Stanisław Godlewski

11/8/2022

**Seasonality Analysis on the CPI index in the country of Canada from 2000 to 2009 per month.**

This seasonality analysis regarding the CPI index in the country of Canada from 2000 to 2009 was a project assigned in Elementary Economic Forecasting, an undergraduate course instructed at the University of Connecticut. The objective of this project was to understand and display the seasonal patterns of the data being analyzed. There are 120 documented observations in this analysis; with each one representing the months from January 2000 to December 2009. The code for this project was written in R.

Here is the code below:

```
library(readr)
library(seastests)

#Time Series of Consumer Price Index
CPI_1 <- read_csv("CPI.csv")
View(CPI_1)

CPI<-ts(data=CPI_1, start=c(2000,1), frequency=12)


print(CPI)
```

| | CPI |
|---|---|
| 1 | 211.933 |
| 2 | 212.705 |

| | |
|---|---|
| 3 | 212.495 |
| 4 | 212.709 |
| 5 | 213.022 |
| 6 | 214.790 |
| 7 | 214.726 |
| 8 | 215.445 |
| 9 | 215.861 |
| 10 | 216.509 |
| 11 | 217.234 |
| 12 | 217.347 |
| 13 | 217.488 |
| 14 | 217.281 |
| 15 | 217.353 |

| | |
|---|---|
| 16 | 217.403 |
| 17 | 217.290 |
| 18 | 217.199 |
| 19 | 217.605 |
| 20 | 217.923 |
| 21 | 218.275 |
| 22 | 219.035 |
| 23 | 219.590 |
| 24 | 220.472 |
| 25 | 221.187 |
| 26 | 221.898 |
| 27 | 223.046 |
| 28 | 224.093 |

| | |
|---|---|
| 29 | 224.806 |
| 30 | 224.806 |
| 31 | 225.395 |
| 32 | 226.106 |
| 33 | 226.597 |
| 34 | 226.750 |
| 35 | 227.169 |
| 36 | 227.223 |
| 37 | 227.842 |
| 38 | 228.329 |
| 39 | 228.807 |
| 40 | 229.187 |
| 41 | 228.713 |

| | |
|---|---|
| 42 | 228.524 |
| 43 | 228.590 |
| 44 | 229.918 |
| 45 | 231.015 |
| 46 | 231.638 |
| 47 | 231.249 |
| 48 | 231.221 |
| 49 | 231.679 |
| 50 | 232.937 |
| 51 | 232.282 |
| 52 | 231.797 |
| 53 | 231.893 |
| 54 | 232.445 |

| | |
|---|---|
| 55 | 232.900 |
| 56 | 233.456 |
| 57 | 233.544 |
| 58 | 233.669 |
| 59 | 234.100 |
| 60 | 234.719 |
| 61 | 235.288 |
| 62 | 235.547 |
| 63 | 236.028 |
| 64 | 236.468 |
| 65 | 236.918 |
| 66 | 237.231 |
| 67 | 237.498 |

| | |
|---|---|
| 68 | 237.460 |
| 69 | 237.477 |
| 70 | 237.430 |
| 71 | 236.983 |
| 72 | 236.252 |
| 73 | 234.718 |
| 74 | 235.236 |
| 75 | 236.005 |
| 76 | 236.156 |
| 77 | 236.974 |
| 78 | 237.684 |
| 79 | 238.053 |
| 80 | 238.028 |

| | |
|---|---|
| 81 | 237.506 |
| 82 | 237.781 |
| 83 | 238.016 |
| 84 | 237.817 |
| 85 | 237.833 |
| 86 | 237.469 |
| 87 | 238.038 |
| 88 | 238.827 |
| 89 | 239.464 |
| 90 | 240.167 |
| 91 | 240.150 |
| 92 | 240.602 |
| 93 | 241.051 |

| | |
|---|---|
| 94 | 241.691 |
| 95 | 242.029 |
| 96 | 242.772 |
| 97 | 243.780 |
| 98 | 243.961 |
| 99 | 243.749 |
| 100 | 244.051 |
| 101 | 243.962 |
| 102 | 244.182 |
| 103 | 244.390 |
| 104 | 245.297 |
| 105 | 246.418 |
| 106 | 246.587 |

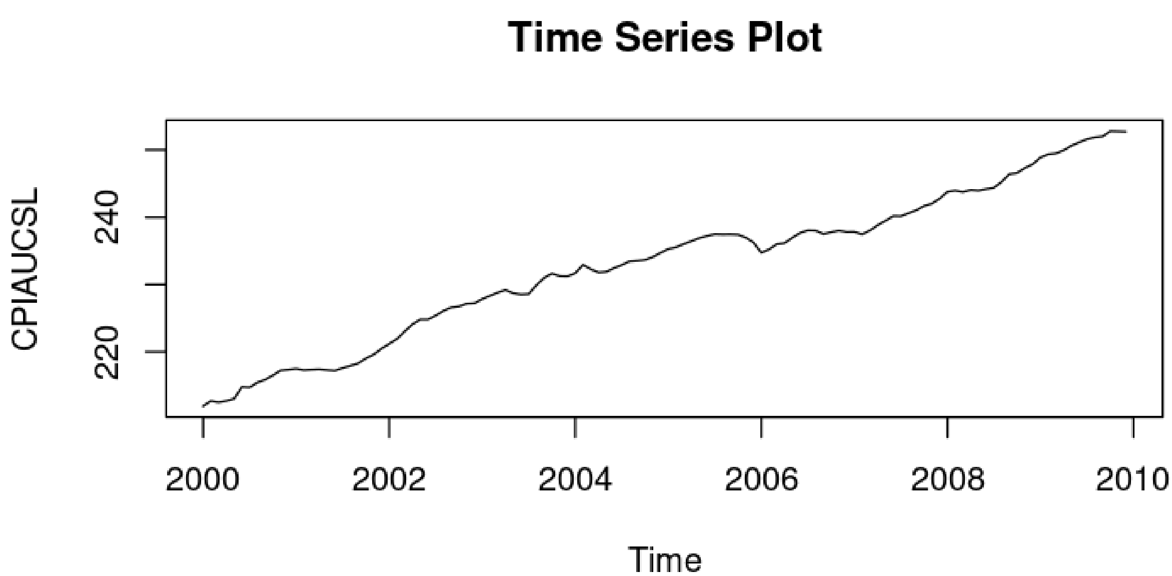
| | |
|---|---|
| 107 | 247.332 |
| 108 | 247.901 |
| 109 | 248.884 |
| 110 | 249.369 |
| 111 | 249.498 |
| 112 | 249.956 |
| 113 | 250.646 |
| 114 | 251.134 |
| 115 | 251.597 |
| 116 | 251.879 |
| 117 | 252.010 |
| 118 | 252.794 |
| 119 | 252.760 |

| | |
|---|---|
| 120 | 252.723 |

View(CPI)

#Time series plot
plot(CPI,main ="Time Series Plot")

View(ts)




time <- seq(1, 120, by=1)
View(time)

#Linear Model

CPIAUCSL<-data.frame(time, CPI)
linear.mod <- lm(CPI~time, data=CPIAUCSL)
print(summary(linear.mod))

```
Call:
lm(formula = CPI ~ time, data = CPIAUCSL)
```

```
Residuals:
    Min      1Q  Median      3Q     Max 
-3.5292 -1.7505 -0.2766  1.6960  3.2218 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 2.140e+02  3.419e-01  626.12   <2e-16 ***
time        3.134e-01  4.904e-03   63.91   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.861 on 118 degrees of freedom
Multiple R-squared:  0.9719, Adjusted R-squared:  0.9717 
F-statistic:  4085 on 1 and 118 DF,  p-value: < 2.2e-16
```

Y = 2.140e+02 + 3.134e-01(time)

#Quadratic Model

time2 <- time^2
quad.mod <- lm(formula = CPI ~ time2, data = CPIAUSCL)
print(summary(quad.mod))

```
Call:
lm(formula = CPI ~ time2, data = CPIAUSCL)

Residuals:
   Min     1Q Median     3Q    Max 
-9.436 -1.932 -0.523  3.514  5.583 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 2.214e+02  5.163e-01  428.77   <2e-16 ***
time2       2.395e-03  7.934e-05   30.18   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.76 on 118 degrees of freedom
Multiple R-squared:  0.8853, Adjusted R-squared:  0.8844 
F-statistic: 911.1 on 1 and 118 DF,  p-value: < 2.2e-16
```

Y = 2.214e+02 + 2.395e-03(time)

#Exponential Model

time.e <- log(time)
exp.model <- lm(formula = CPI ~ time.e, data = CPIAUSCL)
print(summary(exp.model))

```
Call:
lm(formula = CPI ~ time.e, data = CPIAUSCL)

Residuals:
   Min     1Q Median     3Q    Max 
-6.414 -2.768 -1.314  2.002 20.106 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 191.8274     1.7363  110.48   <2e-16 ***
time.e       10.7936     0.4421   24.41   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.515 on 118 degrees of freedom
Multiple R-squared:  0.8347, Adjusted R-squared:  0.8333 
F-statistic:   596 on 1 and 118 DF,  p-value: < 2.2e-16
```

Y = 191.8274 + 10.7936 (time.e)

#Linear-log-quadratic trend estimation with seasonal dummies
mond <- seasonaldummy(CPI)
logls.seas <- data.frame(cbind(CPI, time, mond))
linear.seas.mod <- lm(CPI~ mond + time, data = logls.seas)
print(summary(linear.seas.mod))

```
Call:
lm(formula = CPI ~ mond + time, data = logls.seas)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.3933 -1.7956 -0.4292  1.6890  3.3416 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) 214.188575   0.702953 304.698   <2e-16 ***
mondJan      -0.338813   0.871893  -0.389    0.698    
mondFeb      -0.241784   0.871573  -0.277    0.782    
mondMar      -0.297856   0.871283  -0.342    0.733    
```

```
mondApr      -0.276227   0.871024  -0.317    0.752
mondMay      -0.285099   0.870796  -0.327    0.744
mondJun      -0.150670   0.870598  -0.173    0.863
mondJul      -0.189442   0.870430  -0.218    0.828
mondAug       0.018586   0.870292   0.021    0.983
mondSep       0.069615   0.870186   0.080    0.936
mondOct       0.169643   0.870109   0.195    0.846
mondNov       0.114472   0.870064   0.132    0.896
time          0.312972   0.005153  60.740   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.945 on 107 degrees of freedom
Multiple R-squared:  0.9722, Adjusted R-squared:  0.969
F-statistic: 311.5 on 12 and 107 DF,  p-value: < 2.2e-16
```

Y = 214.188575-0.338813(mondJan) - .241784(mondFeb) -.297856(mondMar) -.276227(mondApr)-.285099(mondMay) - .150670(mondJun) -.189442(mondJul) + .018586(mondAug) + .069615(mondSep) + .169643(mondOct) + .114472(mondNov) + .312972(time)

```
mond <- seasonaldummy(CPI)
time2 <- time^2
logqls.seas <- data.frame(cbind(CPI,time,time2, mond))
quad.seas.mod <- lm(formula = CPI ~ time + time2 + mond, data = logqls.seas)
print(summary(quad.seas.mod))
```

```
Call:
lm(formula = CPI ~ time + time2 + mond, data = logqls.seas)

Residuals:
    Min      1Q  Median      3Q     Max
-3.7197 -1.4446  0.0825  1.5978  2.7369

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.128e+02  7.657e-01 277.880  < 2e-16 ***
time         3.837e-01  1.953e-02  19.653  < 2e-16 ***
time2       -5.849e-04  1.563e-04  -3.743 0.000296 ***
```

```
mondJan     -3.388e-01  8.233e-01  -0.412 0.681511
mondFeb     -2.476e-01  8.230e-01  -0.301 0.764084
mondMar     -3.084e-01  8.227e-01  -0.375 0.708533
mondApr     -2.903e-01  8.225e-01  -0.353 0.724854
mondMay     -3.015e-01  8.223e-01  -0.367 0.714619
mondJun     -1.682e-01  8.221e-01  -0.205 0.838259
mondJul     -2.070e-01  8.219e-01  -0.252 0.801657
mondAug      2.210e-03  8.218e-01   0.003 0.997859
mondSep      5.558e-02  8.217e-01   0.068 0.946201
mondOct      1.591e-01  8.216e-01   0.194 0.846811
mondNov      1.086e-01  8.216e-01   0.132 0.895065
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.837 on 106 degrees of freedom
Multiple R-squared:  0.9754, Adjusted R-squared:  0.9724
F-statistic: 323.6 on 13 and 106 DF,  p-value: < 2.2e-16
```

Y = 2.128e+02 -3.388e-01(mondJan) -2.476e-01(mondFeb) -3.084e-01(mondMar) -2.903e-01(mondApr)-3.015e-01(mondMay)-1.682e-01(mondJun) -2.070e-01(mondJul) + 2.210e-03(mondAug) + 5.558e-02(mondSep) + 1.591e-01(mondOct) + 1.086e-01(mondNov) + 3.837e-01(time) -5.849e-04(time2)

```
time.e <- log(time)
exp.model <- lm(formula = CPI ~ time.e, data = CPIAUSCL)
print(summary(exp.model))
logels.seas <- data.frame(cbind(CPI,time,time.e, mond))
exp.seas.mod <- lm(formula = CPI ~ time + time.e + mond, data = logels.seas)
print(summary(exp.seas.mod))
```

```
Call:
lm(formula = CPI ~ time + time.e + mond, data = logels.seas)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.5997 -1.3549  0.1793  1.4193  2.5213 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  2.095e+02  1.129e+00 185.516  < 2e-16 ***
time         2.668e-01  1.033e-02  25.828  < 2e-16 ***
time.e       1.934e+00  3.861e-01   5.009  2.2e-06 ***
mondJan      3.429e-02  7.912e-01   0.043    0.966    
mondFeb      7.143e-04  7.889e-01   0.001    0.999    
mondMar     -1.287e-01  7.879e-01  -0.163    0.871    
```

```
mondApr     -1.562e-01  7.873e-01  -0.198    0.843
mondMay     -2.003e-01  7.869e-01  -0.255    0.800
mondJun     -9.207e-02  7.866e-01  -0.117    0.907
mondJul     -1.504e-01  7.864e-01  -0.191    0.849
mondAug      4.302e-02  7.863e-01   0.055    0.956
mondSep      8.348e-02  7.862e-01   0.106    0.916
mondOct      1.762e-01  7.861e-01   0.224    0.823
mondNov      1.166e-01  7.861e-01   0.148    0.882
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.758 on 106 degrees of freedom
Multiple R-squared:  0.9775, Adjusted R-squared:  0.9747
F-statistic: 354.2 on 13 and 106 DF,  p-value: < 2.2e-16
```

Y = 2.095e+02 + 3.429e-02(mondJan)+ 7.143e-04(mondFeb) -1.287e-01(mondMar) -1.562e-01(mondApr)-2.003e-01(mondMay)-9.207e-02(mondJun) -1.504e-01(mondJul) + 4.302e-02(mondAug) + 8.348e-02 (mondSep) + 1.762e-01(mondOct) + 1.166e-01(mondNov) + 2.668e-01(time)+ 1.934e+00(time.e)

While the Exponential Model has the lowest residual median out of all three models in terms of residuals, it has the highest range of residuals out of all three models. This means that this model does not accurately predict all the CPIS accurately for each observation as well as the Linear Model. The Linear Model has the lowest residual range out of all the models. This means that the predicted CPI for each observation in the Linear Model is the least off from the actual outcome. In terms of the plots, the Linear Model seems to be the most aligned with the actual results. Given that the Linear Model also yielded the highest R2, the predicted outcomes of the Linear Model and the actual results from the dataset have the strongest correlation with eachother out of all three models.

#TimeFrame

time_new <- seq(1, 120, by=1)
View(time_new)

CPIAUSCL<-data.frame(time_new, CPI)
print(CPIAUSCL)

| | CPI |
|---|---|
| 1 | 211.933 |
| 2 | 212.705 |
| 3 | 212.495 |
| 4 | 212.709 |
| 5 | 213.022 |
| 6 | 214.790 |
| 7 | 214.726 |
| 8 | 215.445 |
| 9 | 215.861 |
| 10 | 216.509 |
| 11 | 217.234 |
| 12 | 217.347 |

| | |
|---|---|
| 13 | 217.488 |
| 14 | 217.281 |
| 15 | 217.353 |
| 16 | 217.403 |
| 17 | 217.290 |
| 18 | 217.199 |
| 19 | 217.605 |
| 20 | 217.923 |
| 21 | 218.275 |
| 22 | 219.035 |
| 23 | 219.590 |
| 24 | 220.472 |
| 25 | 221.187 |

| | |
|---|---|
| 26 | 221.898 |
| 27 | 223.046 |
| 28 | 224.093 |
| 29 | 224.806 |
| 30 | 224.806 |
| 31 | 225.395 |
| 32 | 226.106 |
| 33 | 226.597 |
| 34 | 226.750 |
| 35 | 227.169 |
| 36 | 227.223 |
| 37 | 227.842 |
| 38 | 228.329 |

| | |
|---|---|
| 39 | 228.807 |
| 40 | 229.187 |
| 41 | 228.713 |
| 42 | 228.524 |
| 43 | 228.590 |
| 44 | 229.918 |
| 45 | 231.015 |
| 46 | 231.638 |
| 47 | 231.249 |
| 48 | 231.221 |
| 49 | 231.679 |
| 50 | 232.937 |
| 51 | 232.282 |

| | |
|---|---|
| 52 | 231.797 |
| 53 | 231.893 |
| 54 | 232.445 |
| 55 | 232.900 |
| 56 | 233.456 |
| 57 | 233.544 |
| 58 | 233.669 |
| 59 | 234.100 |
| 60 | 234.719 |
| 61 | 235.288 |
| 62 | 235.547 |
| 63 | 236.028 |
| 64 | 236.468 |

| | |
|---|---|
| 65 | 236.918 |
| 66 | 237.231 |
| 67 | 237.498 |
| 68 | 237.460 |
| 69 | 237.477 |
| 70 | 237.430 |
| 71 | 236.983 |
| 72 | 236.252 |
| 73 | 234.718 |
| 74 | 235.236 |
| 75 | 236.005 |
| 76 | 236.156 |
| 77 | 236.974 |

| | |
|---|---|
| 78 | 237.684 |
| 79 | 238.053 |
| 80 | 238.028 |
| 81 | 237.506 |
| 82 | 237.781 |
| 83 | 238.016 |
| 84 | 237.817 |
| 85 | 237.833 |
| 86 | 237.469 |
| 87 | 238.038 |
| 88 | 238.827 |
| 89 | 239.464 |
| 90 | 240.167 |

| | |
|---|---|
| 91 | 240.150 |
| 92 | 240.602 |
| 93 | 241.051 |
| 94 | 241.691 |
| 95 | 242.029 |
| 96 | 242.772 |
| 97 | 243.780 |
| 98 | 243.961 |
| 99 | 243.749 |
| 100 | 244.051 |
| 101 | 243.962 |
| 102 | 244.182 |
| 103 | 244.390 |

| | |
|---|---|
| 104 | 245.297 |
| 105 | 246.418 |
| 106 | 246.587 |
| 107 | 247.332 |
| 108 | 247.901 |
| 109 | 248.884 |
| 110 | 249.369 |
| 111 | 249.498 |
| 112 | 249.956 |
| 113 | 250.646 |
| 114 | 251.134 |
| 115 | 251.597 |
| 116 | 251.879 |

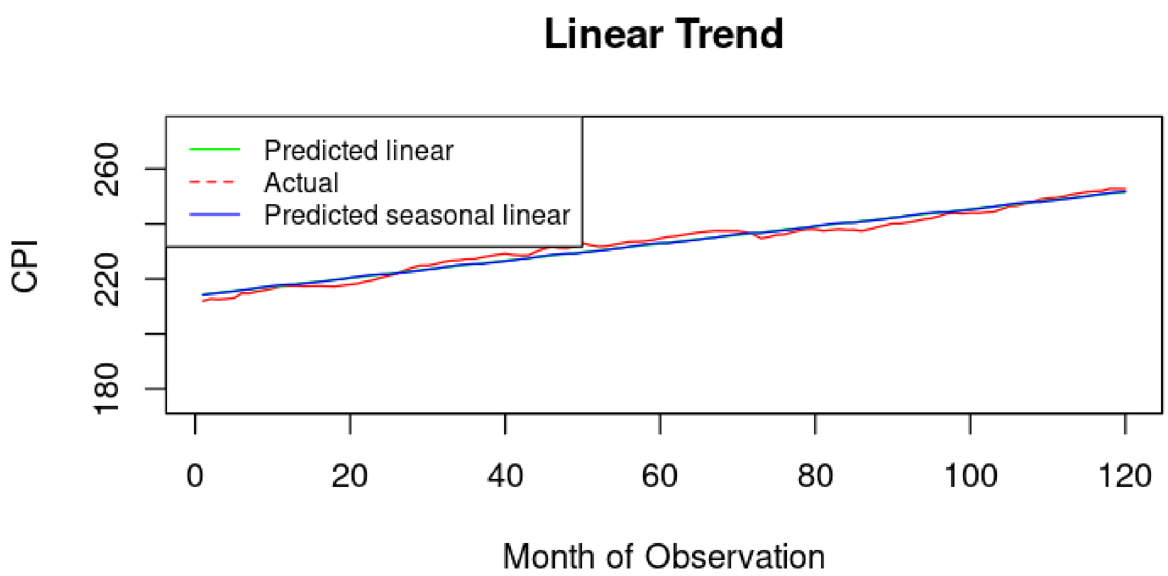
| | |
|---|---|
| 117 | 252.010 |
| 118 | 252.794 |
| 119 | 252.760 |
| 120 | 252.723 |

linear.mod <- lm(CPI~time_new, data=CPIAUSCL)
print(summary(linear.mod))

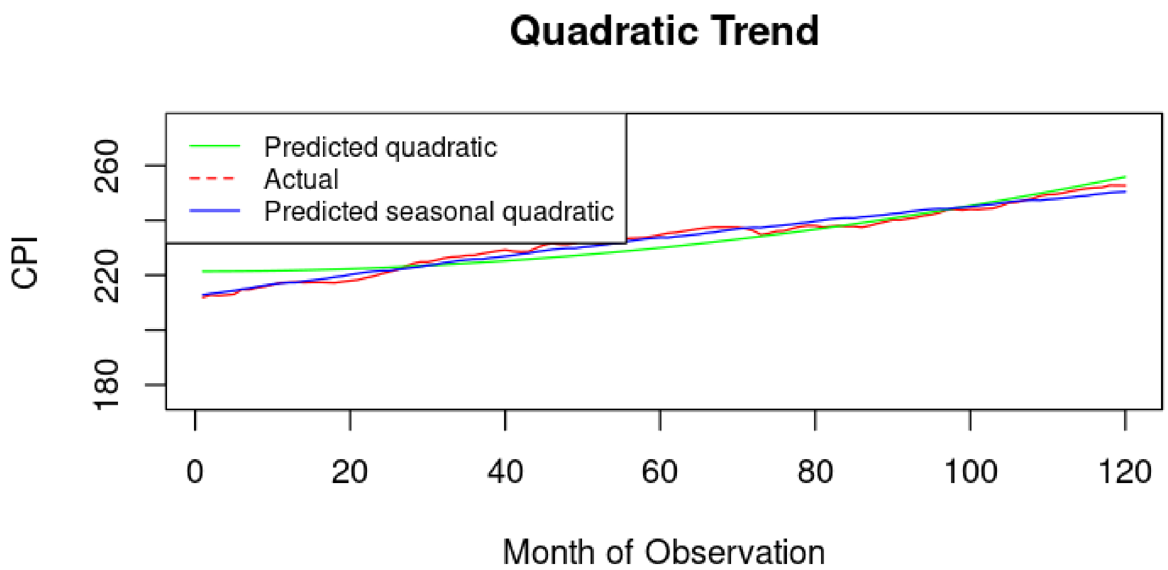
```
Call:
lm(formula = CPI ~ time_new, data = CPIAUSCL)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.5292 -1.7505 -0.2766  1.6960  3.2218 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 2.140e+02  3.419e-01  626.12   <2e-16 ***
time_new    3.134e-01  4.904e-03   63.91   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.861 on 118 degrees of freedom
Multiple R-squared:  0.9719, Adjusted R-squared:  0.9717 
F-statistic:  4085 on 1 and 118 DF,  p-value: < 2.2e-16
```

#Linear Graph
CPA_10.act <- CPIAUSCL$CPI
CPA_10.pred <- predict(linear.mod, CPIAUSCL)
CPA_10.resid <- CPA_10.act - CPA_10.pred
CPA_10.resid.plot <- CPA_10.resid + 200
CPA_10.pred.seas <- predict(linear.seas.mod, CPIAUSCL)
CPA_10.resid.seas <- CPA_10.act - CPA_10.pred.seas
CPA_10.pred.seas.plot <- CPA_10.resid.seas + 200

```
plot(CPA_10.act, type="l", axes=TRUE, main ="Linear Trend", ylim=c(175,275), xlab="Month of Observation", ylab="CPI", col="red")
lines(CPA_10.pred, col="green")
lines(CPA_10.pred.seas, col="blue")
legend("topleft", legend=c("Predicted linear", "Actual", 'Predicted seasonal linear'),
      col=c("green",'red' ,"blue"), lty=1:2, cex=0.8)
```

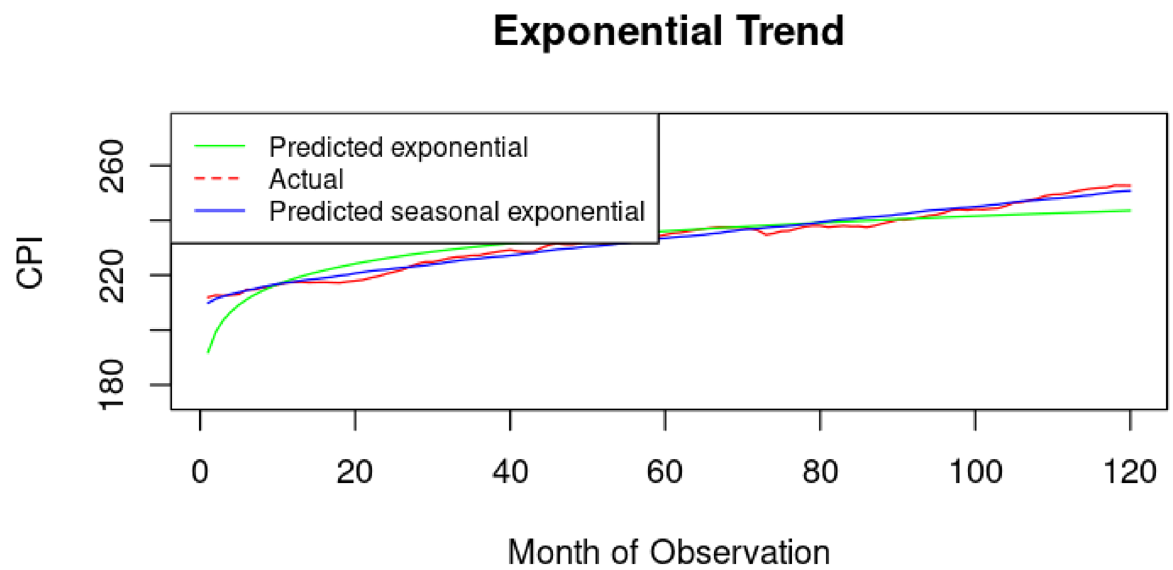



```
#Quadratic Graph

time2 <- time^2
quad.mod <- lm(formula = CPI ~ time2, data = CPIAUSCL)
print(summary(quad.mod))
CPIAUSCL.act <-  CPIAUSCL$CPI
CPIAUSCL.pred <- predict(quad.mod,CPIAUSCL)
CPIAUSCL.resid <-  CPIAUSCL.act- CPIAUSCL.pred
CPIAUSCL.resid.plot <-  CPIAUSCL.resid + 200
CPIAUSCL.pred.seas <- predict(quad.seas.mod, CPIAUSCL)
CPIAUSCL.resid.seas <- CPIAUSCL.act - CPIAUSCL.pred.seas
CPIAUSCL.pred.seas.plot <- CPIAUSCL.resid.seas + 200
plot( CPIAUSCL.act, type="l", axes=TRUE, main ="Quadratic Trend", ylim=c(175,275),
xlab="Month of Observation", ylab="CPI", col="red")
lines( CPIAUSCL.pred, col="green")
lines(CPIAUSCL.pred.seas, col="blue")
legend("topleft", legend=c("Predicted quadratic", "Actual", 'Predicted seasonal quadratic'),
      col=c("green",'red' ,"blue"), lty=1:2, cex=0.8)
```




#Exponential Graph

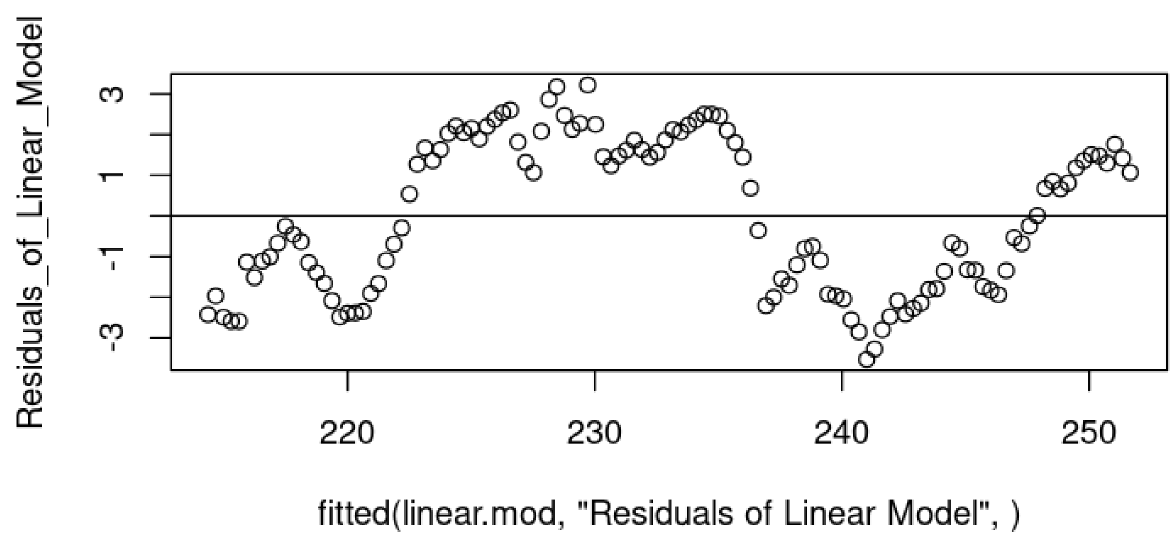
```
CPIAUSCL_1.act <-  CPIAUSCL$CPI
CPIAUSCL_1.pred <- predict(exp.model, CPIAUSCL)
CPIAUSCL_1.resid <- CPIAUSCL_1.act - CPIAUSCL_1.pred
CPIAUSCL_1.resid.plot <-  CPIAUSCL_1.resid + 200
CPIAUSCL_1.pred.seas <- predict(exp.seas.mod, CPIAUSCL)
CPIAUSCL_1.resid.seas <- CPIAUSCL_1.act - CPIAUSCL_1.pred.seas
CPIAUSCL_1.pred.seas.plot <- CPIAUSCL_1.resid.seas + 200
plot( CPIAUSCL_1.act, type="l", axes=TRUE, main ="Exponential Trend", ylim=c(175,275),
xlab="Month of Observation", ylab="CPI", col="red")
lines( CPIAUSCL_1.pred, col="green")
lines( CPIAUSCL_1.pred.seas, col="blue")
legend("topleft", legend=c("Predicted exponential", "Actual", 'Predicted seasonal exponential'),
    col=c("green",'red' ,"blue"), lty=1:2, cex=0.8)
```




#Linear Residuals

Residuals_of_Linear_Model <- resid(linear.mod)
plot(fitted(linear.mod, "Residuals of Linear Model",), Residuals_of_Linear_Model)

abline(0,0)

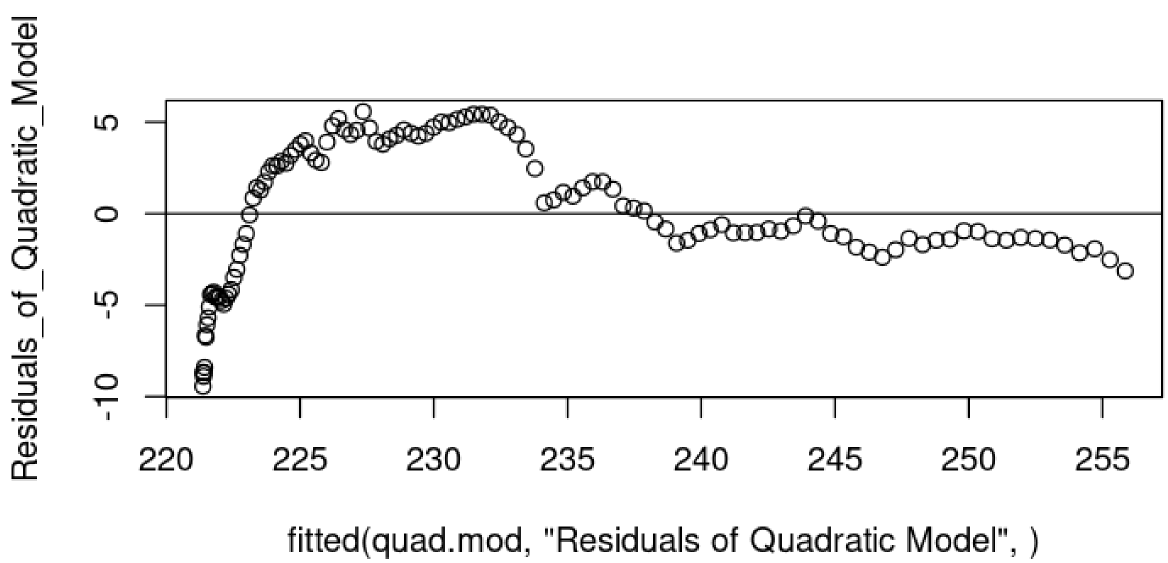


As one can observe this residual plot, one may notice some sharp spikes in the fitted model, which form a nonlinear pattern. However, the observations are spread out in a "C" shaped pattern; indicating heteroskedasticity and bias.The black dots represent the respective observations while the dotted horizontal line represents the regression model.

#Quadratic Residuals

Residuals_of_Quadratic_Model <- CPIAUSCL.resid
plot(fitted(quad.mod, "Residuals of Quadratic Model",), Residuals_of_Quadratic_Model)

abline(0,0)

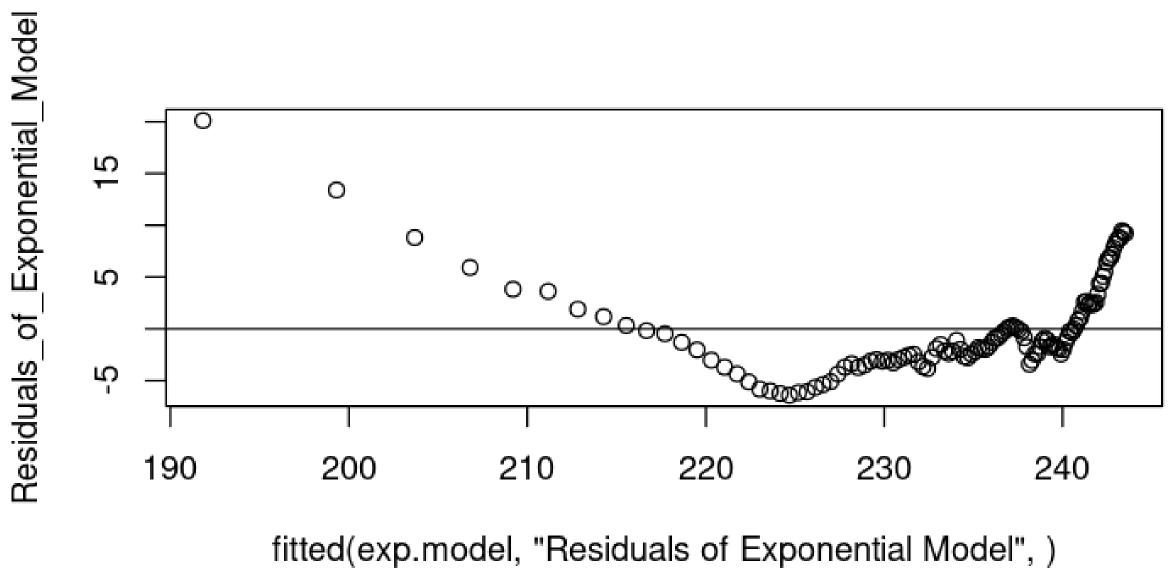



As one can observe this residual plot, one may notice some sharp spikes in the fitted model, which form a nonlinear pattern. However, the observations are spread out in a "C" shaped pattern; indicating heteroskedasticity and bias.The black dots represent the respective observations while the dotted horizontal line represents the regression model.

#Exponential Residuals

Residuals_of_Exponential_Model <- resid(exp.model)
plot(fitted(exp.model, "Residuals of Exponential Model",), Residuals_of_Exponential_Model)

abline(0,0)




As one can observe, there is a double yet subtle hump along the lowess line in the residual plot. In addition, their residuals are quite spread out in no particular pattern which indicates constant variance and homoscedasticity. The black dots are the respective residuals while the dotted horizontal line represents the regression model.

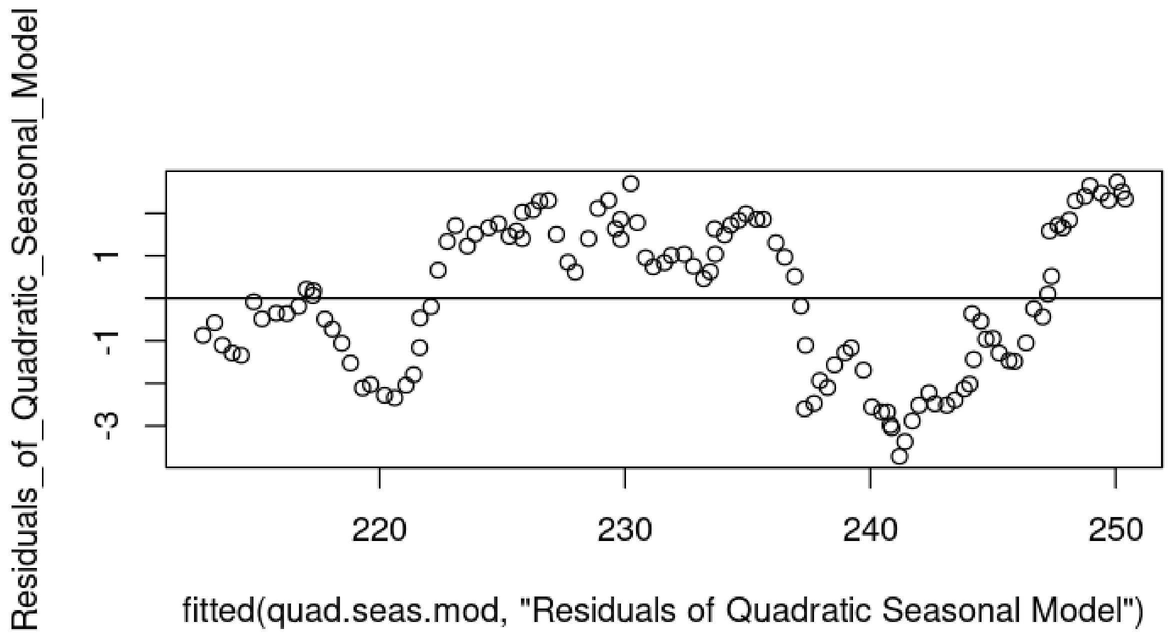
#Seasonal Quadratic Residuals

Residuals_of_Quadratic_Seasonal_Model <- resid(quad.seas.mod)

typeof(Residuals_of_Quadratic_Seasonal_Model)

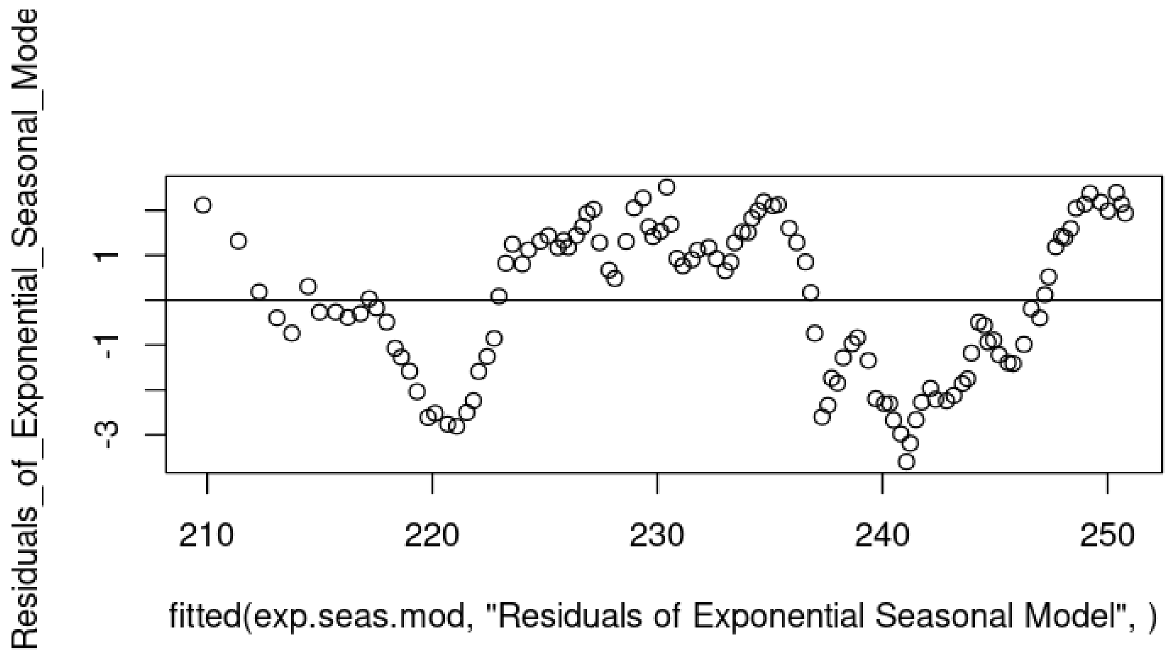
plot(fitted(quad.seas.mod, "Residuals of Quadratic Seasonal Model"), Residuals_of_Quadratic_Seasonal_Model)

abline(0,0)




As one can observe, there is a double yet subtle hump along the lowess line in the residual plot. In addition, their residuals are quite spread out in no particular pattern which indicates constant variance and homoscedasticity. The black dots are the respective residuals while the dotted horizontal line represents the regression model.

```
#Seasonal Exponential Residuals

Residuals_of_Exponential_Seasonal_Model <- resid(exp.seas.mod)

plot(fitted(exp.seas.mod, "Residuals of Exponential Seasonal Model",),
Residuals_of_Exponential_Seasonal_Model)
abline(0,0)
```




As one can observe, there is a double yet subtle hump along the lowess line in the residual plot. In addition, their residuals are quite spread out in no particular pattern which indicates constant variance and homoscedasticity. The black dots are the respective residuals while the dotted horizontal line represents the regression model.

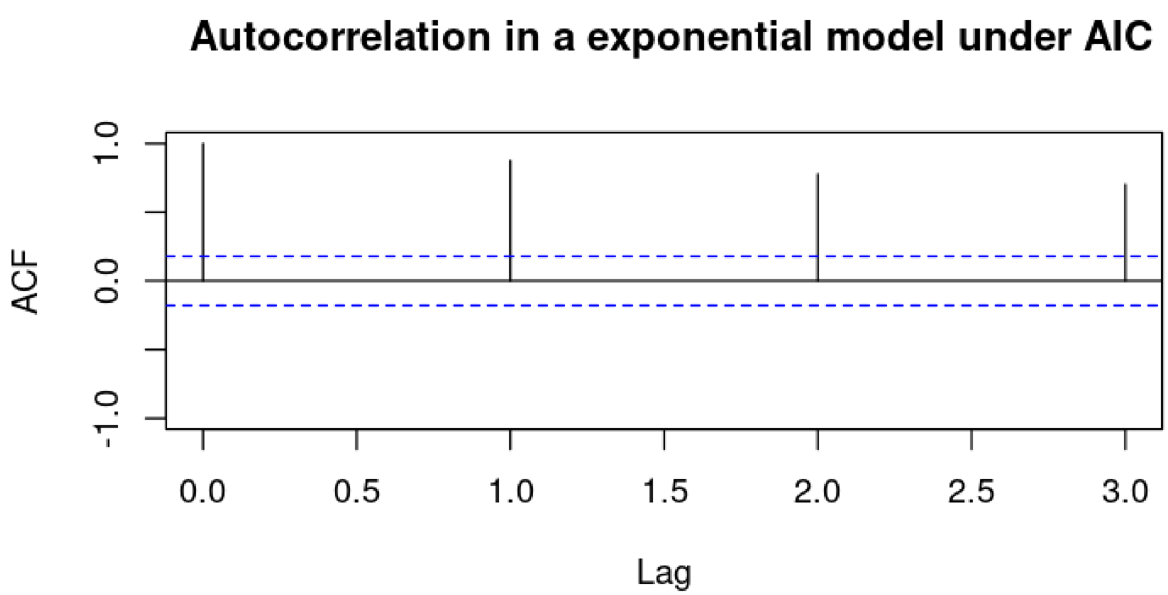
```
#Seasonal Linear Residuals

Residuals_of_Linear_Seasonal_Model <- resid(linear.seas.mod)
plot(fitted(linear.seas.mod, "Residuals of Linear Seasonal Model",),
Residuals_of_Linear_Seasonal_Model)
abline(0,0)

#AIC vs BIC
AIC <- AIC(linear.mod,quad.mod,exp.model,quad.seas.mod,linear.seas.mod,exp.seas.mod)
BIC <- BIC(linear.mod,quad.mod,exp.model,quad.seas.mod,linear.seas.mod,exp.seas.mod)
mod.selection <- data.frame(cbind(AIC,BIC[,2]))
names(mod.selection)[3] <- "BIC"
print(mod.selection)
```

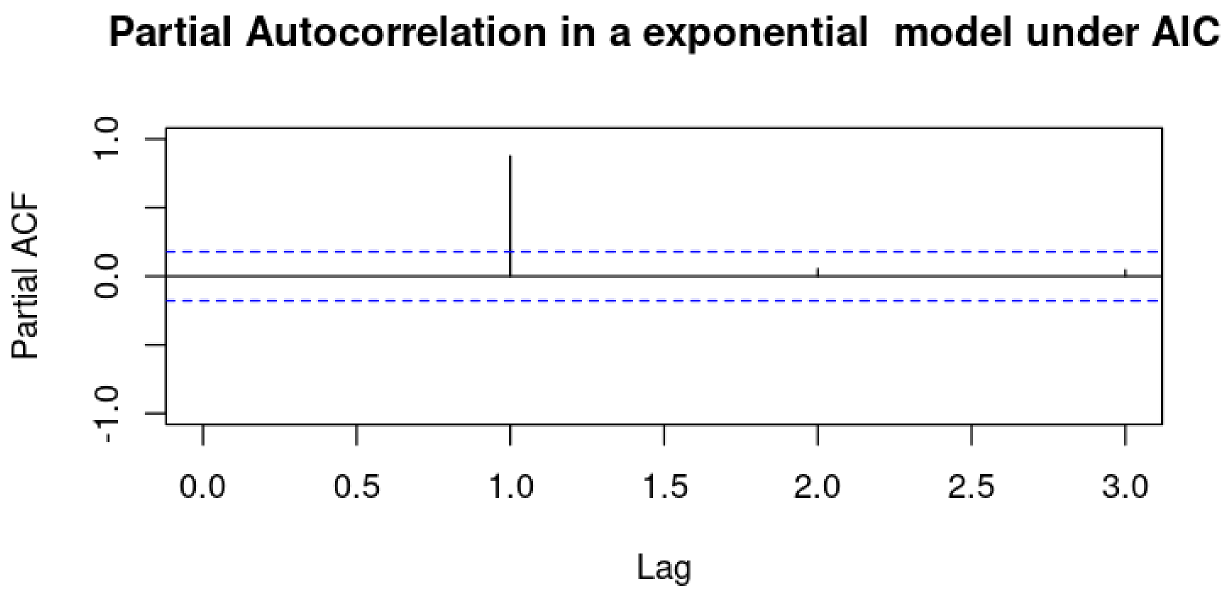
| | df | AIC | BIC |
|---|---|---|---|
| linear.mod | 3 | 493.5627 | 501.9251 |
| quad.mod | 3 | 662.4199 | 670.7823 |
| exp.model | 3 | 706.2842 | 714.6467 |
| quad.seas.mod | 15 | 501.6160 | 543.4284 |

```
linear.seas.mod 14 514.5087 553.5336
exp.seas.mod    15 491.0178 532.8301
```

#Exponential auto and partial correlation plots

```
plot(acf(resisual_exp, plot = F), xlim = c(0,3), ylim = c(-1, 1), main = " Autocorrelation in a exponential model under AIC")
plot(acf(resisual_exp, plot = F, type =  "p"), xlim = c(0,3),ylim = c(-1,1), main = "Partial Autocorrelation in a exponential  model under AIC")
```

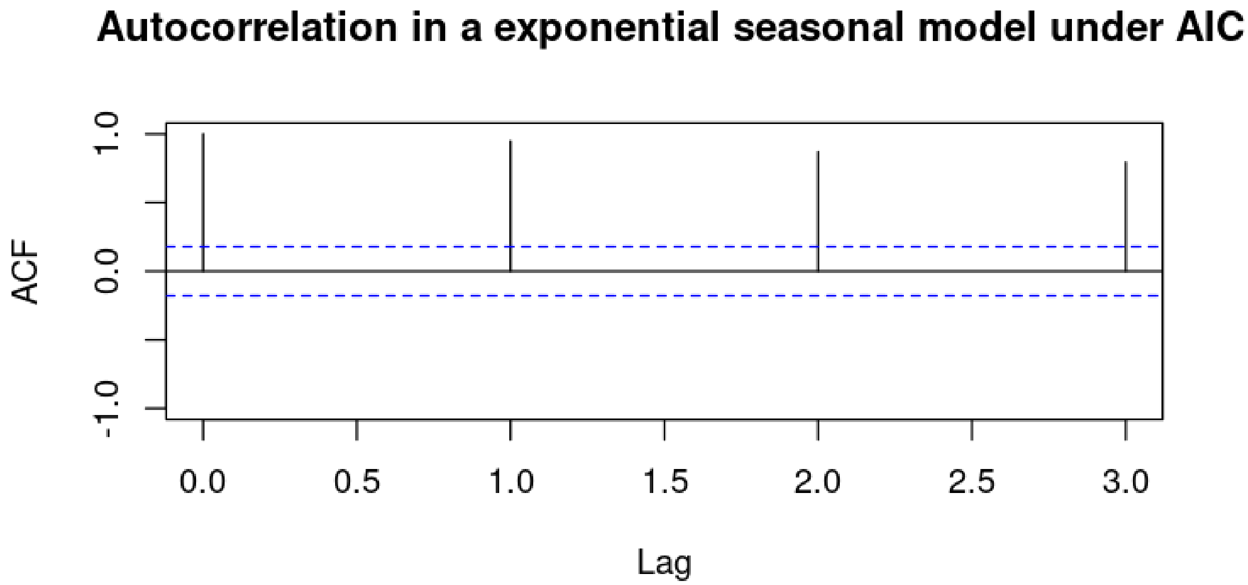


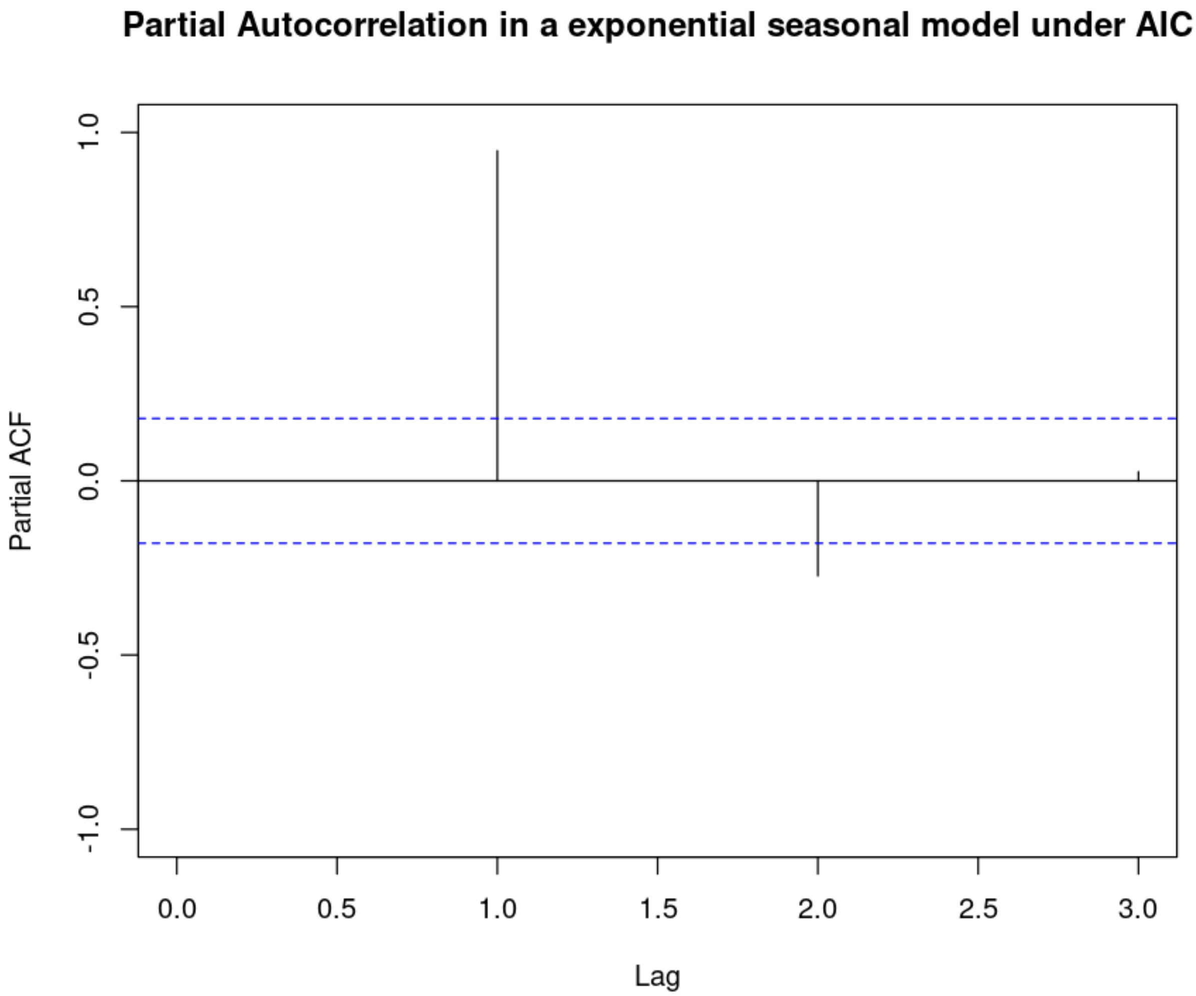



As one can observe, there is statistical significance in the partial and autocorrelation plots, indicating that the lags directly beforehand influenced the CPI values.

The plots below graphically summarize the strength of a relationship with an observation in a time series with observations at prior time steps.The autocorrelation with lag zero always equals 1, because this represents the autocorrelation between each term and itself. Each spike that rises above or falls below the shaded blue shapes is considered to be statistically significant. This means the spike has a value that is significantly different from zero. If a spike is significantly different from zero, that is evidence of autocorrelation. A spike that's close to zero is evidence against autocorrelation.Meanwhile, the lags within partial autocorrelation plots depend on the lag  directly beforehand. SInce the AIC was lower in value than the BIC, the AIC will be used for the correlation plots.

```
plot(acf(resisual_exp_seas, plot = F), xlim = c(0,3), ylim = c(-1, 1), main = " Autocorrelation in a exponential seasonal model under AIC")
plot(acf(resisual_exp_seas, plot = F, type =  "p"), xlim = c(0,3),ylim = c(-1,1), main = "Partial Autocorrelation in a exponential seasonal model under AIC")
```


Autocorrelation in a exponential seasonal model under AIC
ACF
1.0
0.0
-1.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Lag

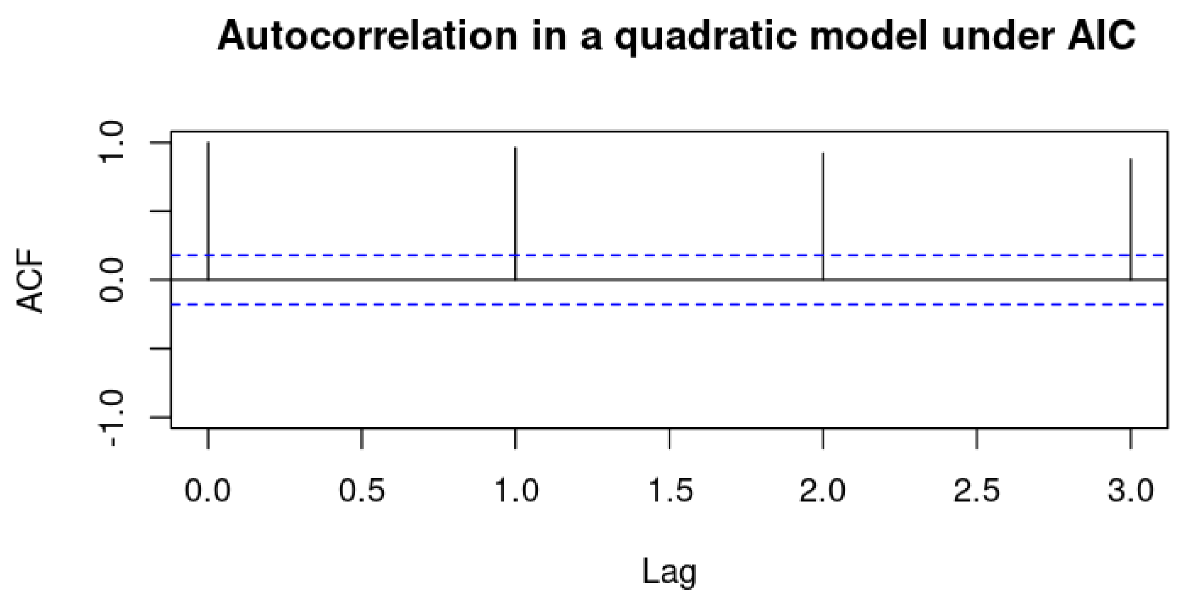


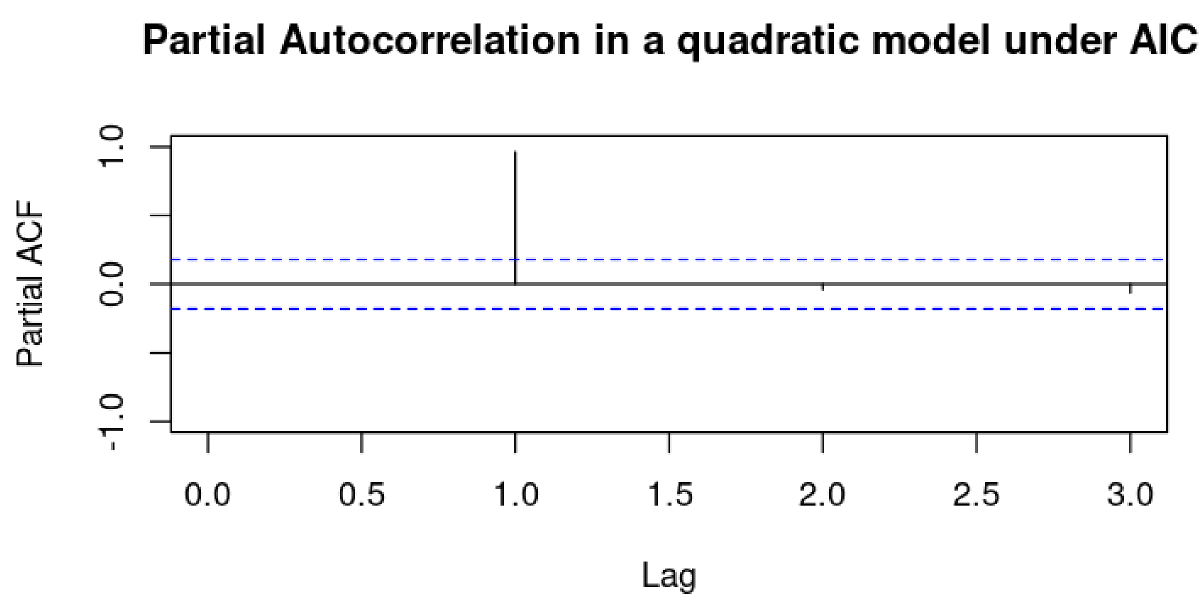

The plots above graphically summarize the strength of a relationship with an observation in a time series with observations at prior time steps.The autocorrelation with lag zero always equals 1, because this represents the autocorrelation between each term and itself. Each spike that rises above or falls below the shaded blue shapes is considered to be statistically significant. This means the spike has a value that is significantly different from zero. If a spike is significantly different from zero, that is evidence of autocorrelation. A spike that's close to zero is evidence against autocorrelation.Meanwhile, the lags within partial autocorrelation plots depend on the lag  directly beforehand. SInce the AIC was lower in value than the BIC, the AIC will be used for the correlation plots.

As one can observe, there is statistical significance in the partial and autocorrelation plots, indicating that the lags directly beforehand influenced the CPI values.

#Quadratic auto and partial correlation plots

plot(acf(resisual_quad, plot = F), xlim = c(0,3), ylim = c(-1, 1), main = " Autocorrelation in a quadratic model under AIC")
plot(acf(resisual_quad, plot = F, type =  "p"), xlim = c(0,3),ylim = c(-1,1), main = "Partial Autocorrelation in a quadratic model under AIC")

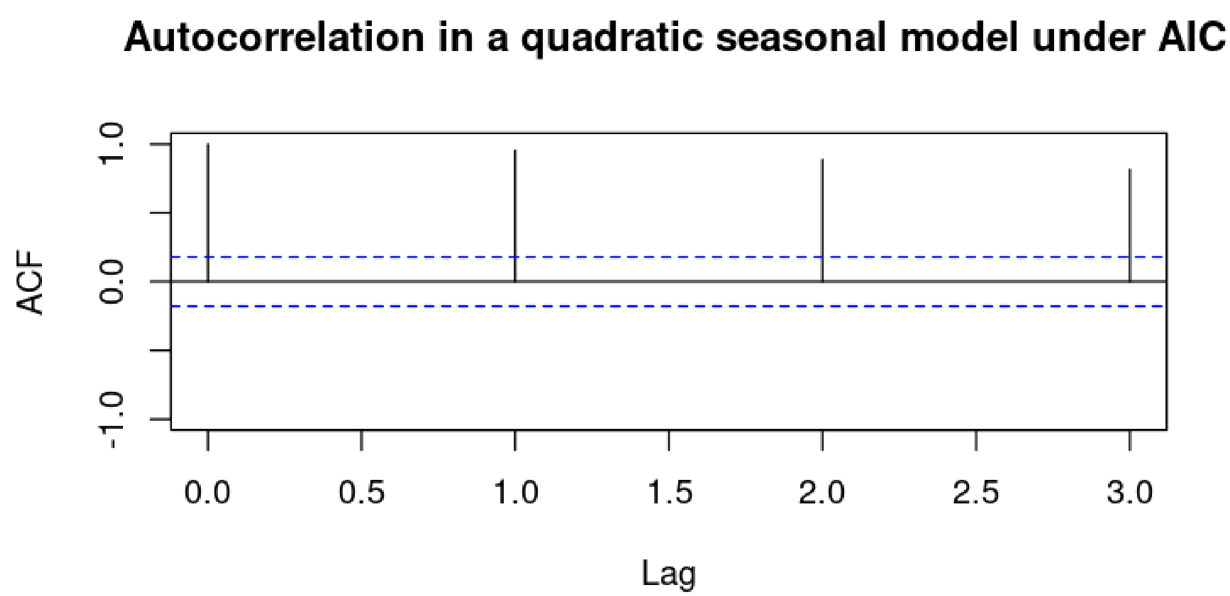


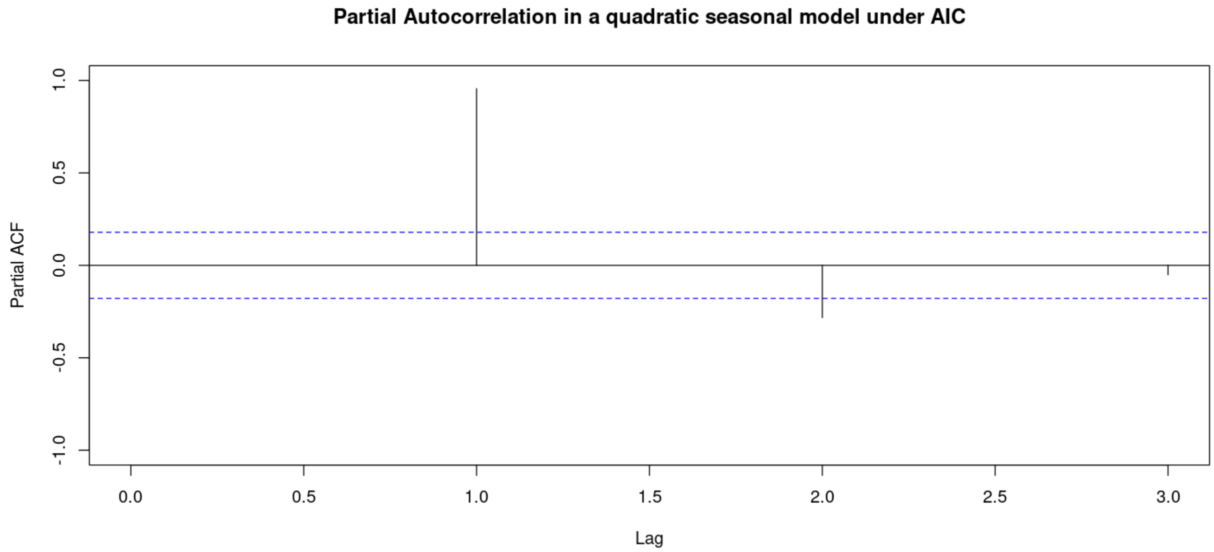



The plots above graphically summarize the strength of a relationship with an observation in a time series with observations at prior time steps.The autocorrelation with lag zero always equals 1, because this represents the autocorrelation between each term and itself. Each spike that rises above or falls below the shaded blue shapes is considered to be statistically significant. This means the spike has a value that is significantly different from zero. If a spike is significantly different from zero, that is evidence of autocorrelation. A spike that's close to zero is evidence against autocorrelation.Meanwhile, the lags within partial autocorrelation plots depend on the lag  directly beforehand. SInce the AIC was lower in value than the BIC, the AIC will be used for the correlation plots.

As one can observe, there is statistical significance in the partial and autocorrelation plots, indicating that the lags directly beforehand influenced the CPI values.

```
plot(acf(Residuals_of_Quadratic_Seasonal_Model, plot = F), xlim = c(0,3), ylim = c(-1, 1), main = " Autocorrelation in a quadratic seasonal model under AIC")
plot(acf(Residuals_of_Quadratic_Seasonal_Model, plot = F, type =  "p"), xlim = c(0,3),ylim = c(-1,1), main = "Partial Autocorrelation in a quadratic seasonal model under AIC")
```

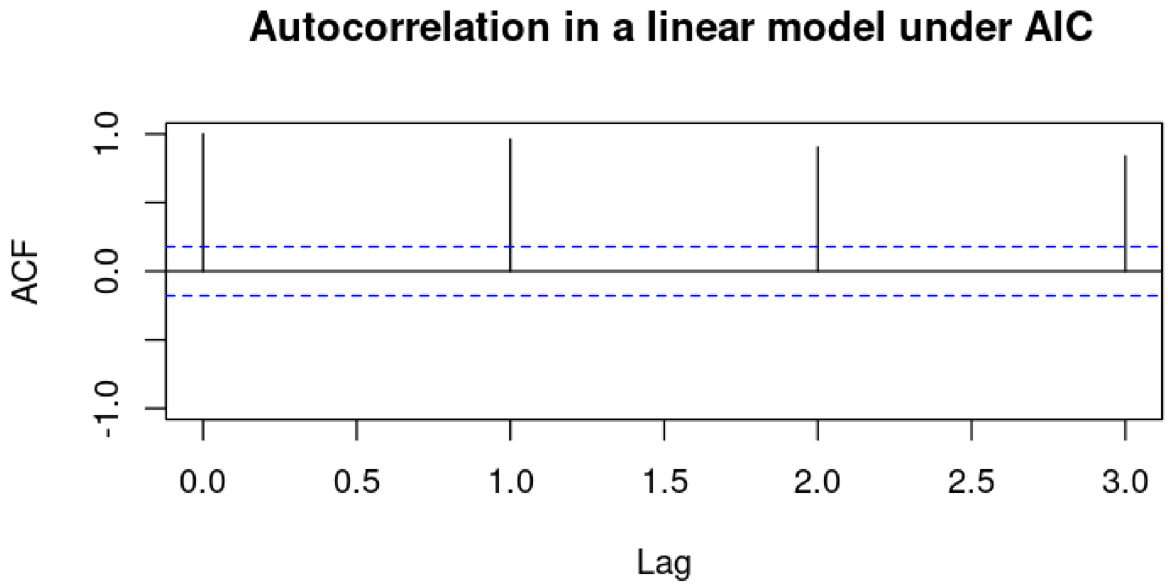


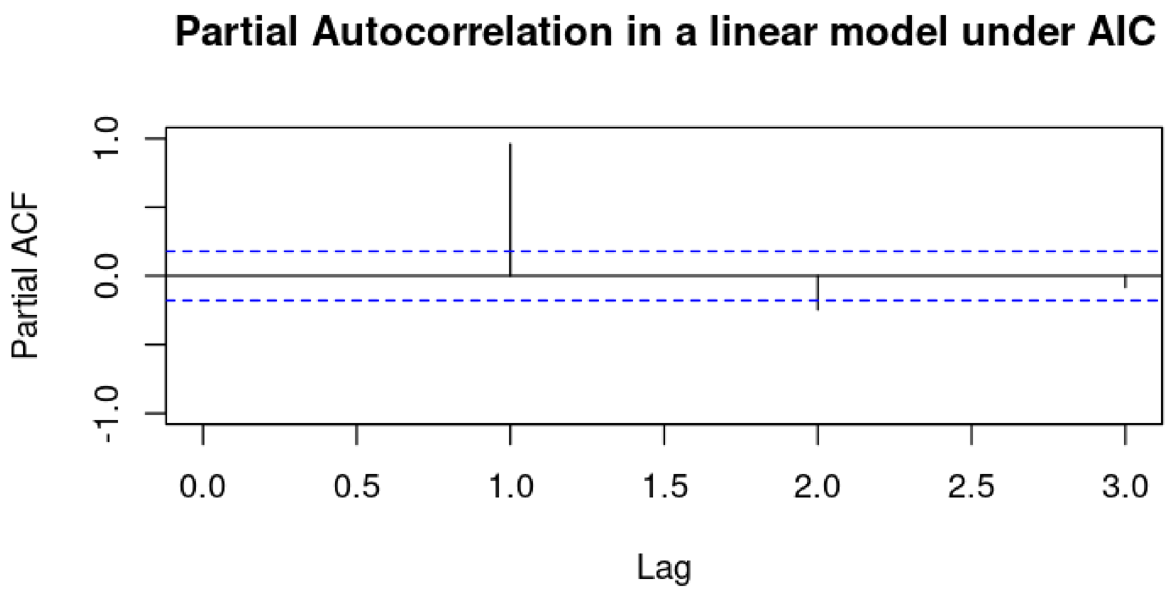



The plots above graphically summarize the strength of a relationship with an observation in a time series with observations at prior time steps.The autocorrelation with lag zero always equals 1, because this represents the autocorrelation between each term and itself. Each spike that rises above or falls below the shaded blue shapes is considered to be statistically significant. This means the spike has a value that is significantly different from zero. If a spike is significantly different from zero, that is evidence of autocorrelation. A spike that's close to zero is evidence against autocorrelation.Meanwhile, the lags within partial autocorrelation plots depend on the lag  directly beforehand. SInce the AIC was lower in value than the BIC, the AIC will be used for the correlation plots.

As one can observe, there is statistical significance in the partial and autocorrelation plots, indicating that the lags directly beforehand influenced the CPI values.

```
#Linear auto and partial correlation plots

plot(acf(resisual_linear, plot = F), xlim = c(0,3), ylim = c(-1, 1), main = " Autocorrelation in a linear model under AIC")
plot(acf(resisual_linear, plot = F, type =  "p"), xlim = c(0,3),ylim = c(-1,1), main = "Partial Autocorrelation in a linear model under AIC")
```

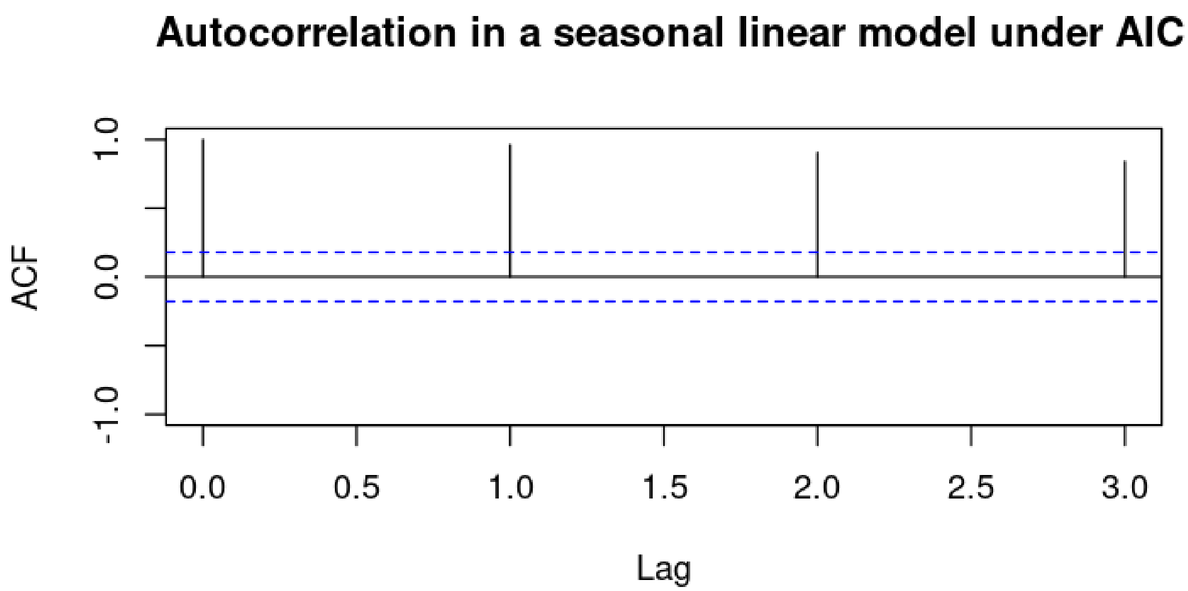


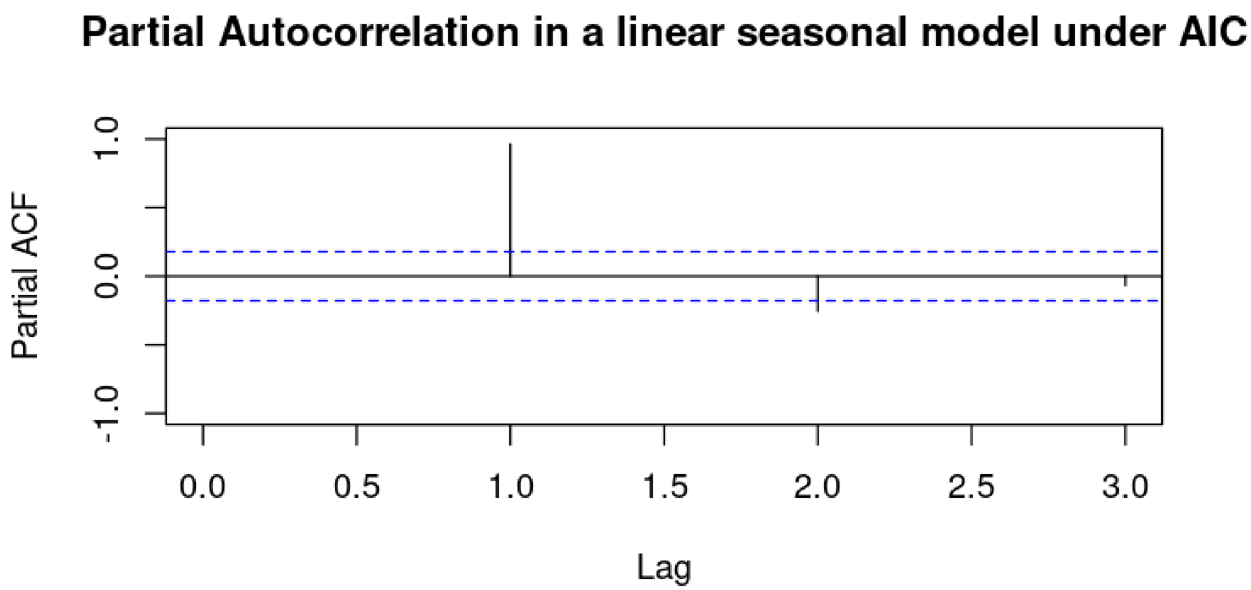




The plots above graphically summarize the strength of a relationship with an observation in a time series with observations at prior time steps.The autocorrelation with lag zero always equals 1, because this represents the autocorrelation between each term and itself. Each spike that rises above or falls below the shaded blue shapes is considered to be statistically significant. This means the spike has a value that is significantly different from zero. If a spike is significantly different from zero, that is evidence of autocorrelation. A spike that's close to zero is evidence against autocorrelation.Meanwhile, the lags within partial autocorrelation plots depend on the lag  directly beforehand. SInce the AIC was lower in value than the BIC, the AIC will be used for the correlation plots.

As one can observe, there is statistical significance in the partial and autocorrelation plots, indicating that the lags directly beforehand influenced the CPI values.

plot(acf(resisual_Lin_seas, plot = F), xlim = c(0,3), ylim = c(-1, 1), main = " Autocorrelation in a seasonal linear model under AIC")
plot(acf(resisual_Lin_seas, plot = F, type =  "p"), xlim = c(0,3),ylim = c(-1,1), main = "Partial Autocorrelation in a linear seasonal model under AIC")






The plots above graphically summarize the strength of a relationship with an observation in a time series with observations at prior time steps.The autocorrelation with lag zero always equals 1, because this represents the autocorrelation between each term and itself. Each spike that rises above or falls below the shaded blue shapes is considered to be statistically significant. This means the spike has a value that is significantly different from zero. If a spike is significantly different from zero, that is evidence of autocorrelation. A spike that's close to zero is evidence against autocorrelation.Meanwhile, the lags within partial autocorrelation plots depend on the lag  directly beforehand. SInce the AIC was lower in value than the BIC, the AIC will be used for the correlation plots.

As one can observe, there is statistical significance in the partial and autocorrelation plots, indicating that the lags directly beforehand influenced the CPI values.

#Seasonality Test

fried(CPI, freq = 12, diff = T, residuals = F, autoarima = T)

```
Test used:  Friedman rank

Test statistic:  4.2
P-value:  0.9636688
```

Given that the p value of the Friedman test, an algorithm that determines whether a time series contains seasonality or trends, was 0.9636688; the null hypothesis has been rejected. However, since the alpha threshold of .05 is utilized in this algorithm, it is statistically significant that CPI values within the country of Canada from the years 2000 to 2009 did not have any indication of seasonality or trends within its values.As such, the alternative hypothesis being that the CPI index in Canada from 2000 to 2009 was seasonally adjusted is supported by the Freidman test.