Centro de Estatística Aplicada

Gustavo Kanno[1]
Rodrigo Marcel Araujo[2]
Victor Ribeiro Baião Decanini[3]

Junho de 2021

Sumário

[1]Número USP: 9795810
[2]Número USP: 9299208
[3]Número USP: 9790502

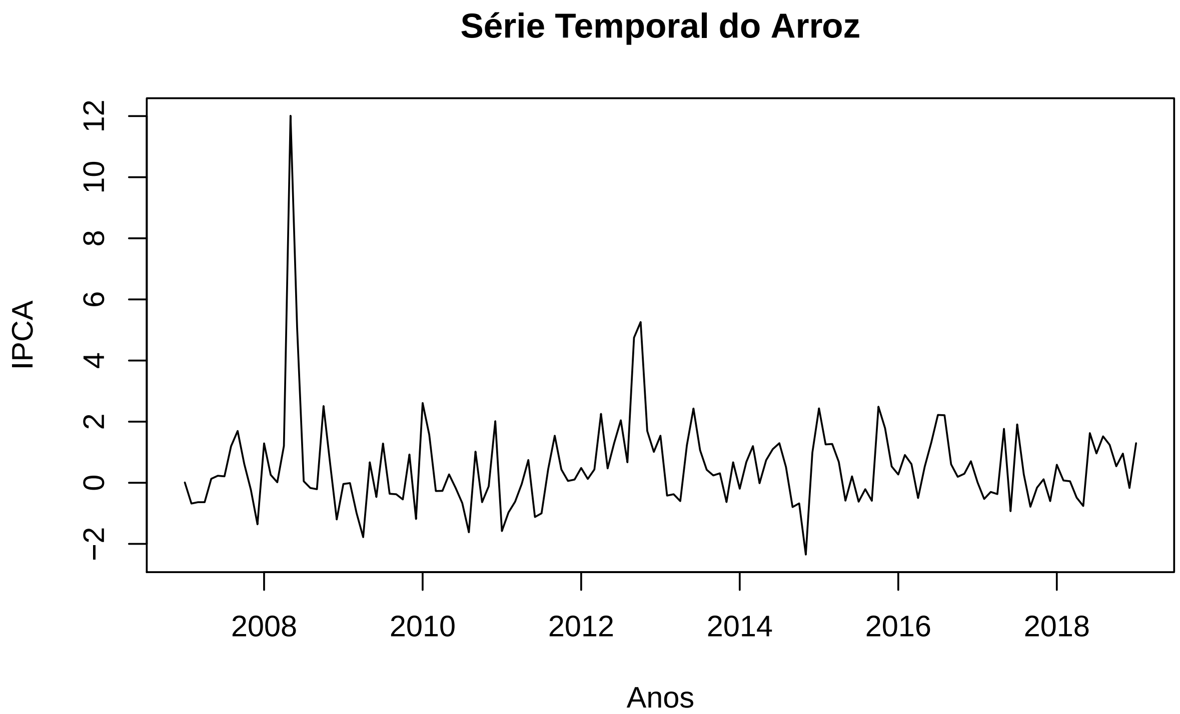
Análise Descritiva

```r
#setwd("C:\\Users\\Rodrigo Araujo\\Documents\\IME-USP\\CEA 1\\dados")
#data = read_xlsx("IPCA_DADOS_AGRUPADOS.xlsx", sheet = 1)

setwd("C:\\Users\\Rodrigo Araujo\\Documents\\IME-USP\\CEA 1\\dados")
data = read_xlsx("IPCA_DADOS_AGRUPADOS.xlsx", sheet = 1)

data$Data <- as.Date(data$Data)
head(data)
```

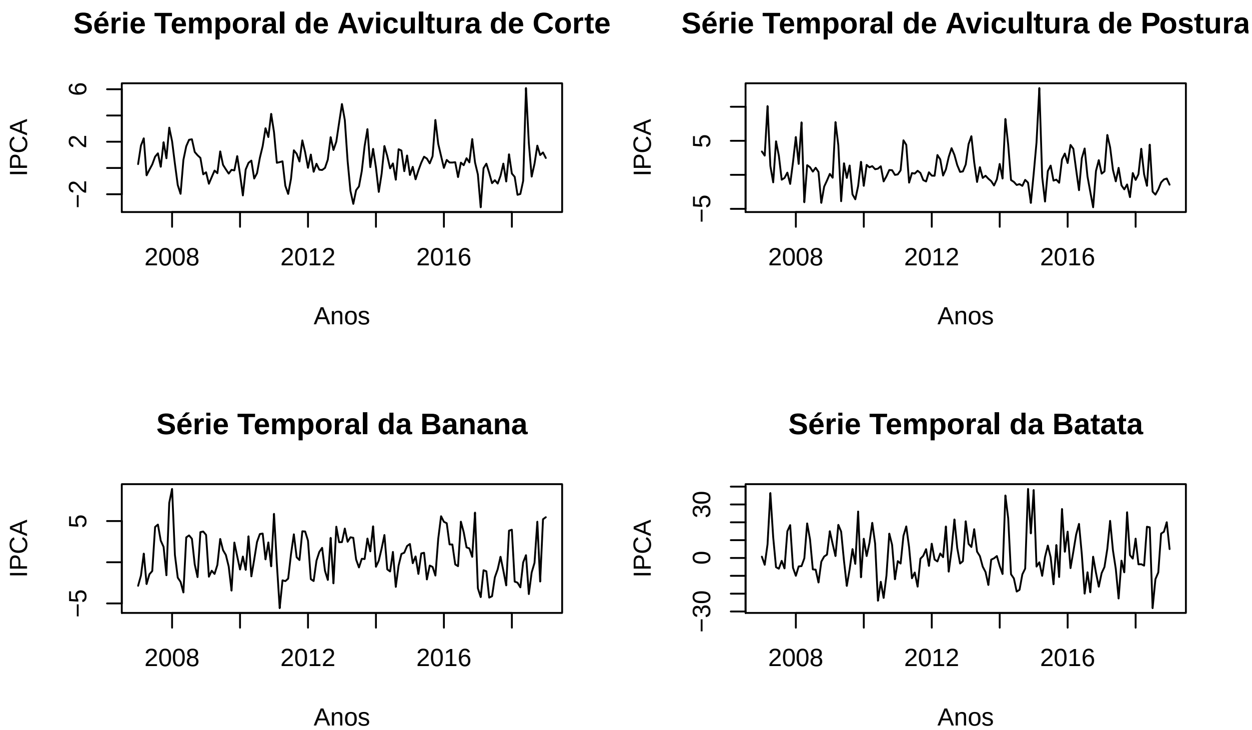
```
## # A tibble: 6 x 24
##   Data        Arroz `Avicultura de ~ `Avicultura de ~ Banana Batata
##   <date>      <dbl>            <dbl>            <dbl>  <dbl>  <dbl>
## 1 2007-01-01  0.01             0.295             3.43  -2.86   0.75
## 2 2007-02-01 -0.68             1.71              2.82  -1.62  -3.83
## 3 2007-03-01 -0.635            2.26             10.1    1.05   7.61
## 4 2007-04-01 -0.635           -0.56              1.31  -2.65  36.4
## 5 2007-05-01  0.13            -0.13             -1.11  -1.46  11.6
## 6 2007-06-01  0.230            0.27              4.93  -1.07  -5.17
## # ... with 18 more variables: Bovinocultura <dbl>, `Cacau e produtos` <dbl>,
## #   Café <dbl>, Cebola <dbl>, `Complexo soja` <dbl>, `Complexo
## #   sucroalc.` <dbl>, Feijão <dbl>, Frutas <dbl>, Hortícolas <dbl>,
## #   Indefinido <dbl>, `Laranja e citros` <dbl>, Lácteos <dbl>, Mandioca <dbl>,
## #   Milho <dbl>, Pescado <dbl>, Suinocultura <dbl>, Tomate <dbl>, Trigo <dbl>
```

```r
zt2 <- ts(data[,2], frequency = 12, start = 2007, end = 2019)
zt3 <- ts(data[,3], frequency = 12, start = 2007, end = 2019)
zt4 <- ts(data[,4], frequency = 12, start = 2007, end = 2019)
zt5 <- ts(data[,5], frequency = 12, start = 2007, end = 2019)
zt6 <- ts(data[,6], frequency = 12, start = 2007, end = 2019)
zt7 <- ts(data[,7], frequency = 12, start = 2007, end = 2019)
zt8 <- ts(data[,8], frequency = 12, start = 2007, end = 2019)
zt9 <- ts(data[,9], frequency = 12, start = 2007, end = 2019)
zt10 <- ts(data[,10], frequency = 12, start = 2007, end = 2019)
zt11 <- ts(data[,11], frequency = 12, start = 2007, end = 2019)

zt12 <- ts(data[,12], frequency = 12, start = 2007, end = 2019)
zt13 <- ts(data[,13], frequency = 12, start = 2007, end = 2019)
zt14 <- ts(data[,14], frequency = 12, start = 2007, end = 2019)
zt15 <- ts(data[,15], frequency = 12, start = 2007, end = 2019)
zt16 <- ts(data[,16], frequency = 12, start = 2007, end = 2019)
zt17 <- ts(data[,17], frequency = 12, start = 2007, end = 2019)
zt18 <- ts(data[,18], frequency = 12, start = 2007, end = 2019)
zt19 <- ts(data[,19], frequency = 12, start = 2007, end = 2019)
zt20 <- ts(data[,20], frequency = 12, start = 2007, end = 2019)
zt21 <- ts(data[,21], frequency = 12, start = 2007, end = 2019)
zt22 <- ts(data[,22], frequency = 12, start = 2007, end = 2019)
zt23 <- ts(data[,23], frequency = 12, start = 2007, end = 2019)
zt24 <- ts(data[,24], frequency = 12, start = 2007, end = 2019)
```

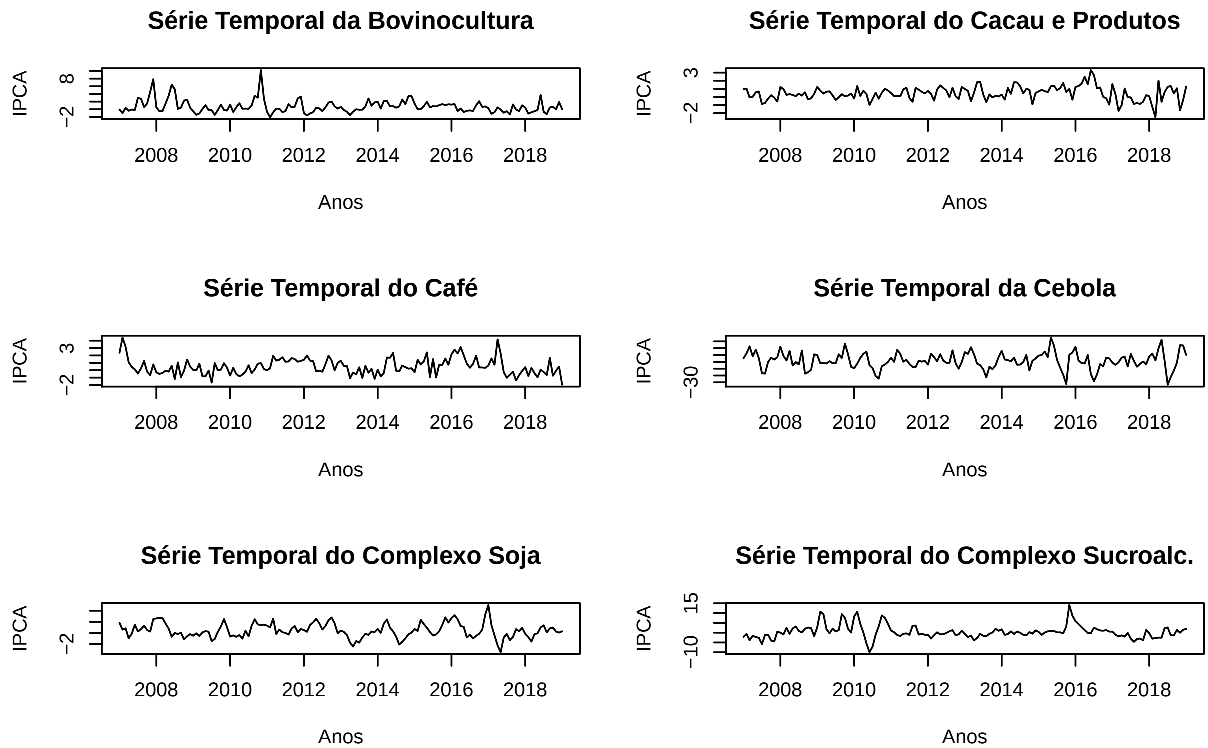
```r
plot(zt2,main="Série Temporal do Arroz", xlab= "Anos", ylab="IPCA")
```



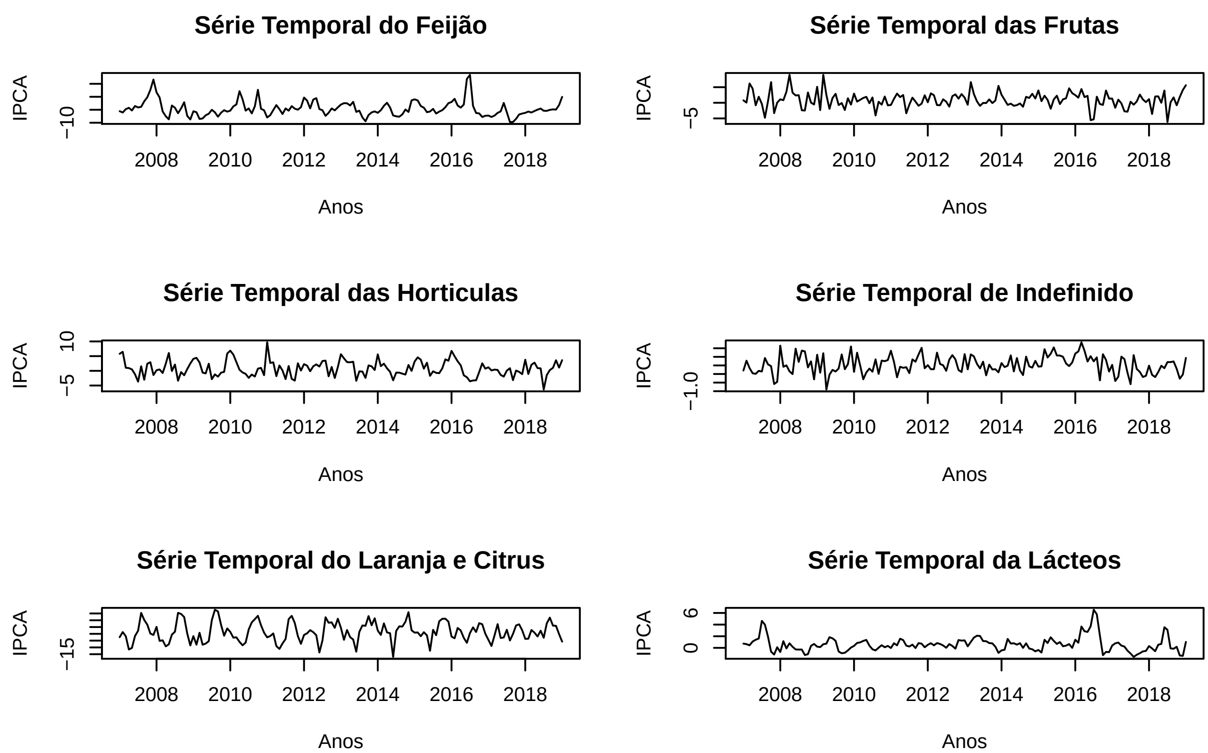

```r
par(mfrow = c(2, 2))
plot(zt3,main="Série Temporal de Avicultura de Corte", xlab= "Anos", ylab="IPCA")
plot(zt4,main="Série Temporal de Avicultura de Postura", xlab= "Anos", ylab="IPCA")
plot(zt5,main="Série Temporal da Banana", xlab= "Anos", ylab="IPCA")
plot(zt6,main="Série Temporal da Batata", xlab= "Anos", ylab="IPCA")
```



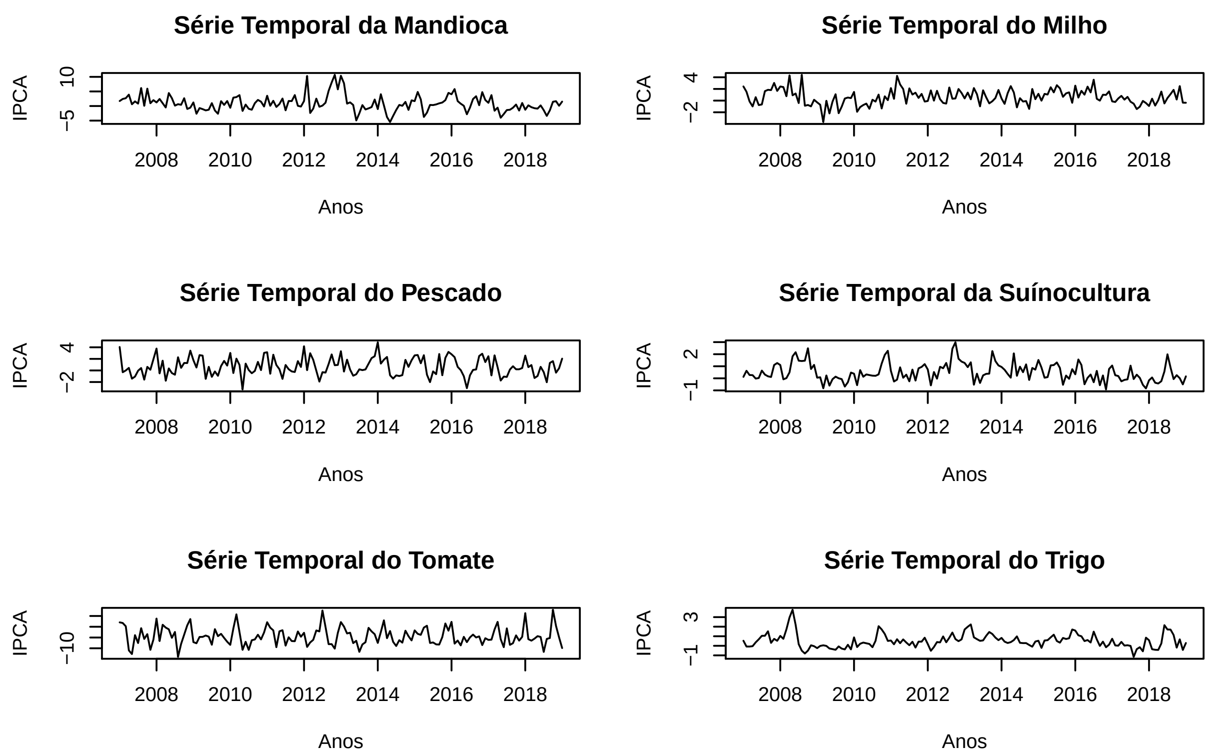

```r
par(mfrow = c(3, 2))


plot(zt7,main="Série Temporal da Bovinocultura", xlab= "Anos", ylab="IPCA")
plot(zt8,main="Série Temporal do Cacau e Produtos", xlab= "Anos", ylab="IPCA")
plot(zt9,main="Série Temporal do Café", xlab= "Anos", ylab="IPCA")
plot(zt10,main="Série Temporal da Cebola", xlab= "Anos", ylab="IPCA")
plot(zt11,main="Série Temporal do Complexo Soja", xlab= "Anos", ylab="IPCA")
plot(zt12,main="Série Temporal do Complexo Sucroalc.", xlab= "Anos", ylab="IPCA")
```



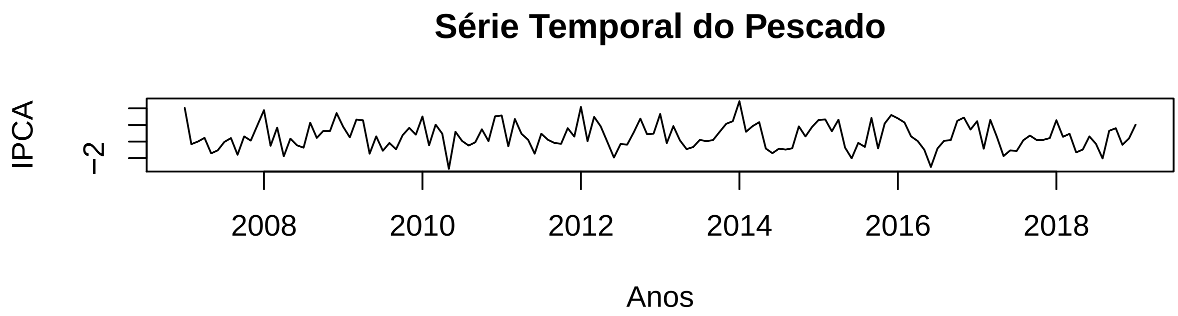

```r
par(mfrow = c(3, 2))


plot(zt13,main="Série Temporal do Feijão", xlab= "Anos", ylab="IPCA")
plot(zt14,main="Série Temporal das Frutas", xlab= "Anos", ylab="IPCA")
plot(zt15,main="Série Temporal das Horticulas", xlab= "Anos", ylab="IPCA")
plot(zt16,main="Série Temporal de Indefinido", xlab= "Anos", ylab="IPCA")
plot(zt17,main="Série Temporal do Laranja e Citrus", xlab= "Anos", ylab="IPCA")
plot(zt18,main="Série Temporal da Lácteos", xlab= "Anos", ylab="IPCA")
```

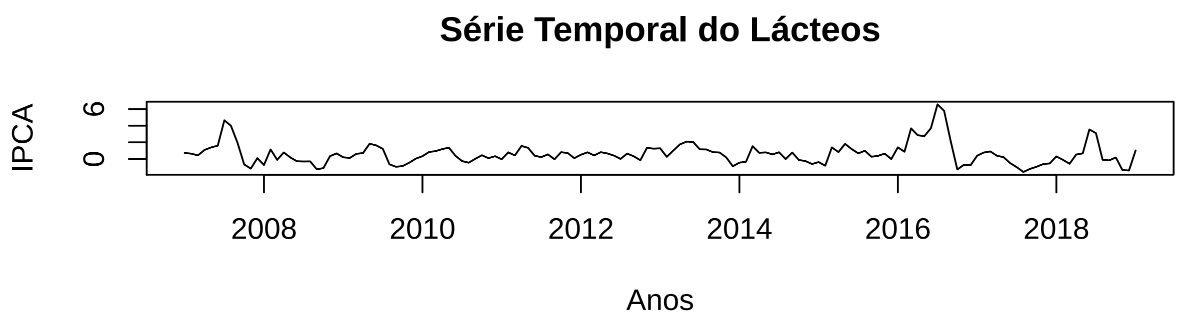



```
par(mfrow = c(3, 2))


plot(zt19,main="Série Temporal da Mandioca", xlab= "Anos", ylab="IPCA")
plot(zt20,main="Série Temporal do Milho", xlab= "Anos", ylab="IPCA")
plot(zt21,main="Série Temporal do Pescado", xlab= "Anos", ylab="IPCA")
plot(zt22,main="Série Temporal da Suínocultura", xlab= "Anos", ylab="IPCA")
plot(zt23,main="Série Temporal do Tomate", xlab= "Anos", ylab="IPCA")
plot(zt24,main="Série Temporal do Trigo", xlab= "Anos", ylab="IPCA")
```



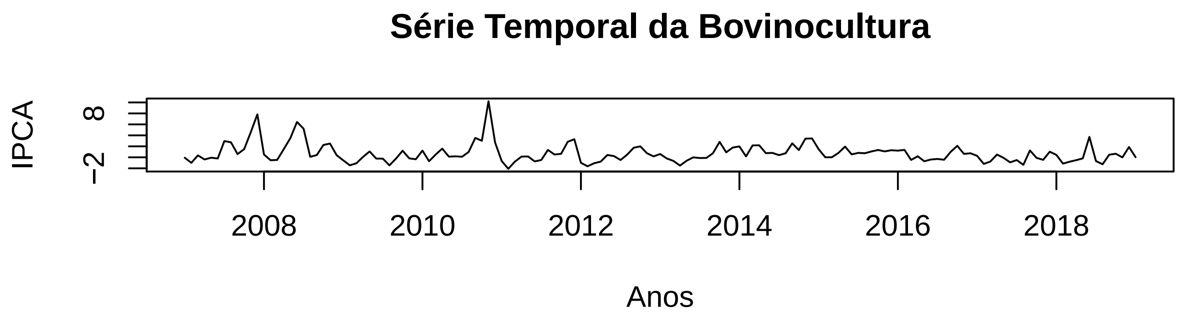

```r
par(mfrow = c(2, 1))
plot(zt21,main="Série Temporal do Pescado", xlab= "Anos", ylab="IPCA")
plot(zt18,main="Série Temporal do Lácteos", xlab= "Anos", ylab="IPCA")
```

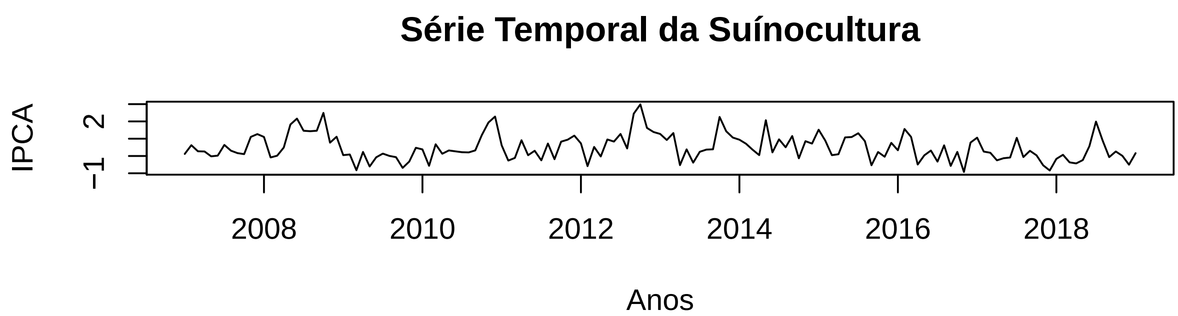




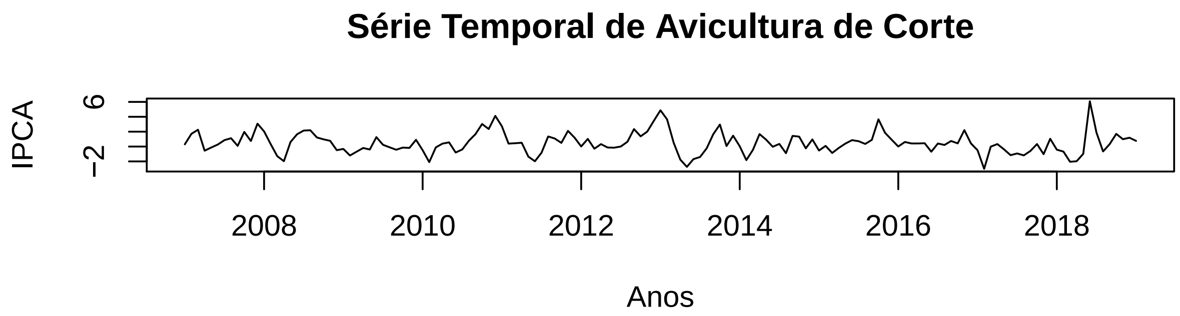

```r
#900#650
```

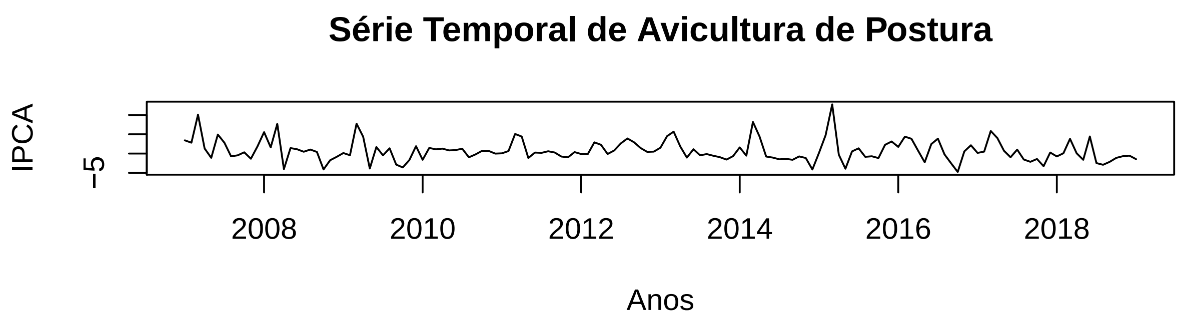
```r
par(mfrow = c(2, 1))
plot(zt7,main="Série Temporal da Bovinocultura", xlab= "Anos", ylab="IPCA")
plot(zt22,main="Série Temporal da Suínocultura", xlab= "Anos", ylab="IPCA")
```





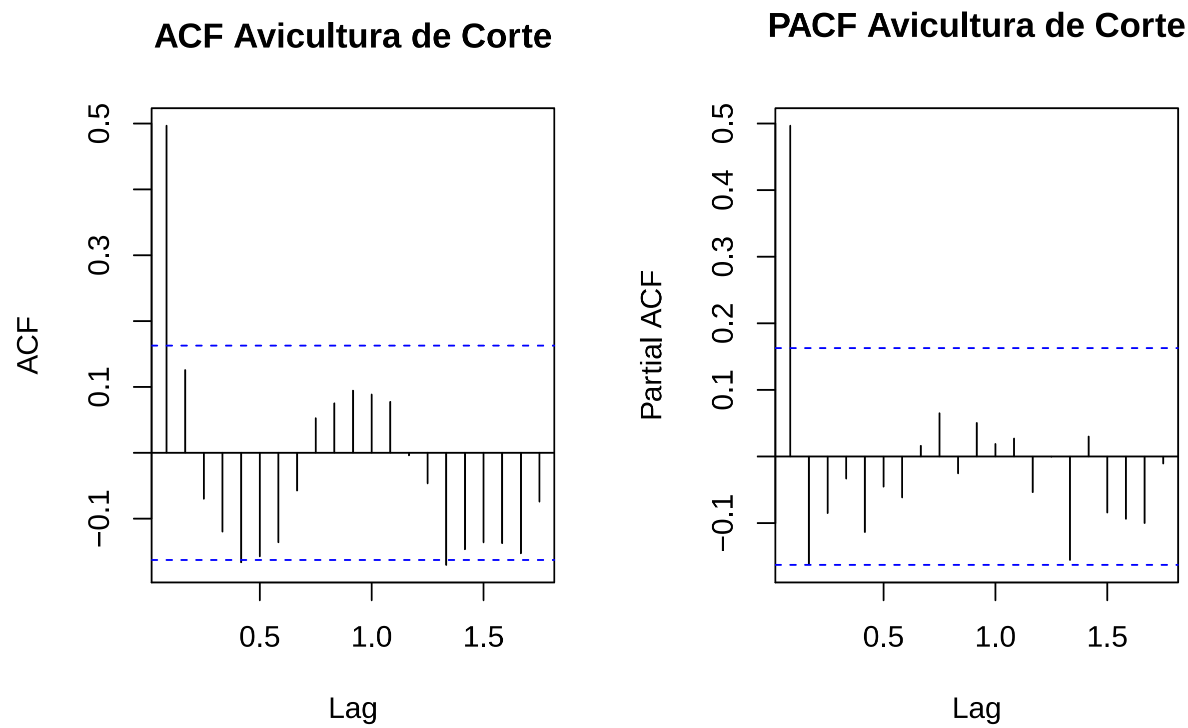

```
par(mfrow = c(2, 1))
plot(zt3,main="Série Temporal de Avicultura de Corte", xlab= "Anos", ylab="IPCA")
plot(zt4,main="Série Temporal de Avicultura de Postura", xlab= "Anos", ylab="IPCA")
```



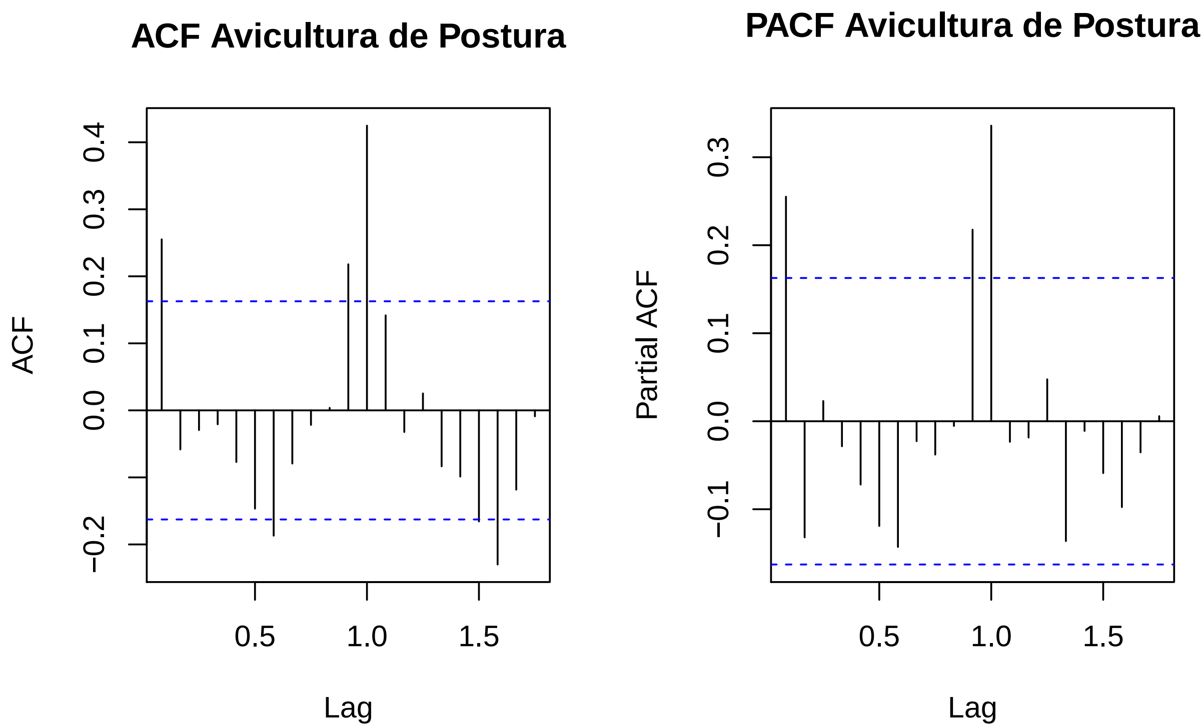


Funções de Autocorrelações

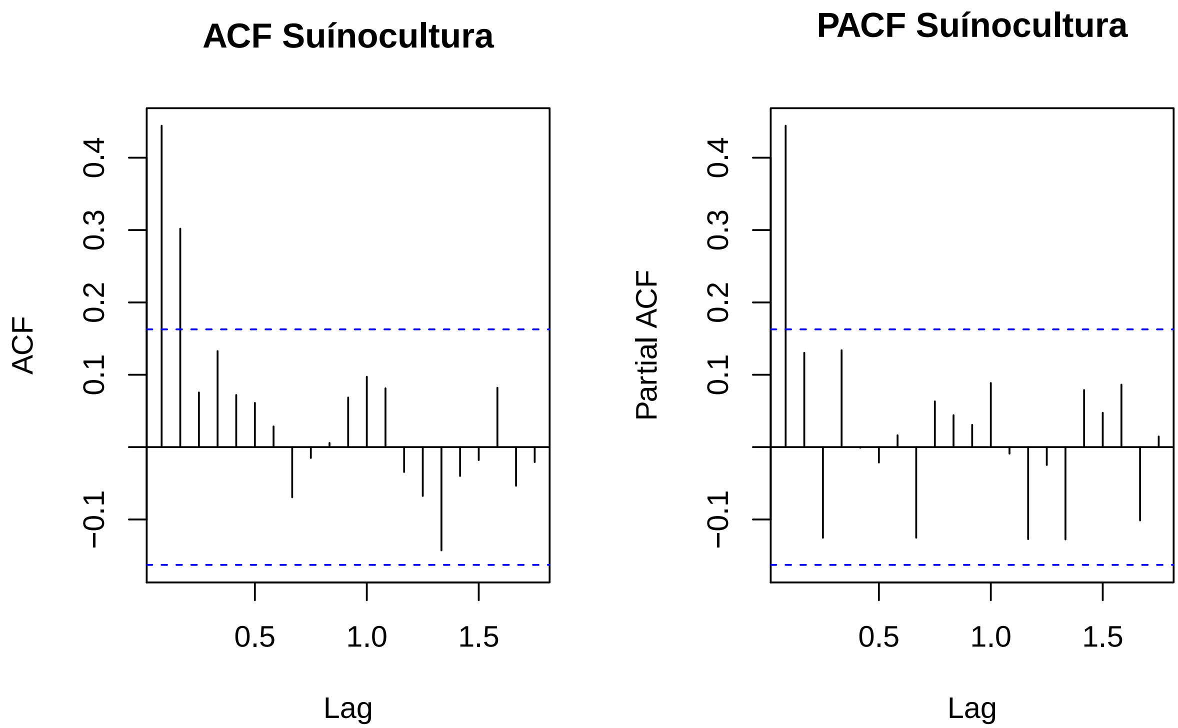
```r
par(mfrow = c(1, 2))
acf(zt3, main="ACF Avicultura de Corte")
pacf(zt3, main="PACF Avicultura de Corte")
```



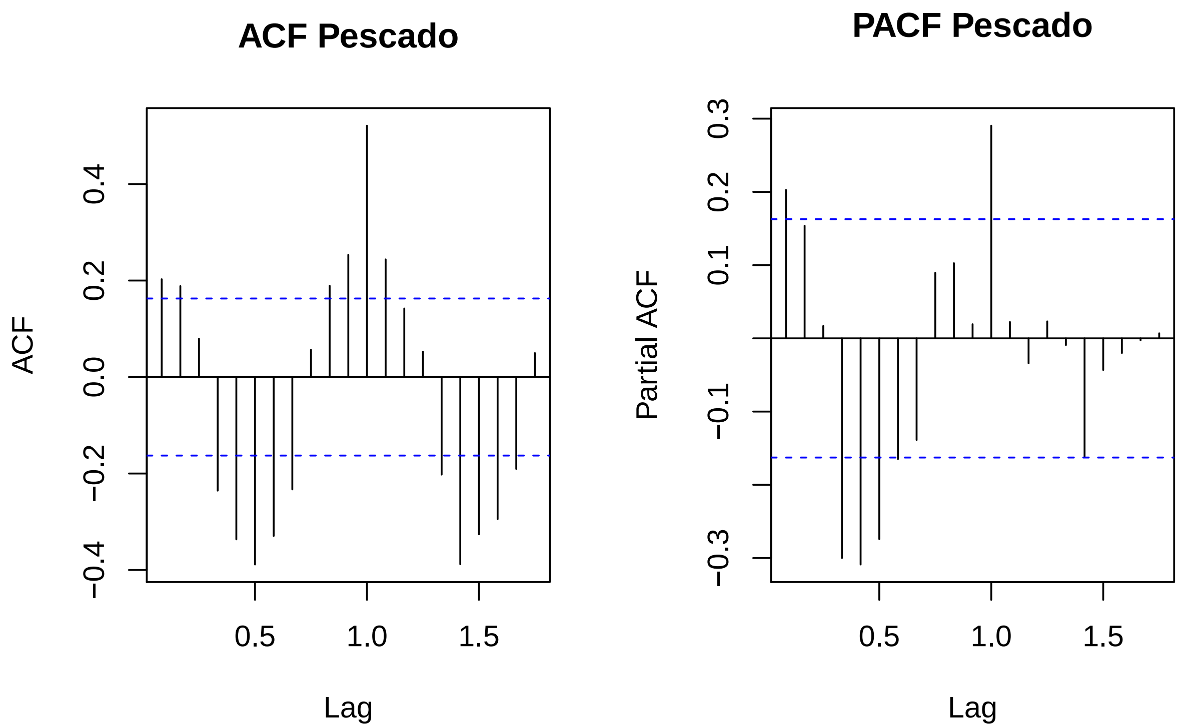

```
par(mfrow = c(1, 2))
acf(zt4, main="ACF Avicultura de Postura")
pacf(zt4, main="PACF Avicultura de Postura")
```



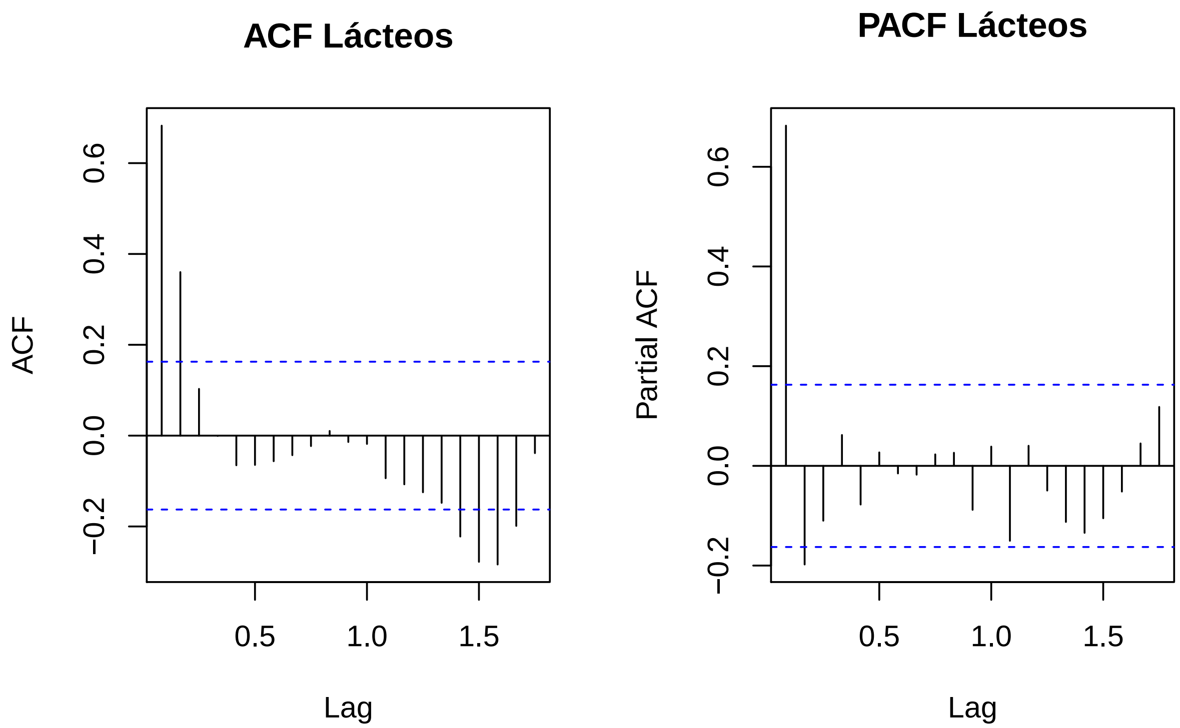

```r
par(mfrow = c(1, 2))
acf(zt22, main="ACF Suínocultura")
pacf(zt22, main="PACF Suínocultura")
```



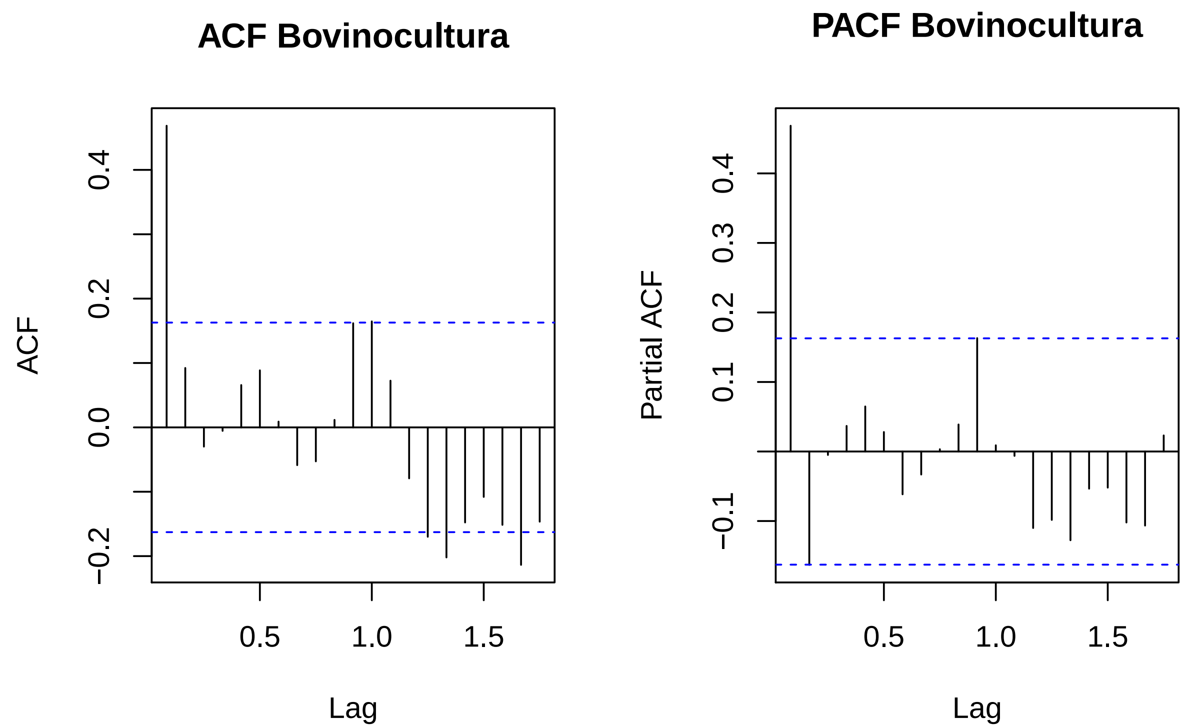

```r
par(mfrow = c(1, 2))
acf(zt21, main="ACF Pescado")
pacf(zt21, main="PACF Pescado")
```




```r
par(mfrow = c(1, 2))
acf(zt18, main="ACF Lácteos")
pacf(zt18, main="PACF Lácteos")
```

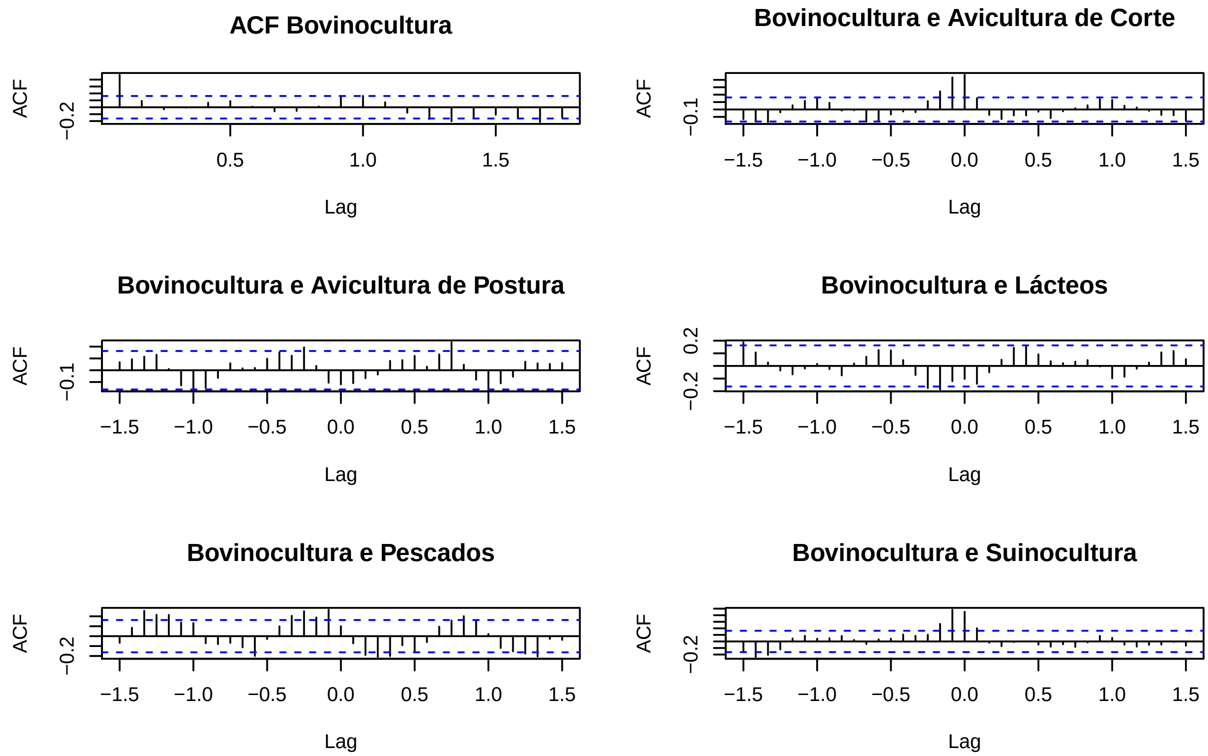



```r
par(mfrow = c(1, 2))
acf(zt7, main="ACF Bovinocultura")
pacf(zt7, main="PACF Bovinocultura")
```

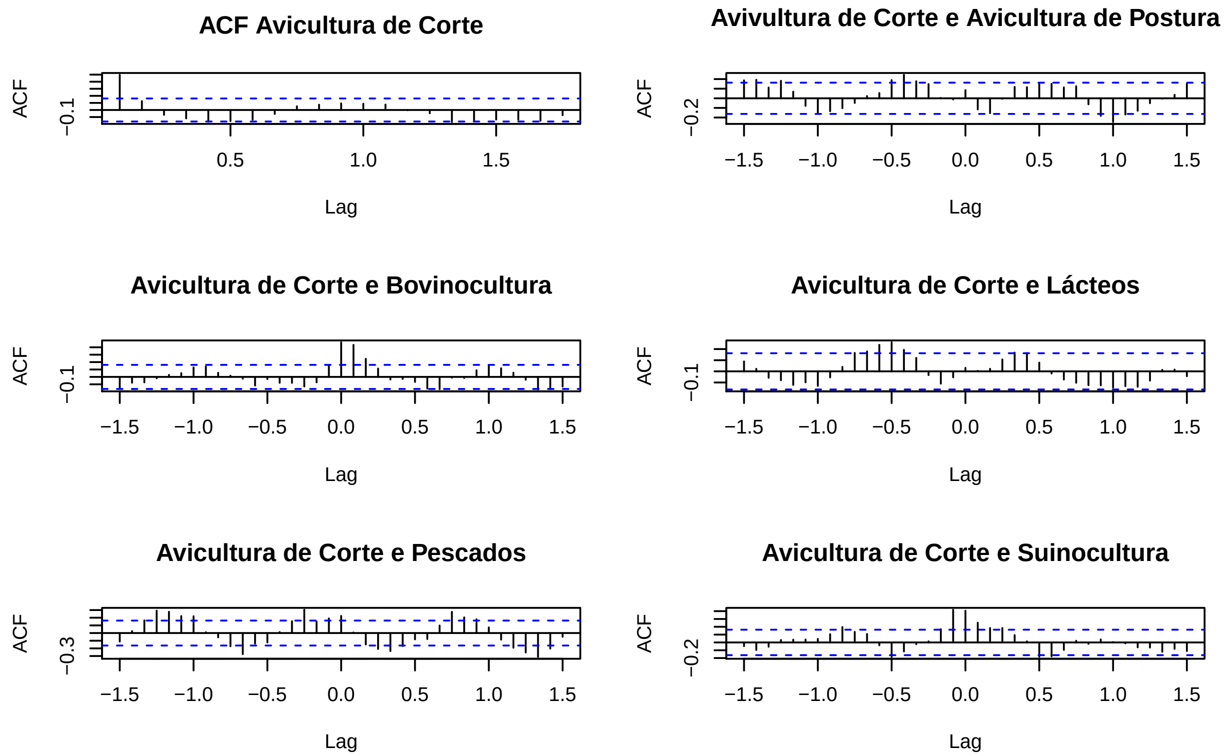


Análise Correlação Cruzada

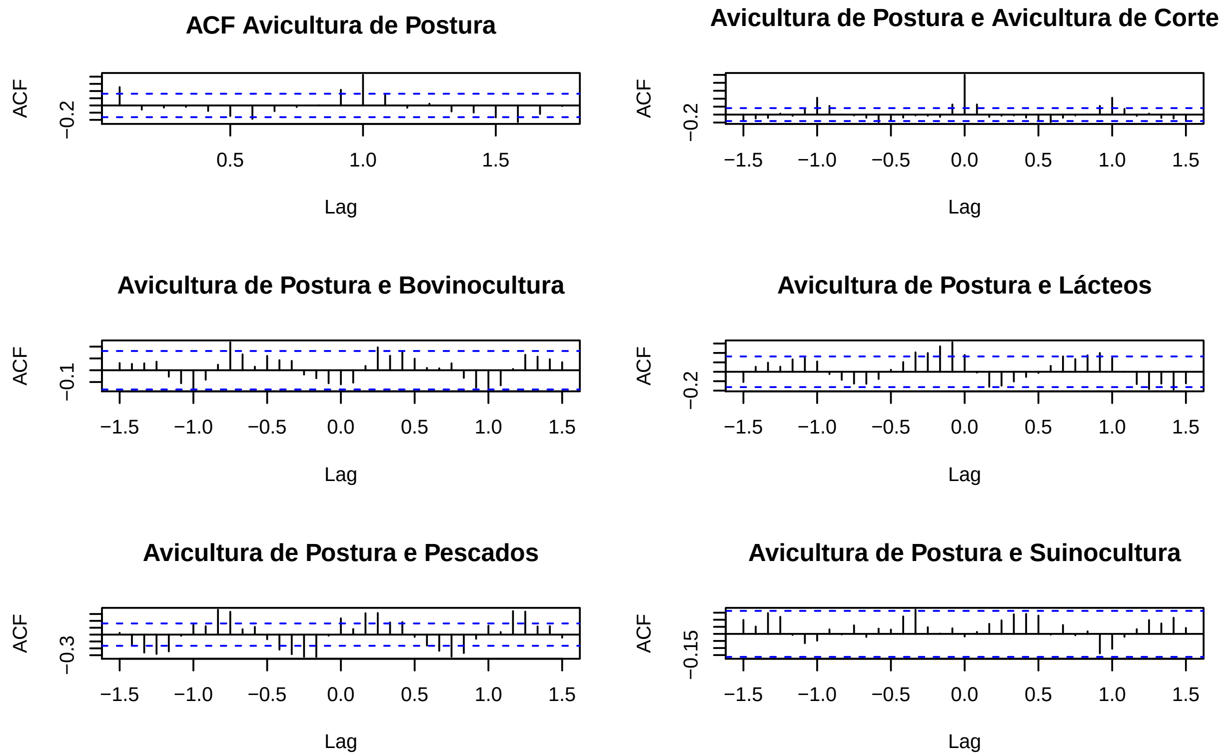
```
#Correlaões cruzadas da Bovincultura
par(mfrow = c(3,2))
acf(zt7,main="ACF Bovinocultura")
ccf(zt7,zt3,main="Bovinocultura e Avicultura de Corte")
ccf(zt7,zt4,main="Bovinocultura e Avicultura de Postura")
ccf(zt7,zt18,main="Bovinocultura e Lácteos")
ccf(zt7,zt21,main="Bovinocultura e Pescados")
ccf(zt7,zt22,main="Bovinocultura e Suinocultura")
```



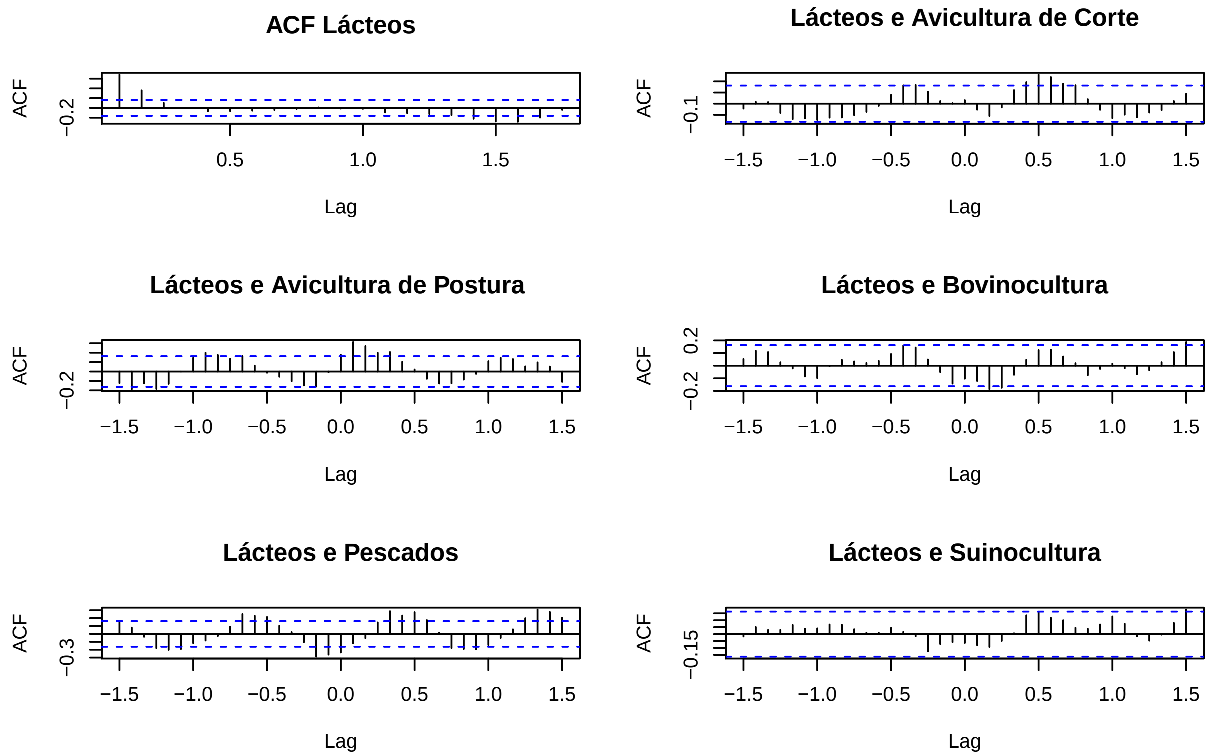

```r
#Correlações cruzadas da Avicultura de Corte
par(mfrow = c(3,2))
acf(zt3,main="ACF Avicultura de Corte")
ccf(zt3,zt4,main="Avivultura de Corte e Avicultura de Postura")
ccf(zt3,zt7,main="Avicultura de Corte e Bovinocultura")
ccf(zt3,zt18,main="Avicultura de Corte e Lácteos")
ccf(zt3,zt21,main="Avicultura de Corte e Pescados")
ccf(zt3,zt22,main="Avicultura de Corte e Suinocultura")
```



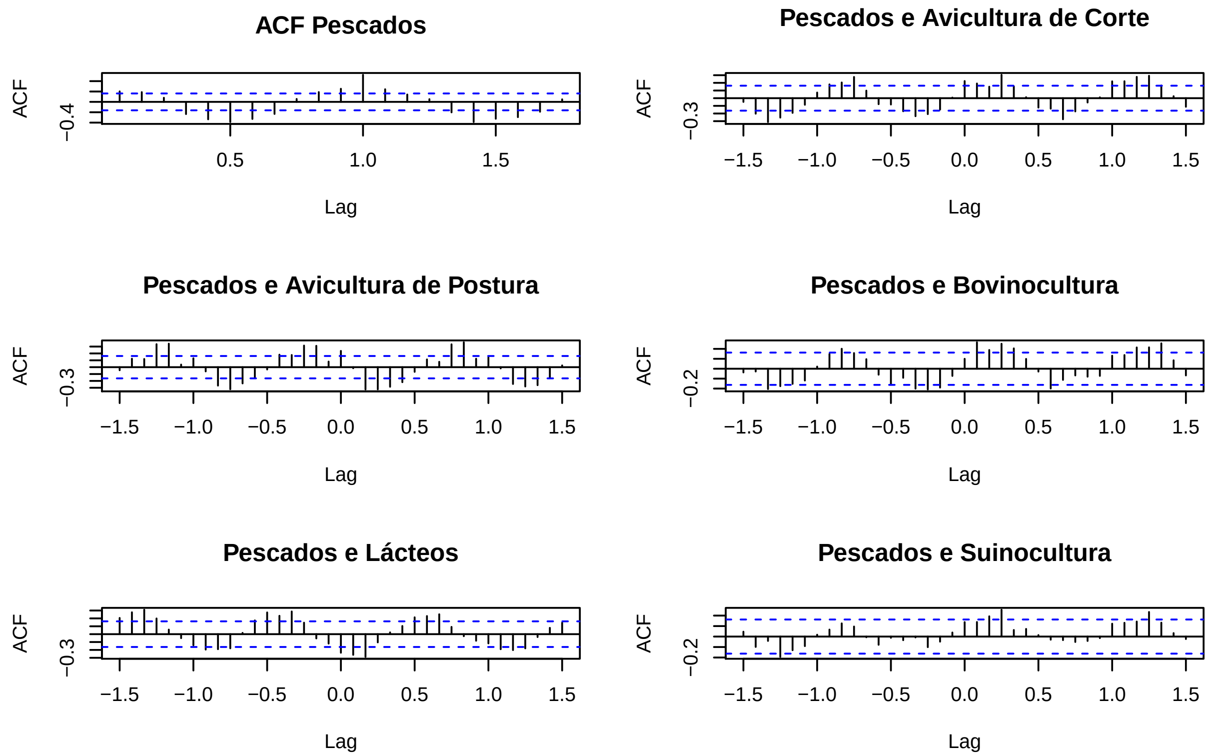

```r
#Correlações cruzadas da Avicultura de Postura
par(mfrow = c(3,2))
acf(zt4,main="ACF Avicultura de Postura")
ccf(zt4,zt4,main="Avicultura de Postura e Avicultura de Corte")
ccf(zt4,zt7,main="Avicultura de Postura e Bovinocultura")
ccf(zt4,zt18,main="Avicultura de Postura e Lácteos")
ccf(zt4,zt21,main="Avicultura de Postura e Pescados")
ccf(zt4,zt22,main="Avicultura de Postura e Suinocultura")
```



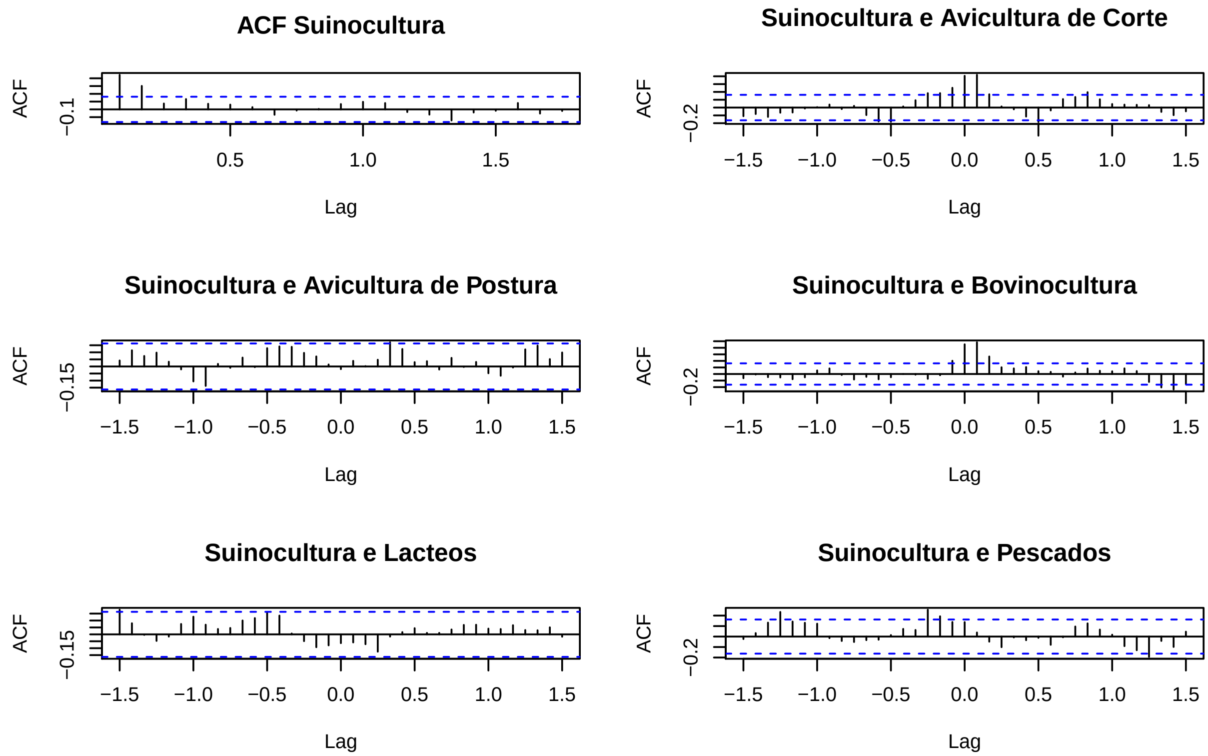

```r
#Correlações cruzadas dos Lácteos
par(mfrow = c(3,2))
acf(zt18,main="ACF Lácteos")
ccf(zt18,zt3,main="Lácteos e Avicultura de Corte")
ccf(zt18,zt4,main="Lácteos e Avicultura de Postura ")
ccf(zt18,zt7,main="Lácteos e Bovinocultura")
ccf(zt18,zt21,main="Lácteos e Pescados")
ccf(zt18,zt22,main="Lácteos e Suinocultura")
```




```r
# Correlações cruzadas dos Pescados
par(mfrow = c(3,2))
acf(zt21,main="ACF Pescados")
ccf(zt21,zt3,main="Pescados e Avicultura de Corte")
ccf(zt21,zt4,main="Pescados e Avicultura de Postura")
ccf(zt21,zt7,main="Pescados e Bovinocultura")
ccf(zt21,zt18,main="Pescados e Lácteos")
ccf(zt21,zt22,main="Pescados e Suinocultura")
```




```r
#Correlações cruzadas da Suinocultura
par(mfrow = c(3,2))
acf(zt22,main="ACF Suinocultura")
ccf(zt22,zt3,main="Suinocultura e Avicultura de Corte")
ccf(zt22,zt4,main="Suinocultura e Avicultura de Postura")
ccf(zt22,zt7,main="Suinocultura e Bovinocultura")
ccf(zt22,zt18,main="Suinocultura e Lacteos")
ccf(zt22,zt21,main="Suinocultura e Pescados")
```




```r
combinar = function(df,coluna,nome,lag){
  n = nrow(df)
  pre = rep(NA,lag)
  newcol = c(pre,coluna)
  for (k in 1:lag){
    df = rbind(df,rep(NA,ncol(df)))
  }
  df[nome] = newcol
  return (df)
}
```

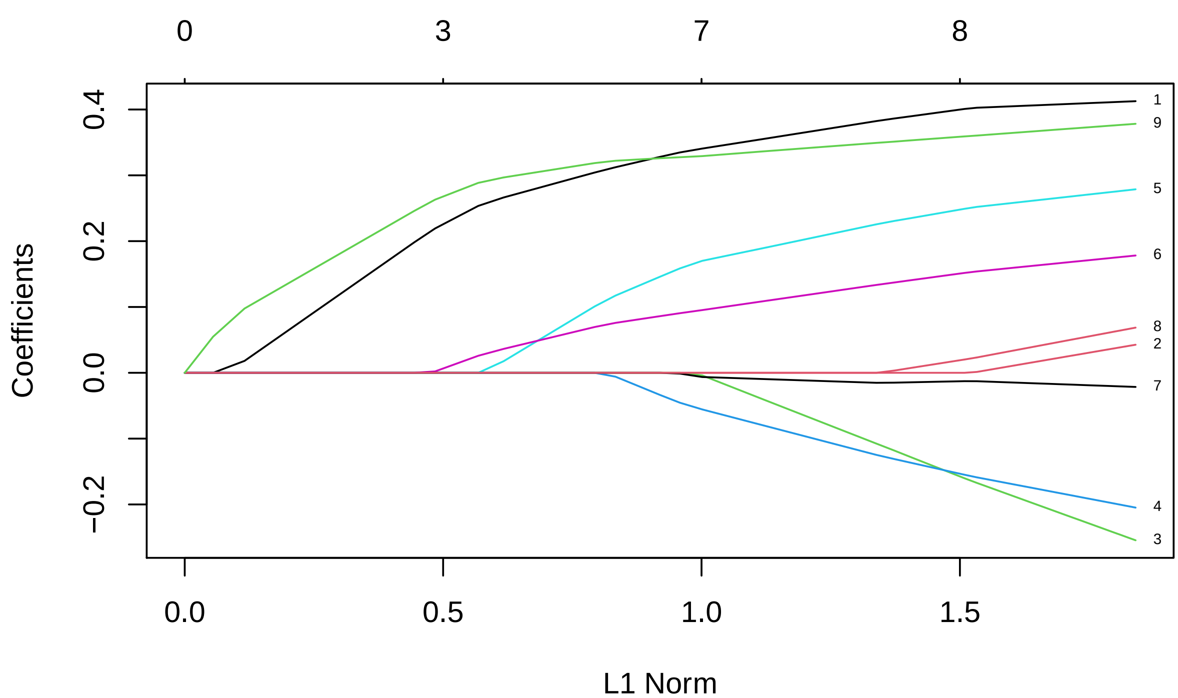
```r
data_cut = data[,c("Bovinocultura","Avicultura de Corte","Avicultura de Postura","Pescado","Lácteos","Su

df1<- combinar(data_cut, data_cut$`Avicultura de Postura`, 'avp9', 9)
df1 <- combinar(df1, df1$Pescado, 'p3', 3)
df1 <- combinar(df1, df1$Pescado, 'p10', 10)
df1 <- combinar(df1, df1$Bovinocultura, 'b1', 1)


df2 <- na.omit(df1)
```

Regressão LASSO

```r
library(glmnet)
```

```
## Loading required package: Matrix

##
## Attaching package: 'Matrix'

## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack

## Loaded glmnet 4.0-2
```

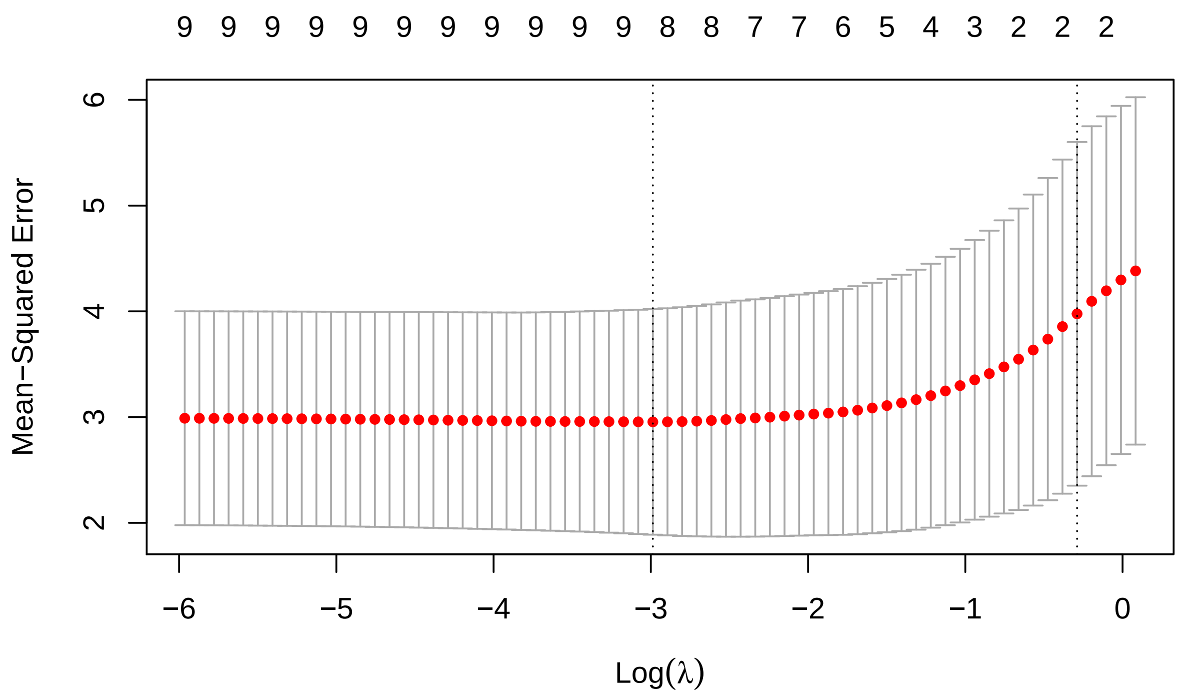
```r
x = model.matrix(Bovinocultura~.,df2)[,-1]
y = df2$Bovinocultura

grid = 10^seq(10,-2, length = 100)
lasso.mod = glmnet(x,y,alpha = 1)
plot(lasso.mod, xvar = "norm", label = TRUE)
```



```r
set.seed(123)

cv.lasso <- cv.glmnet(x, y, alpha = 1, family = "gaussian")
print(cv.lasso)
```

```
##
## Call:  cv.glmnet(x = x, y = y, alpha = 1, family = "gaussian")
##
## Measure: Mean-Squared Error
##
##     Lambda Measure    SE Nonzero
## min 0.0504   2.955 1.067       8
## 1se 0.7489   3.976 1.625       2
```

```r
plot(cv.lasso)
```




```r
cv.lasso$lambda.min
```

```
## [1] 0.05043405
```

```r
cv.lasso$lambda.1se
```

```
## [1] 0.7489297
```

```r
coef(cv.lasso, cv.lasso$lambda.min)
```

```
## 10 x 1 sparse Matrix of class "dgCMatrix"
##                                  1
## (Intercept)              0.31262107
## `Avicultura de Corte`    0.39843764
## `Avicultura de Postura`  .
## Pescado                 -0.15329263
## Lácteos                 -0.15094180
## Suinocultura             0.24605653
## avp9                     0.14932952
## p3                      -0.01311084
## p10                      0.01739267
## b1                       0.35784156
```

```r
coef(cv.lasso, cv.lasso$lambda.1se)
```

```
## 10 x 1 sparse Matrix of class "dgCMatrix"
##                                 1
## (Intercept)             0.68566274
## `Avicultura de Corte`   0.08952851
## `Avicultura de Postura` .
## Pescado                 .
## Lácteos                 .
## Suinocultura            .
## avp9                    .
## p3                      .
## p10                     .
## b1                      0.15648256
```

Regressão classifica no contexto de Séries Temporais

```r
set.seed(1234)
summary(fit <- lm(y~x))
```

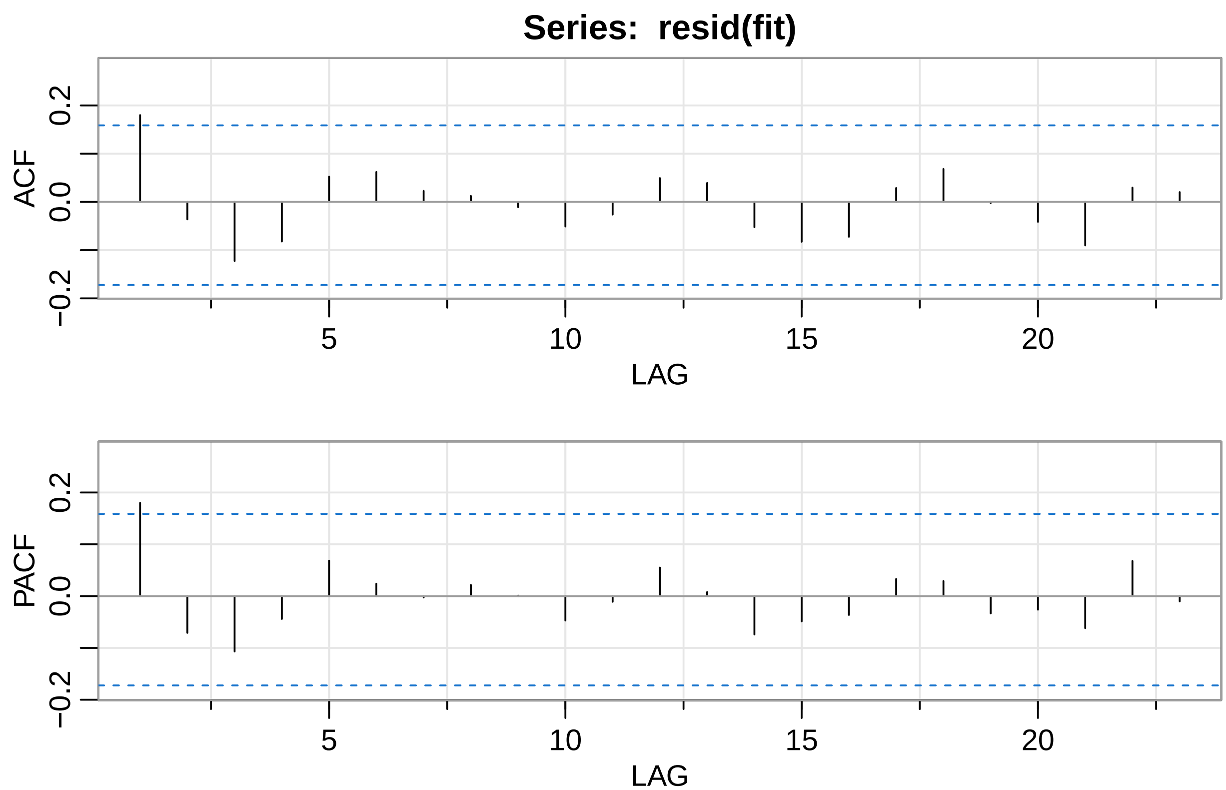
```
##
## Call:
## lm(formula = y ~ x)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.5314 -0.9189 -0.0157  0.5586  8.5757
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)               0.28536    0.20364   1.401 0.163405
## x`Avicultura de Corte`    0.41328    0.11349   3.642 0.000384 ***
## x`Avicultura de Postura`  0.04542    0.06035   0.753 0.452982
```

```
## xPescado                -0.26037    0.11194  -2.326 0.021498 *
## xLácteos                -0.20785    0.12322  -1.687 0.093939 .
## xSuinocultura            0.28048    0.21162   1.325 0.187266
## xavp9                    0.17980    0.05358   3.356 0.001026 **
## xp3                     -0.02202    0.10186  -0.216 0.829147
## xp10                     0.07166    0.10163   0.705 0.481954
## xb1                      0.37950    0.09758   3.889 0.000157 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.636 on 136 degrees of freedom
## Multiple R-squared:  0.4224, Adjusted R-squared:  0.3841
## F-statistic: 11.05 on 9 and 136 DF,  p-value: 8.134e-13
```

```r
library(astsa)
```

```
##
## Attaching package: 'astsa'

## The following objects are masked from 'package:fma':
##
##     chicken, sales

## The following object is masked from 'package:forecast':
##
##     gas

## The following object is masked from 'package:fpp2':
##
##     oil

## The following object is masked from 'package:faraway':
##
##     star

## The following object is masked from 'package:gamlss.data':
##
##     oil
```

```r
acf2(resid(fit))
```

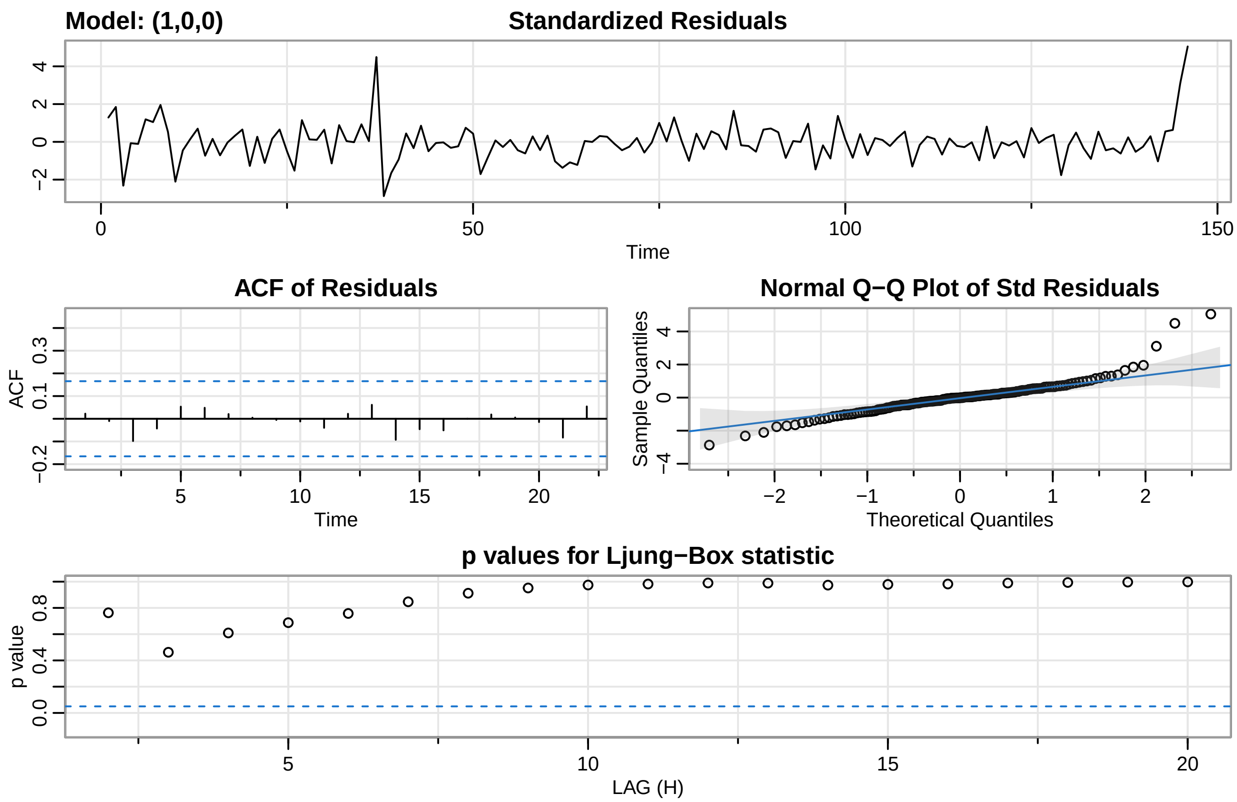



```
##      [,1]  [,2]  [,3]  [,4] [,5] [,6] [,7] [,8]  [,9] [,10] [,11] [,12] [,13]
## ACF  0.18 -0.04 -0.12 -0.08 0.05 0.06 0.02 0.01 -0.01 -0.05 -0.03  0.05  0.04
## PACF 0.18 -0.07 -0.11 -0.04 0.07 0.02 0.00 0.02  0.00 -0.05 -0.01  0.06  0.01
##      [,14] [,15] [,16] [,17] [,18] [,19] [,20] [,21] [,22] [,23]
## ACF  -0.05 -0.08 -0.07  0.03  0.07  0.00 -0.04 -0.09  0.03  0.02
## PACF -0.07 -0.05 -0.04  0.03  0.03 -0.03 -0.03 -0.06  0.07 -0.01
```

```
set.seed(12345)

fit2 <- sarima(y, 1,0,0, xreg=x)
```

```
## initial  value 0.453951
## iter   2 value 0.433468
## iter   3 value 0.420315
## iter   4 value 0.411120
## iter   5 value 0.410309
## iter   6 value 0.410165
## iter   7 value 0.410127
## iter   8 value 0.410120
## iter   9 value 0.410117
## iter  10 value 0.410117
## iter  11 value 0.410116
## iter  12 value 0.410116
## iter  12 value 0.410116
## iter  12 value 0.410116
## final  value 0.410116
```

```
## converged
## initial  value 0.414187
## iter   2 value 0.413832
## iter   3 value 0.413745
## iter   4 value 0.413714
## iter   5 value 0.413707
## iter   6 value 0.413704
## iter   7 value 0.413704
## iter   8 value 0.413704
## iter   9 value 0.413704
## iter  10 value 0.413704
## iter  10 value 0.413704
## iter  10 value 0.413704
## final  value 0.413704
## converged
```




```
fit2
```

```
## $fit
##
## Call:
## stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,
##     Q), period = S), xreg = xreg, transform.pars = trans, fixed = fixed, optim.control = list(trace =
##     REPORT = 1, reltol = tol))
##
## Coefficients:
```

```
##          ar1  intercept  `Avicultura de Corte`  `Avicultura de Postura`
##       0.4691     0.4196                 0.5589                   0.0076
## s.e.  0.1227     0.2895                 0.1131                   0.0540
##       Pescado  Lácteos  Suinocultura    avp9      p3     p10      b1
##       -0.1639  -0.1834        0.3054  0.1548  0.0282  0.1139  0.0712
## s.e.   0.0963   0.1454        0.2012  0.0492  0.0856  0.0844  0.1194
##
## sigma^2 estimated as 2.283:  log likelihood = -267.57,  aic = 559.13
##
## $degrees_of_freedom
## [1] 135
##
## $ttable
##                         Estimate     SE t.value p.value
## ar1                       0.4691 0.1227  3.8246  0.0002
## intercept                 0.4196 0.2895  1.4497  0.1495
## `Avicultura de Corte`     0.5589 0.1131  4.9432  0.0000
## `Avicultura de Postura`   0.0076 0.0540  0.1400  0.8889
## Pescado                  -0.1639 0.0963 -1.7023  0.0910
## Lácteos                  -0.1834 0.1454 -1.2614  0.2093
## Suinocultura              0.3054 0.2012  1.5182  0.1313
## avp9                      0.1548 0.0492  3.1468  0.0020
## p3                        0.0282 0.0856  0.3290  0.7427
## p10                       0.1139 0.0844  1.3496  0.1794
## b1                        0.0712 0.1194  0.5960  0.5521
##
## $AIC
## [1] 3.829668
##
## $AICc
## [1] 3.843162
##
## $BIC
## [1] 4.074896
```