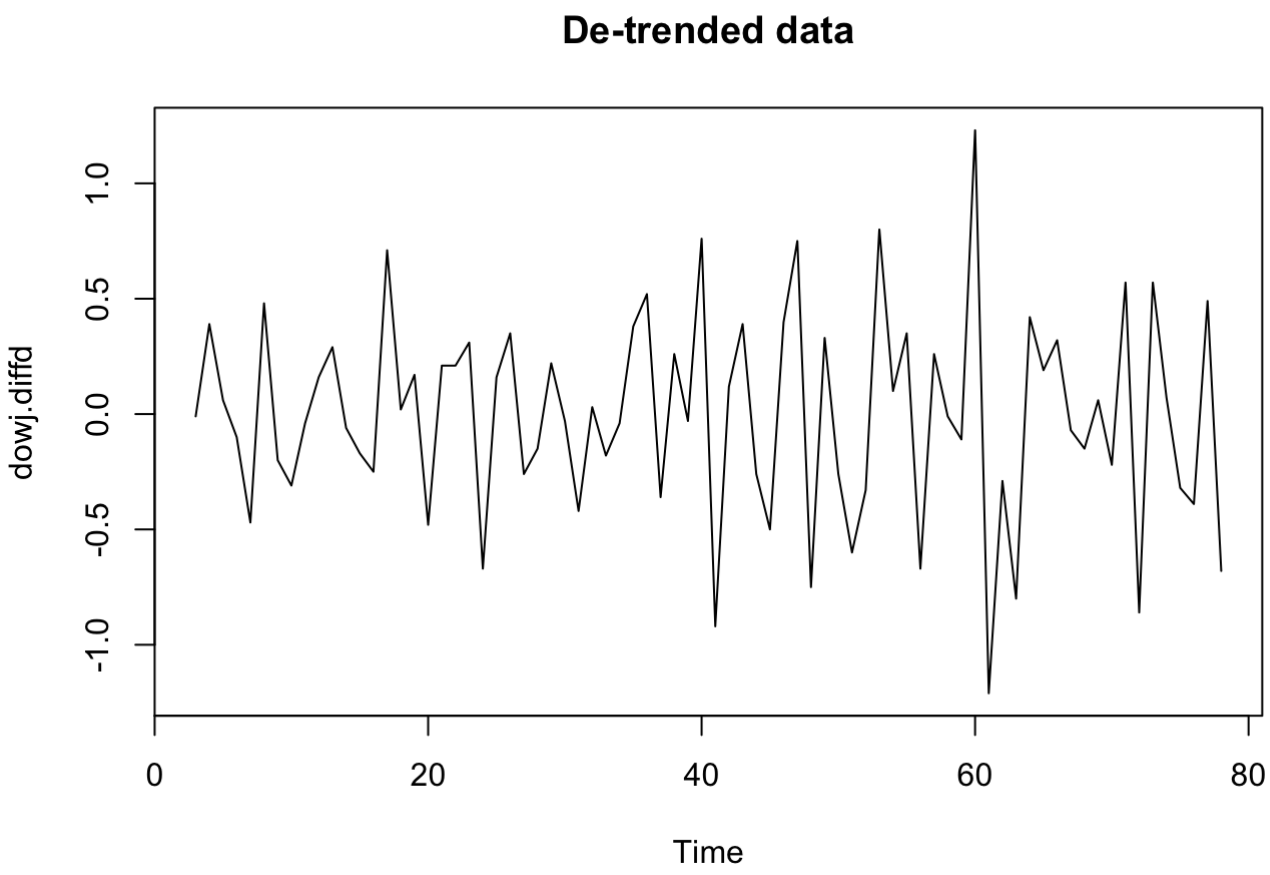




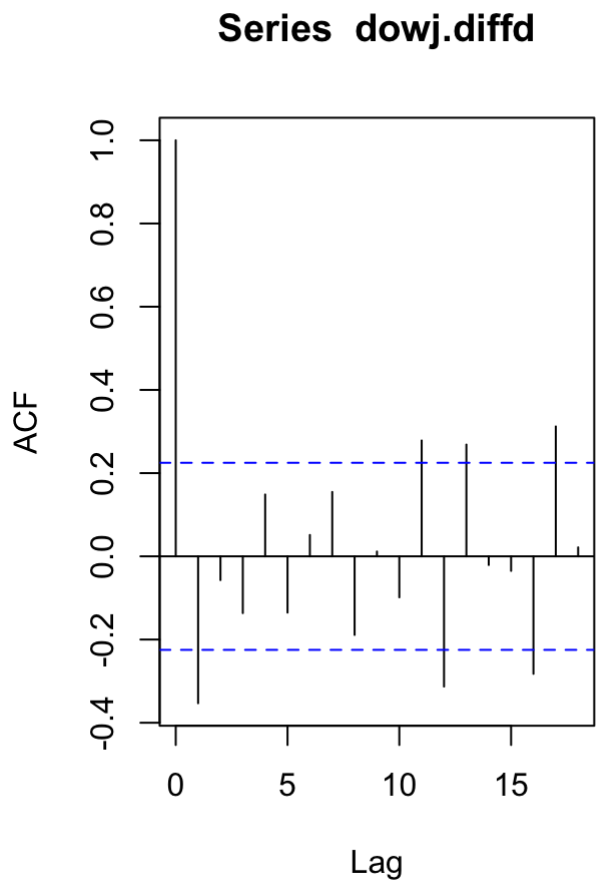

```
var(dowj.diff)
```

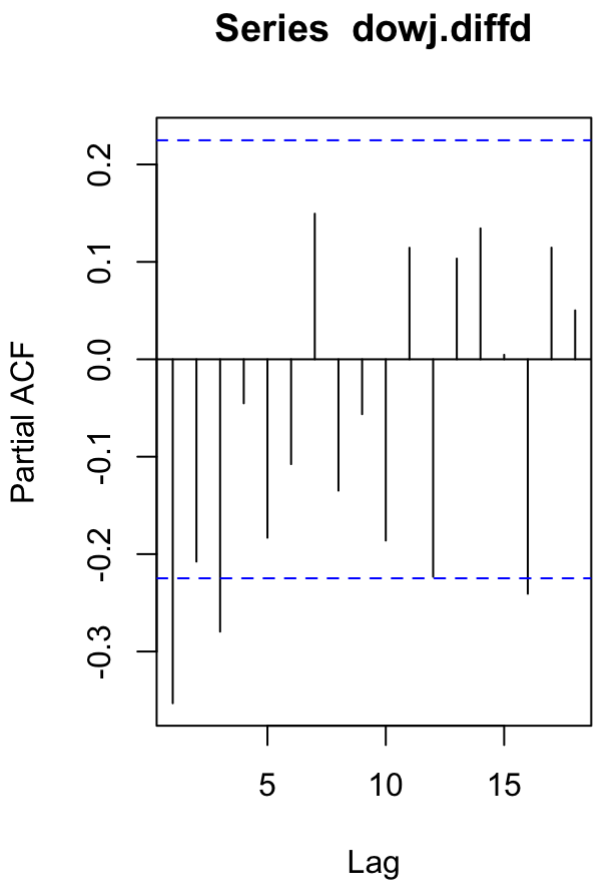
```
## [1] 0.1822866
```

```
dowj.diffd <- diff(dowj.diff,1)
ts.plot(dowj.diffd, main = "De-trended data")
```




```
op <- par(mfrow=c(1,2))
acf(dowj.diffd)
pacf(dowj.diffd)
```






```
var(dowj.diffd)
```

```
## [1] 0.2006806
```

So differencing once at lag 1 is sufficient to make the data stationary.