

A2 Insu Yoon

```
library(ggplot2)
library(forecast)

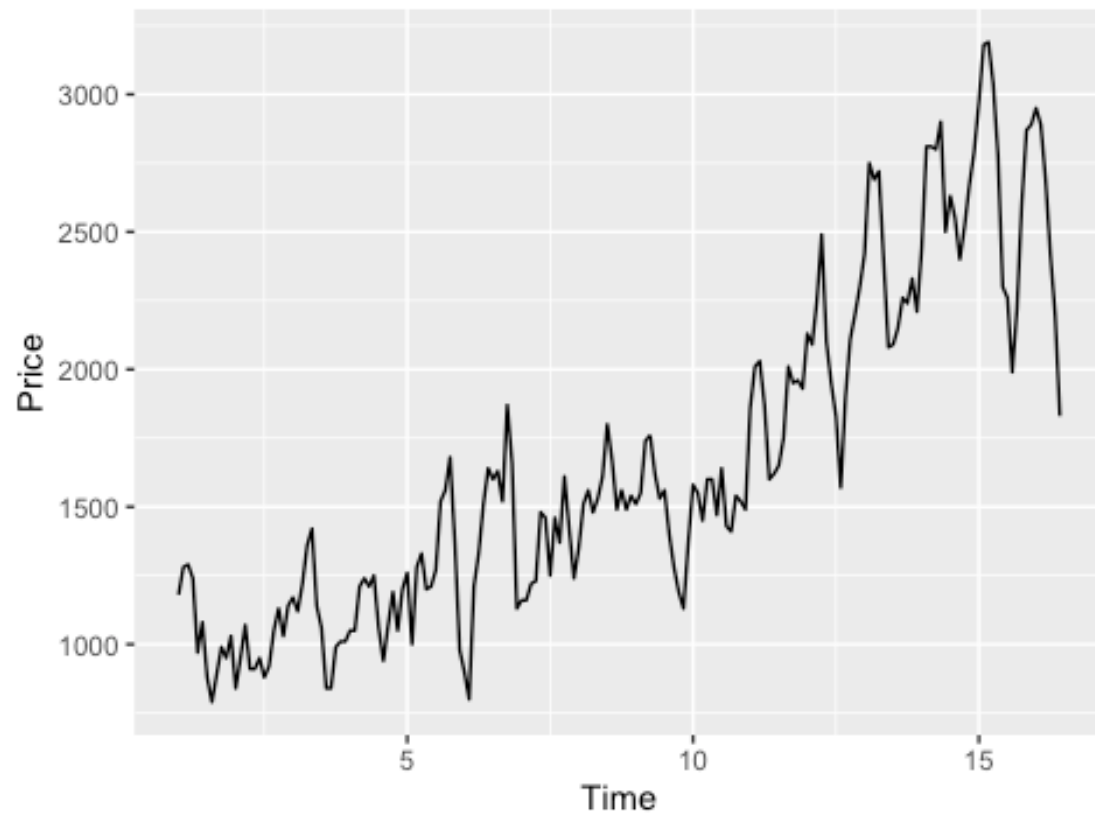
## Registered S3 method overwritten by 'quantmod':
##   method      from
##   as.zoo.data.frame zoo

data<-read.csv("1003180933.csv")
fit = arima( x = data[2], order = c(1,0,1) )
prd = predict( fit, n.ahead = 12 )
write.table( prd$pred, file = "~/Desktop/insu/STA457/A2\\data.csv", sep =
",", # export to csv
            row.names = FALSE, col.names = FALSE)
```

1. A discussion of the characteristics of the time series (e.g. trend, seasonality, stationarity)

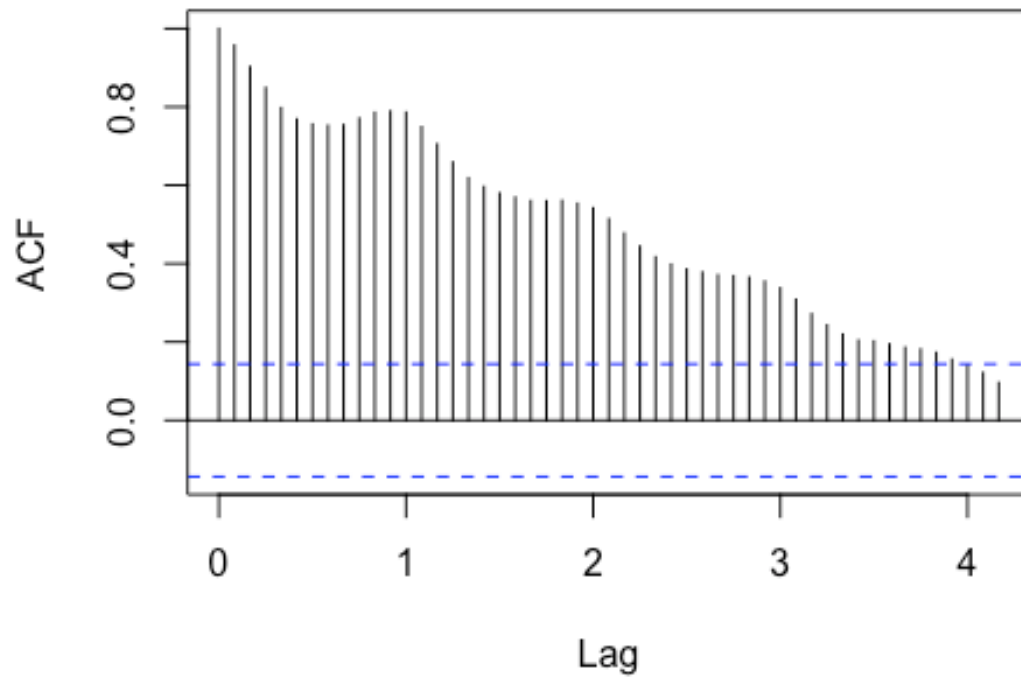
```
y<-ts(data$French.Potato.Prices,frequency=12)
autoplot(y)+
  ggtitle("French Potato price")+
  xlab("Time")+
  ylab("Price")
```

French Potato price

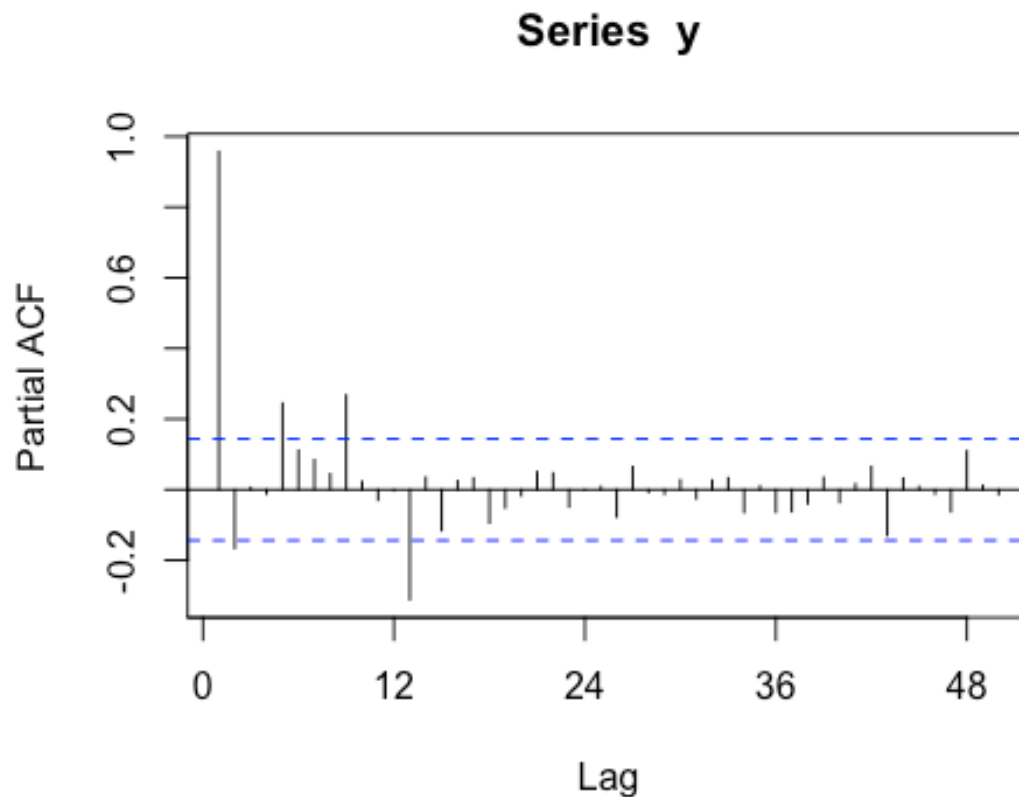


```
acf(y,type="correlation",lag=50,plot = TRUE)
```

Series y



```
Pacf(y,lag.max = 50,plot=TRUE)
```

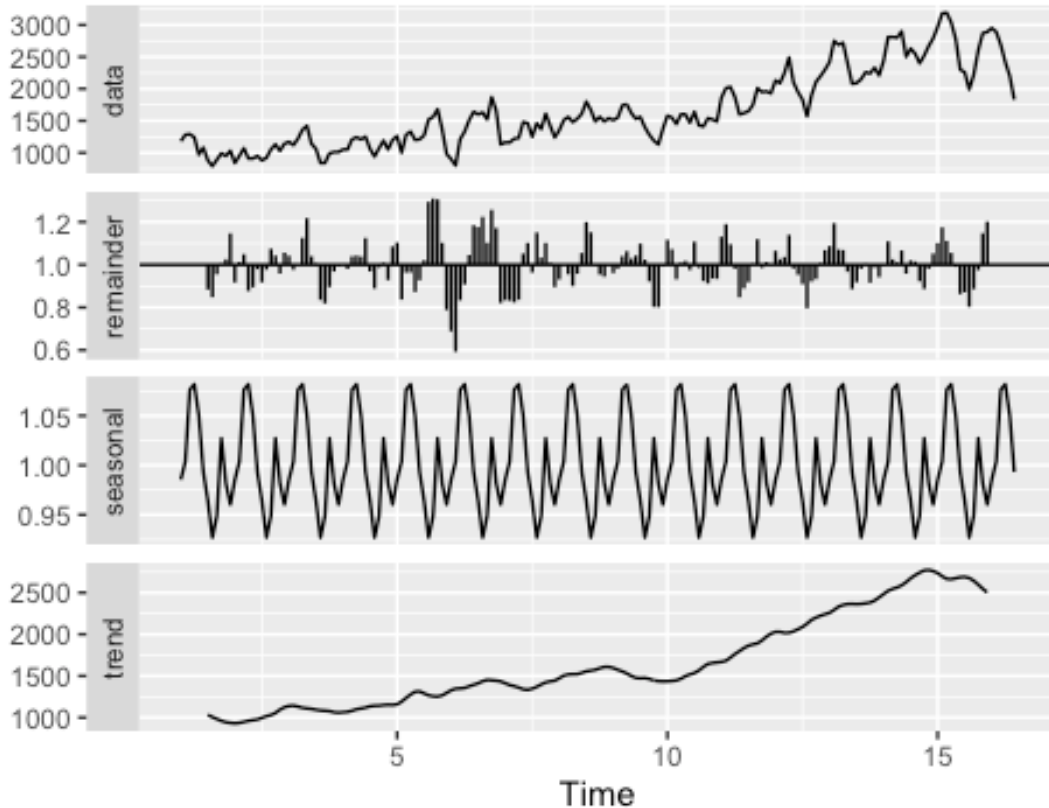


I can see the trend of increasing of French potato price and there is a seasonal pattern that increases in size as the level of series increase and there is a drop at the start of each year. In ACF plot, It is shown that the large and positive acf in the small lags so that we can see that the series have a trend, moreover, the scalloped shape is found in the acf of French potato prices which indicates the seasonality of the series. To check stationarity, the we can check that autocorrelations drop to zero relatively fast for large lags but, in this case, we cannot find it so we can argue that this time series is not stationary.

2. An explanation of any data preprocessing you had to do. Because I can detect both seasonality and trend from the French potato price and the time series model we got is not stationary. We can implement the classical multiplicative decomposition to detrend the mean and the adjust the seasonality.
3. The model which you used.

```
dcmp.y<-decompose(y,type="multiplicative")
autoplot(dcmp.y)+
  ggtitle("Decomposed French Potato price")
```

Decomposed French Potato price



We can capture seasonal pattern has a period of 1, and the trend of the series is a slow increasing linear at the first and more rapid increasing and it falls at the end. We can argue that the potato price is likely to be fallen.

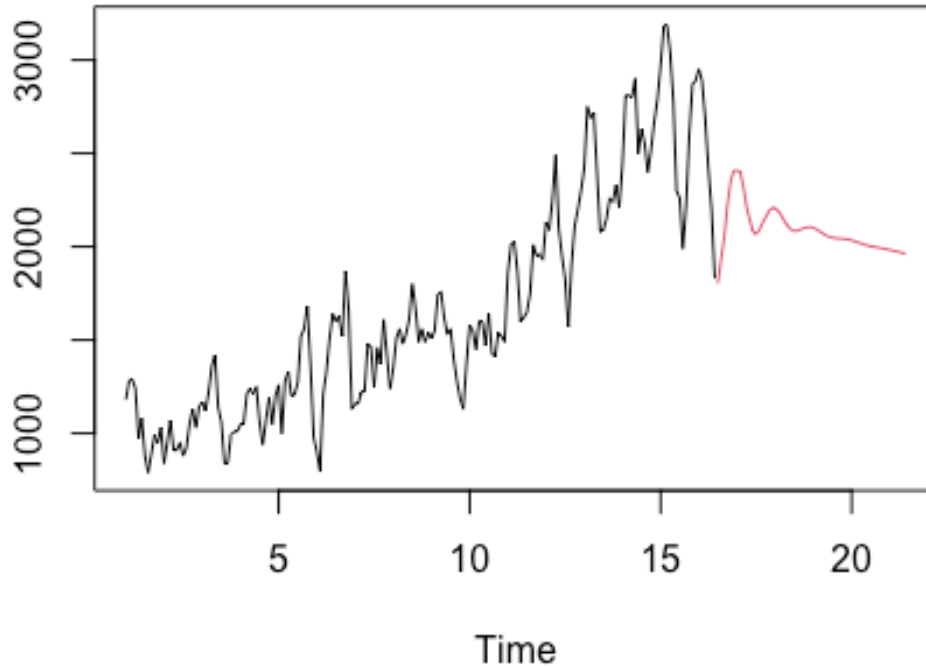
4. A graph of the time series, with your forecasts in a different colour (see graph below for an example)

```
library(signal)

##
## Attaching package: 'signal'

## The following objects are masked from 'package:stats':
##
##   filter, poly

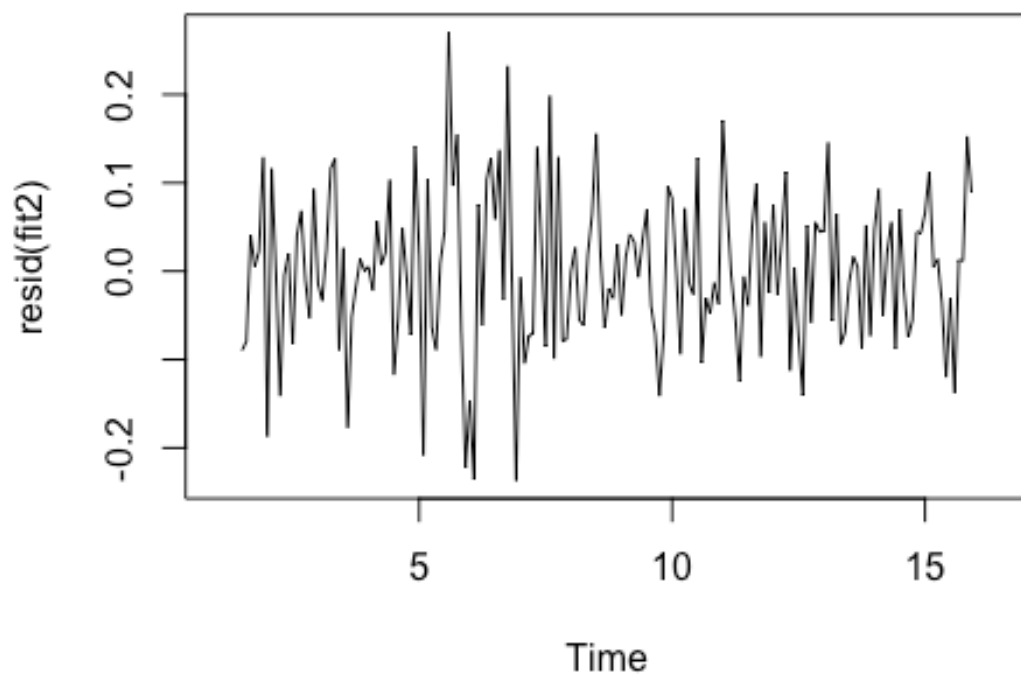
yw.fit <- ar.yw(y, order=10)
potato.forecast = predict(yw.fit, n.ahead=60)
ts.plot(y, potato.forecast$pred, col=1:2, lty=1)
```



5. A

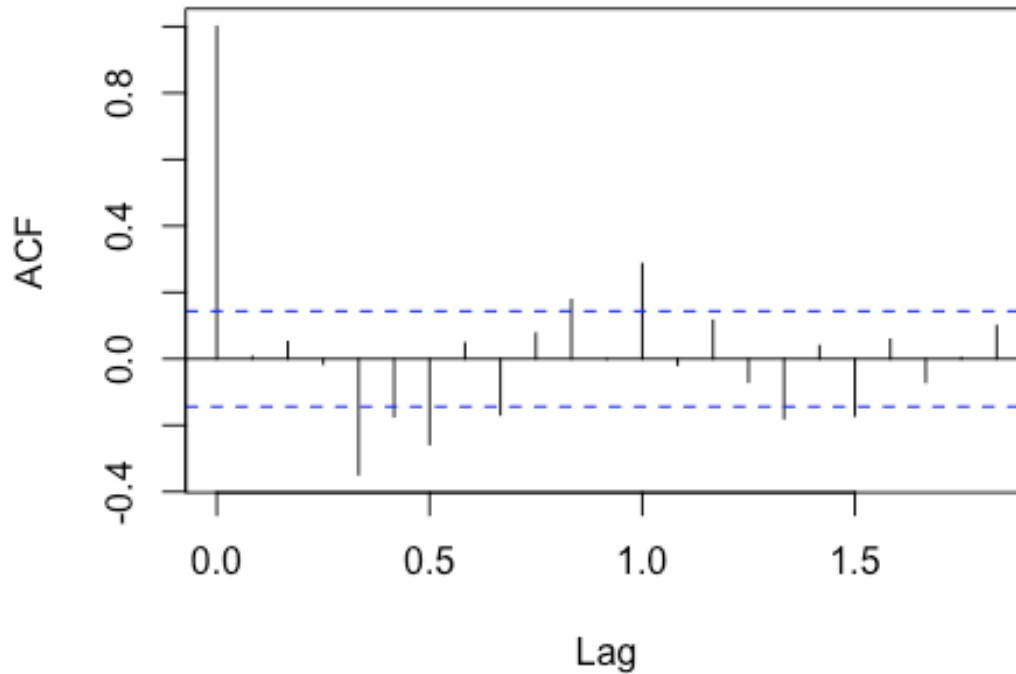
discussion of your model's fit (diagnostics) and limitations. The list above is what your written report must contain, but not an exhaustive list of all that it can contain. If there are any other topics that are worth discussing related to how you forecasted the data, please include them.

```
fit2<-arima(dcmp.y$random,order=c(1,0,1))
num=length(data$French.Potato.Prices)
plot(resid(fit2))
```



```
acf(resid(fit2),na.action = na.pass)
```

Series resid(fit2)



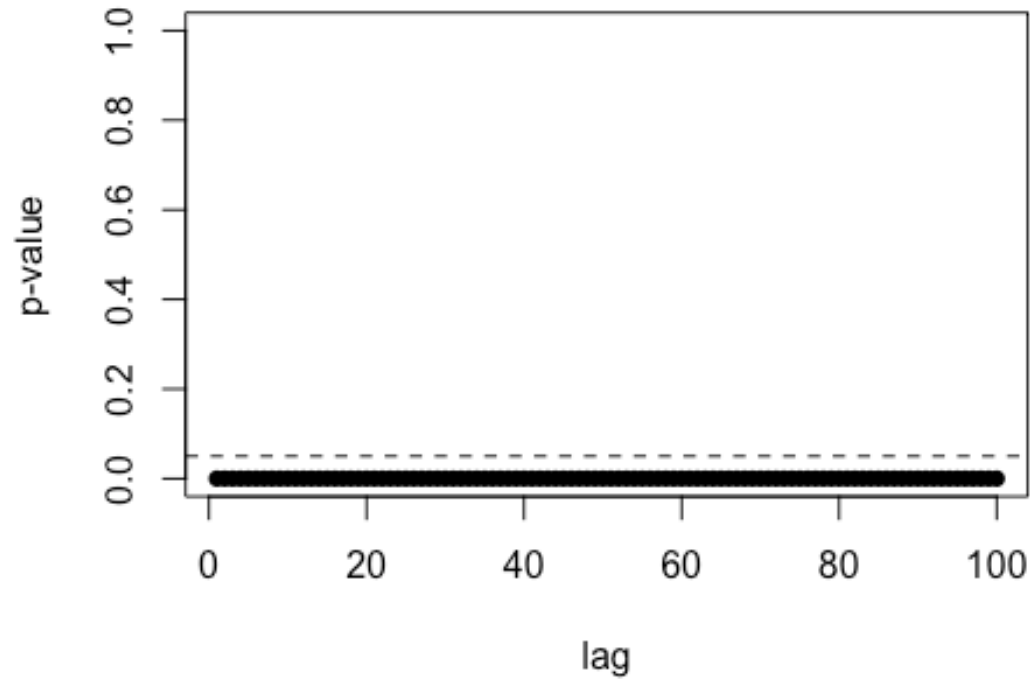
```
AIC(fit2)/num -log(2*pi)
## [1] -3.684567

BIC(fit2)/num-log(2*pi)
## [1] -3.61663

Box.test(dcmp.y$random, lag=50,type="Ljung-Box")
##
## Box-Ljung test
##
## data: dcmp.y$random
## X-squared = 338.52, df = 50, p-value < 2.2e-16

B=NULL; for (i in 1:100)
B = c(B,Box.test(dcmp.y$random,lag = i,type = "Ljung-Box")$p.value)
plot(B,main = "Ljung-Box tests",ylab = "p-value",
      xlab = "lag",pch = 16,
      ylim = c(0,1))
abline(h = .05,lty = 2)
```


Ljung-Box tests



We get low AIC, and BIC values and since the acf of the residuals shows that there is no significant remaining auto-correlation. The low p-value from Ljung-Box test shows that there is serial correlation exists.