Untitled

ERNEST

2024-04-15

Loading of packages

```
library(tseries)

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

library(seastests)
library(forecast)
library(ggplot2)
library(readxl)
```

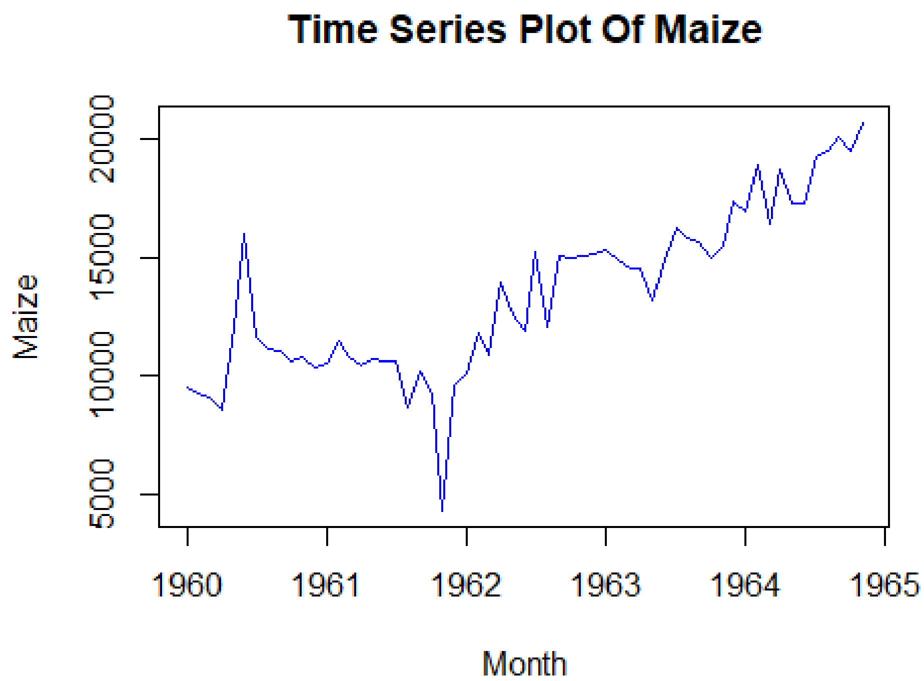
Importing of Data

```
data = read_excel("C:/Users/E. K AGYAPONG/Desktop/FINAL YEAR/Sem 1/Applied
Time Series Analysis (STAT 471)/Lecturer/Practicals_data.xls")
attach(data)
maize = ts(Maize, start = c(1960,1),frequency = 12)
maize

##        Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct   Nov
Dec
## 1960  9536  9283  9039  8537 12062 16026 11640 11089 11047 10643 10758
10358
## 1961 10525 11431 10741 10438 10703 10634 10615  8682 10161  9276  4300
9613
## 1962 10086 11845 10907 13907 12610 11889 15260 12040 15092 14933 15020
15153
## 1963 15285 14850 14557 14578 13150 14900 16272 15794 15613 14994 15437
17371
## 1964 16969 18874 16458 18712 17240 17291 19218 19500 20112 19474 20697
```

Plotting of Data

```
plot(maize, main = "Time Series Plot Of Maize", ylab="Maize", xlab="Month",
col="blue")
```



###Test for Stationarity
ADF Test
H_0:The Maize series data is not stationary
H_1:The Maize series data is Stationary
significance level= 0.05

```
adf.test(maize)

##
##  Augmented Dickey-Fuller Test
##
## data:  maize
## Dickey-Fuller = -1.9609, Lag order = 3, p-value = 0.5905
## alternative hypothesis: stationary
```

Conclusion: Since the p-value(= 0.5905) is greater than the significannce level(=0.05),we fail to reject H_0 and conclude that the maize series data is not stationary.

PP Test
H_0:The Maize series data is not stationary
H_1:The Maize series data is Stationary
significance level= 0.05

```
pp.test(maize)
```

```
##
##  Phillips-Perron Unit Root Test
##
## data:  maize
## Dickey-Fuller Z(alpha) = -25.631, Truncation lag parameter = 3, p-value
## = 0.01151
## alternative hypothesis: stationary
```

Conclusion: Since the p-value(= 0.01151) is less than the significance level(=0.05), we reject H_0 and conclude that the maize series data is stationary.

KPSS Test
H_0:The Maize series data is stationary
H_1:The Maize series data is not Stationary
significance level= 0.05

```
kpss.test(maize)

## Warning in kpss.test(maize): p-value smaller than printed p-value

##
##  KPSS Test for Level Stationarity
##
## data:  maize
## KPSS Level = 1.3192, Truncation lag parameter = 3, p-value = 0.01
```

Conclusion: Since the p-value(= 0.01) is less than the significance level(=0.05), we reject H_0 and conclude that the maize series data is not stationary.
Since ADF and KPSS test confirm that the data is not stationary, we conclude that the Maize series data is not Stationary.

Differencing Of Data

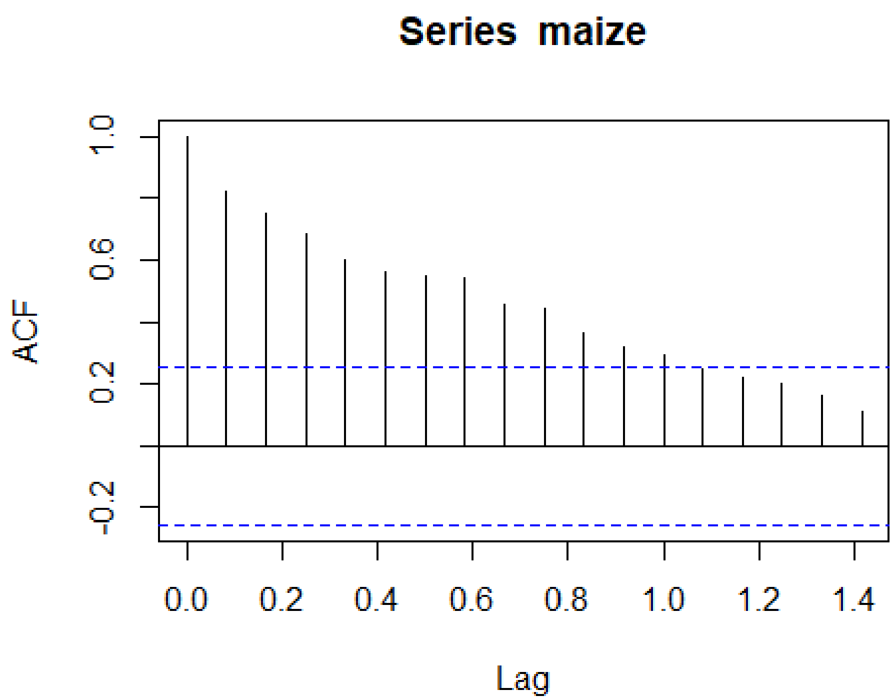
```
ndiffs(maize)

## [1] 1

FirstDiff = diff(maize)
```

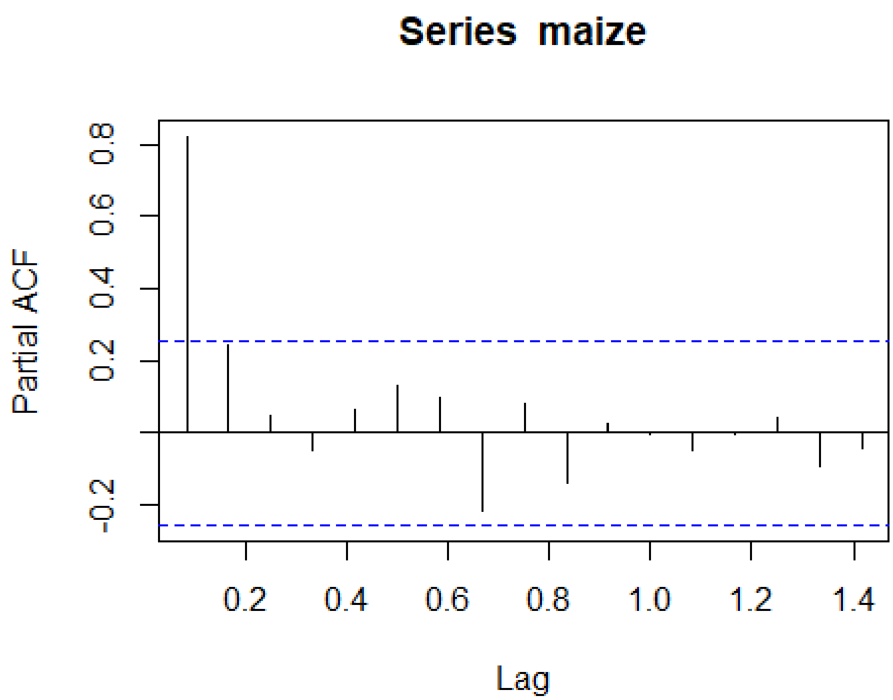
/ Ploting of ACF and PACF

```
acf(maize)
```



The ACF of the Maize series decays exponentially

```
pacf(maize)
```



The PACF of the maize series data cuts off after lag one.
From the Acf and the Pacf plot we conclude that the model for the data is ARIMA(1,1,0)

Test for Seasonarity
H_0:There is no seasonarity component
H_1:There is seasonal component

```
welch(maize, freq = 12)

## Test used:  Kruskall Wallis
##
## Test statistic:  2
## P-value:  0.09404391
```

Conclusion: Since the p-value(=0.09404391),is greater than the significance level we fail to reject H_0 and conclude that there is no seasonal component in the maize series data.

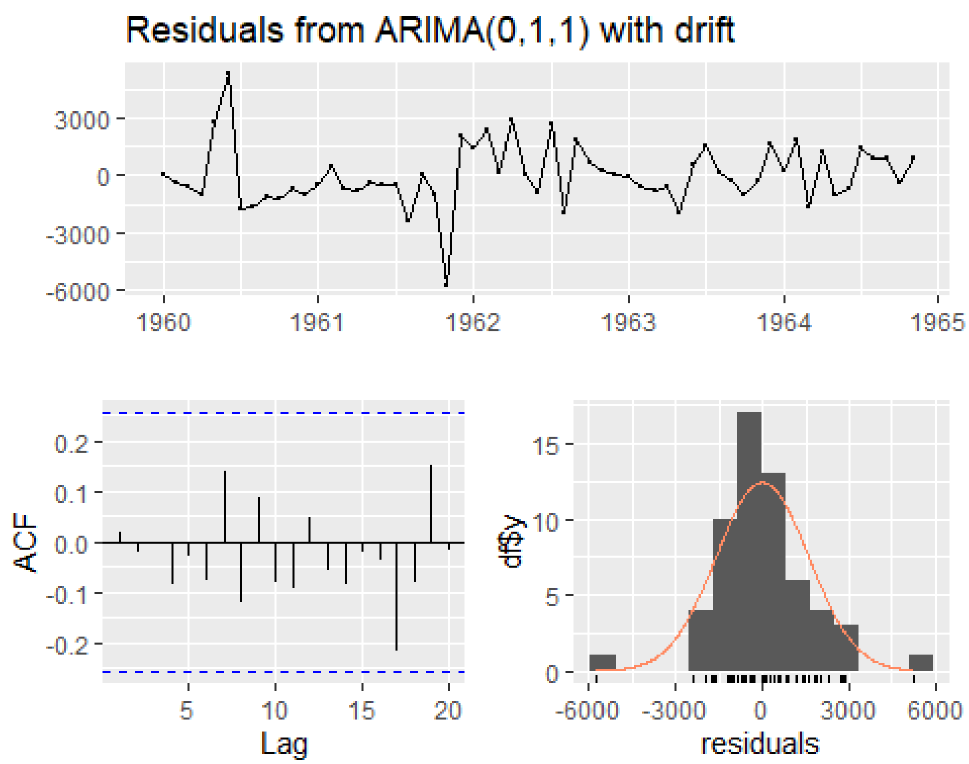
model

```
model1 = Arima(maize, order = c(1,1,0))
model2 = auto.arima(maize, stepwise = FALSE,approximation = FALSE)
model1

## Series: maize
## ARIMA(1,1,0)
```

```
##
## Coefficients:
##           ar1
##       -0.3747
## s.e.   0.1207
##
## sigma^2 = 2866400:  log likelihood = -513.06
## AIC=1030.12   AICc=1030.33   BIC=1034.24
```

```r
checkresiduals(model2)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(0,1,1) with drift
## Q* = 5.123, df = 11, p-value = 0.9251
##
## Model df: 1.   Total lags used: 12
```

Splitting data into train and test data

```r
traindata = window(maize,start = c(1960,1),end = c(1964,1))
testData = window(maize,start = c(1964,2))
traindata
```

```
##         Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct   Nov
Dec
## 1960  9536  9283  9039  8537 12062 16026 11640 11089 11047 10643 10758
```

```
10358
## 1961 10525 11431 10741 10438 10703 10634 10615  8682 10161  9276  4300
9613
## 1962 10086 11845 10907 13907 12610 11889 15260 12040 15092 14933 15020
15153
## 1963 15285 14850 14557 14578 13150 14900 16272 15794 15613 14994 15437
17371
## 1964 16969

testData

##        Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct   Nov
## 1964 18874 16458 18712 17240 17291 19218 19500 20112 19474 20697

welch(traindata)

## Test used:  Kruskall Wallis
##
## Test statistic:  1.5
## P-value:  0.2339556

m1 = auto.arima(traindata,stepwise = FALSE,approximation = FALSE)
m1

## Series: traindata
## ARIMA(0,1,1)
##
## Coefficients:
##           ma1
##       -0.4713
## s.e.   0.1384
##
## sigma^2 = 2979222:  log likelihood = -425.5
## AIC=855   AICc=855.27   BIC=858.75

m2 = Arima(traindata,order = c(1,1,0))
m2

## Series: traindata
## ARIMA(1,1,0)
##
## Coefficients:
##           ar1
##       -0.3361
## s.e.   0.1340
##
## sigma^2 = 3156199:  log likelihood = -426.82
## AIC=857.64   AICc=857.91   BIC=861.38

f1 = forecast(m1,h = length(testData))
f2 = forecast(m2, h = length(testData))
accuracy(f1,testData)
```

```
##                      ME      RMSE       MAE         MPE      MAPE       MASE
## Training set  278.8484 1690.450 1125.740 -0.1651948 10.61235 0.4401930
## Test set     2044.6850 2417.379 2095.668 10.4641743 10.77395 0.8194595
##                       ACF1 Theil's U
## Training set -0.003953328        NA
## Test set      0.310518088  1.582129
```

```
accuracy(f2,testData)
```

```
##                      ME      RMSE       MAE        MPE       MAPE       MASE
## Training set  205.9387 1739.935 1147.436 -0.613419 10.889739 0.4486766
## Test set     1684.9202 2120.330 1805.062  8.537876  9.267864 0.7058250
##                     ACF1 Theil's U
## Training set -0.1002318        NA
## Test set      0.3168790  1.390824
```

```
dat = arima.sim(model = list(oder = c(2,0,2),ar =c(0.25,0.63),ma =
c(0.53,0.43)),n = 60)
auto.arima(dat)
```

```
## Series: dat
## ARIMA(2,0,2) with zero mean
##
## Coefficients:
##          ar1     ar2     ma1     ma2
##       0.1427  0.6011  0.7909  0.5448
## s.e.  0.1342  0.1283  0.1531  0.1166
##
## sigma^2 = 1.066:  log likelihood = -86.38
## AIC=182.76   AICc=183.88   BIC=193.24
```