Atividade 2 - Raíz Unitária

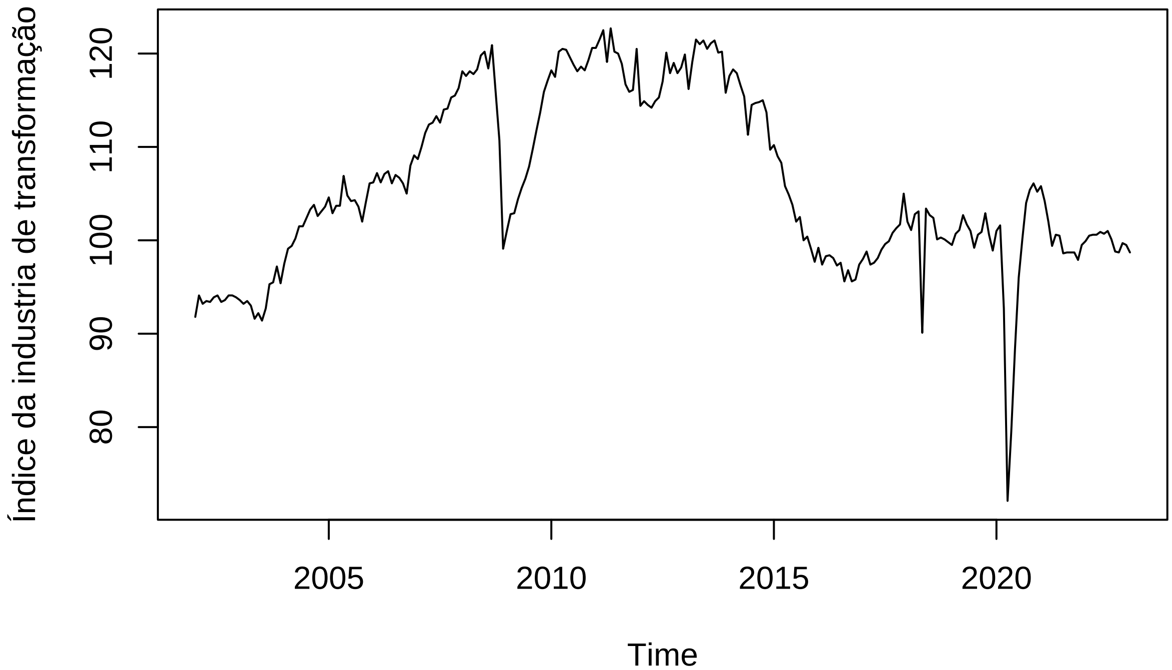
Rodrigo Cabral

21/04/2023

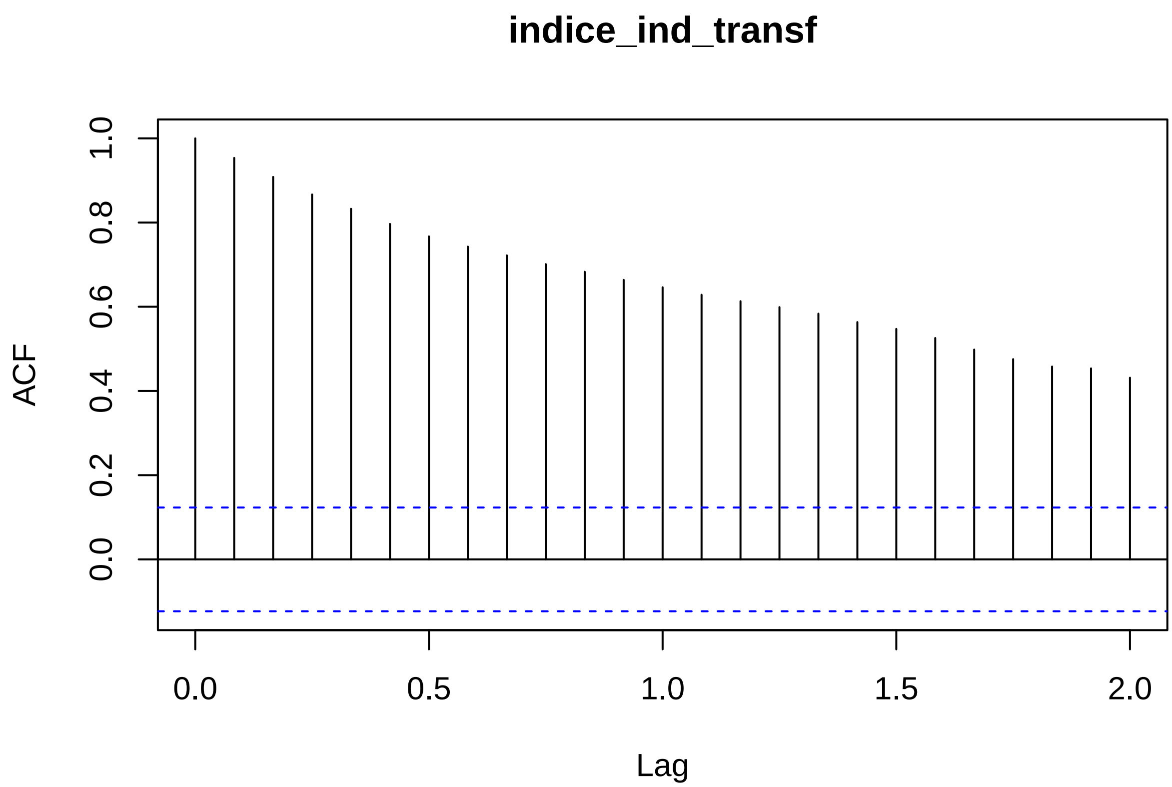
Nesta atividade foi utilizada uma série com o índice de produção da indústria de transformação brasiliera com ajuste sazonal. O período é referente a Jan/02 a Jan/23.

Lendo a série temporal

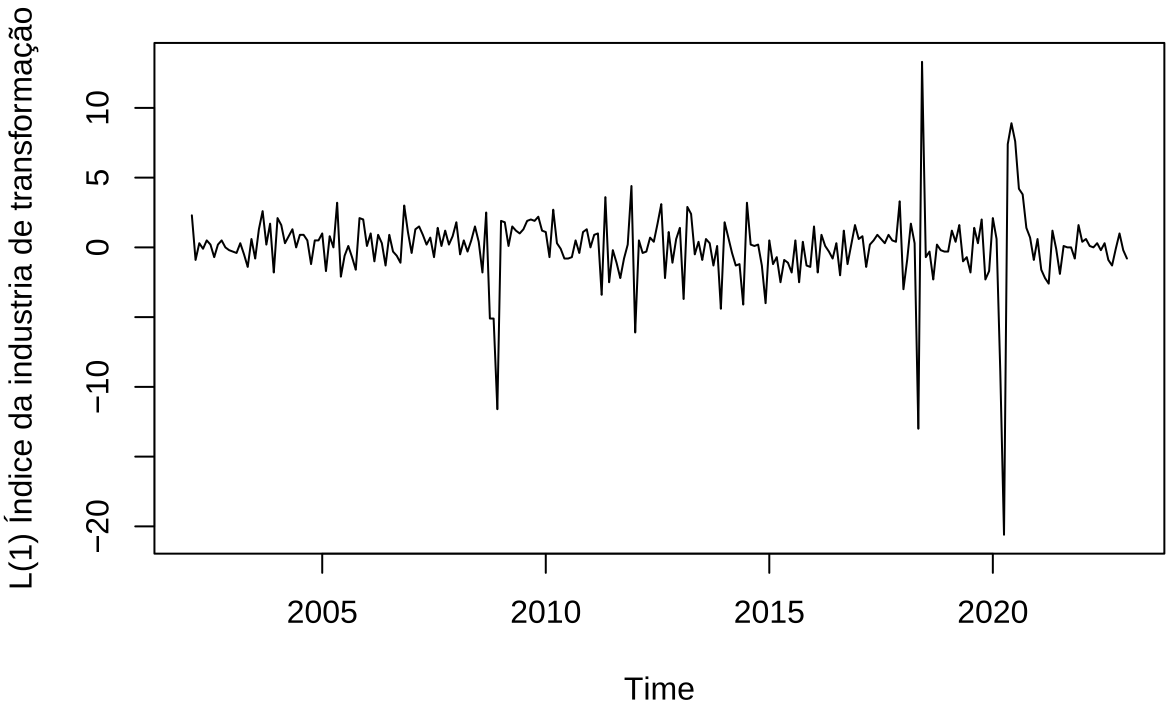
```
library(readxl)

ind=read_excel('indice_industria_transf.xlsx')

ind = ts(ind[,2],start=c(2002,1),freq=12)

plot(ind,ylab='Índice da industria de transformação')
```


Índice da industria de transformação
120
110
100
90
80
2005
2010
2015
2020
Time

```
acf(ind)
```

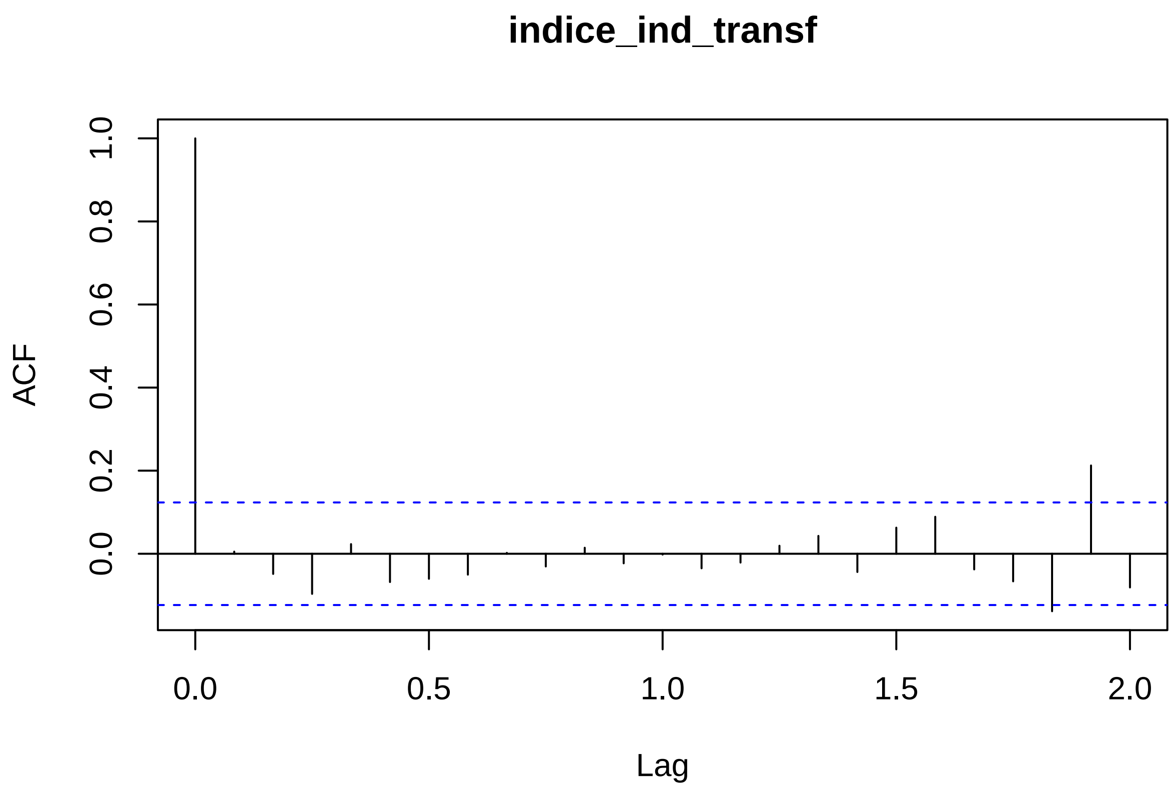



```
#Primeira Diferença
dind=diff(ind)
plot(dind,ylab='L(1) Índice da industria de transformação')
```



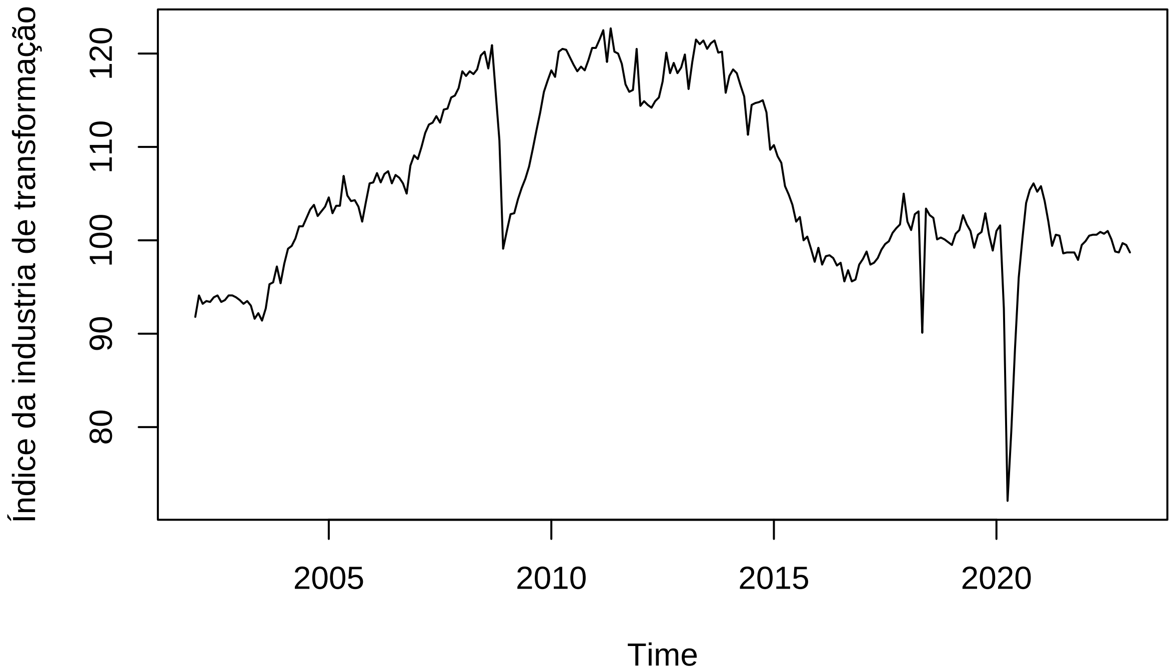
L(1) Índice da industria de transformação
10
5
0
−10
−20
2005
2010
2015
2020
Time


```
acf(dind)
```



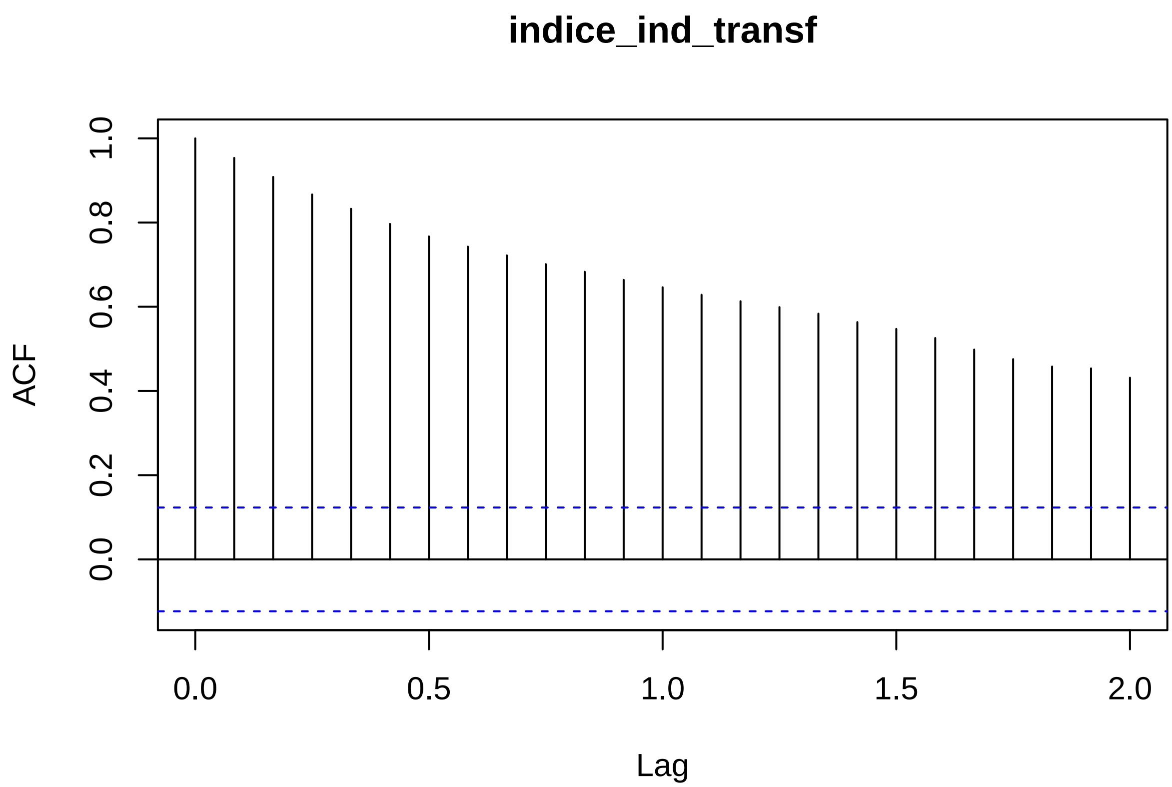
Plotando a série original, a primeira diferença e as respectivas autocorrelações.

```
plot(ind,ylab='Índice da industria de transformação')
```

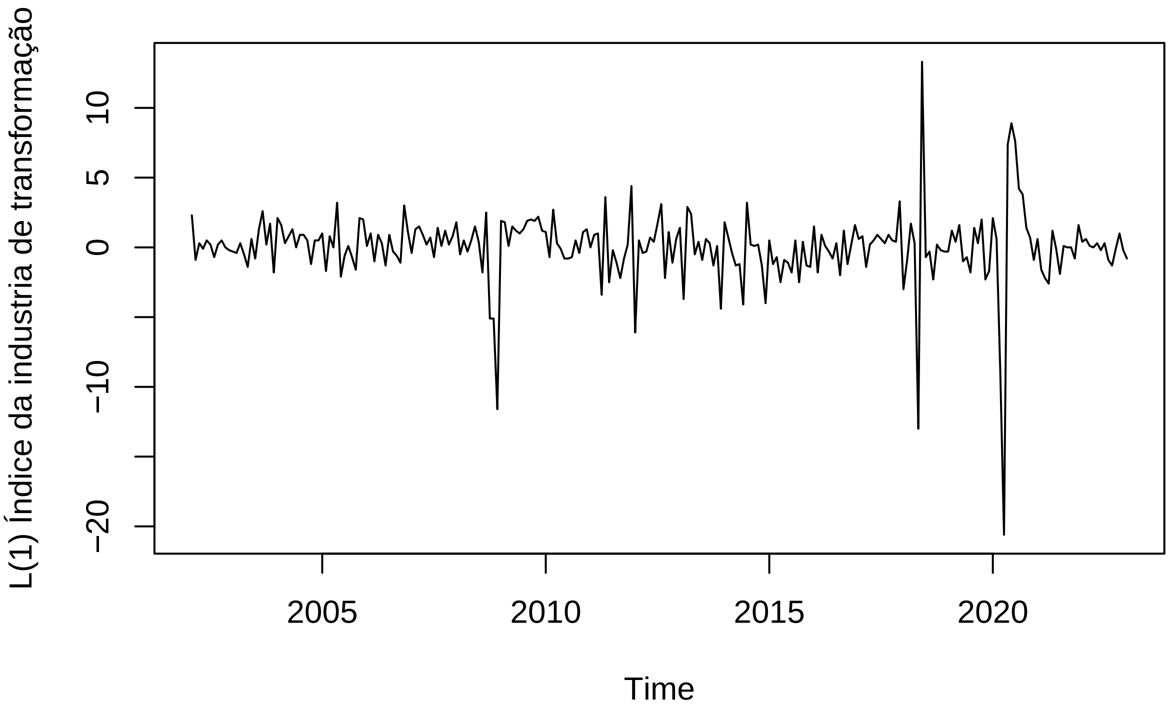


Índice da industria de transformação
80
90
100
110
120
2005
2010
2015
2020
Time

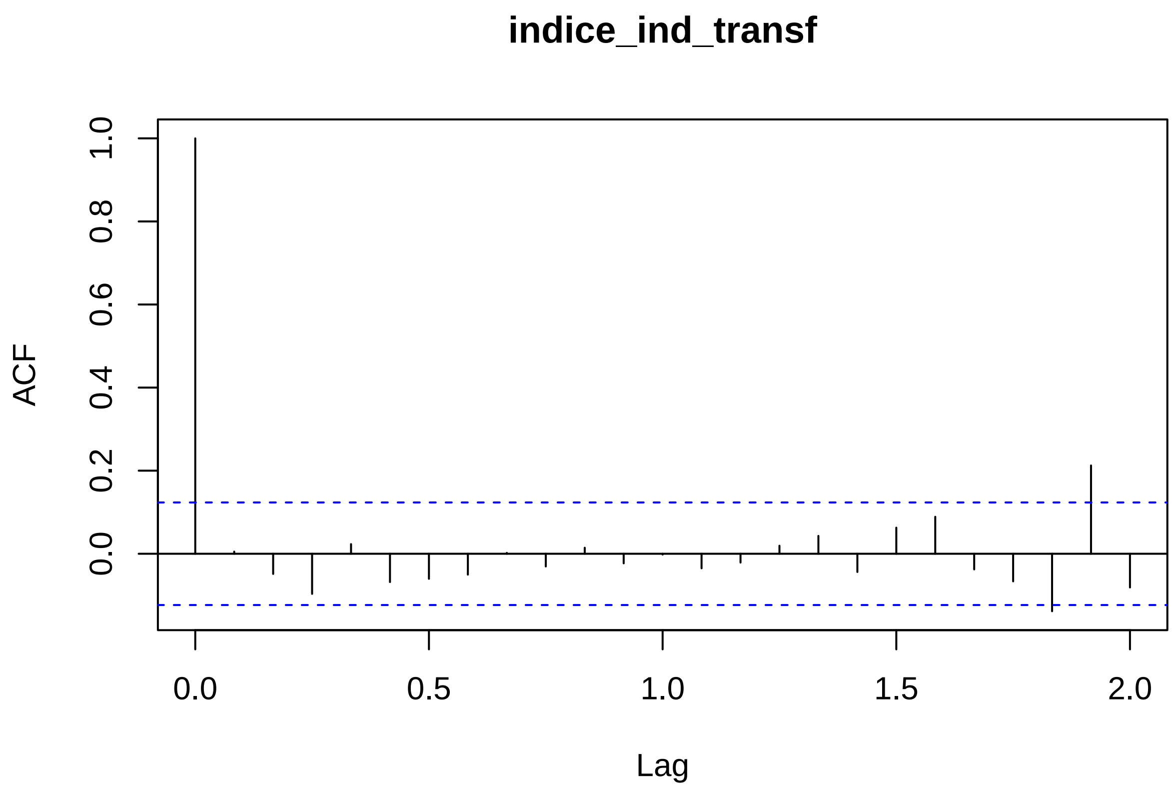

```
acf(ind)
```




```
#Primeira Diferença
dind=diff(ind)
plot(dind,ylab='L(1) Índice da industria de transformação')
```


L(1) Índice da industria de transformação
10
5
0
−10
−20
2005
2010
2015
2020
Time

```
acf(dind)
```



Nota-se que a série original possui uma forte autocorrelação com decaimento aproximandamente linear. Já a série diferenciada 1 vez, nota-se a estacionaridade da série.

Teste de Raiz Unitária

Aqui serão aplicados alguns testes de raiz unitárias. Serão eles:

- Dickey Fuller

- Dickey Fuller Aumentado

- Phillips Perron

- KPSS

- Dickey e PAntula

Vale observar que aparentemente a série não possui uma tendência explícita. Esta observação será considerada na aplicação dos testes.

Teste de RU ADF

Aplicando o teste de ADF puro e depois com lag=1, critério de informação Bayesian e drift.

```
#install.packages('urca')
library(urca)
```

```
## Warning: package 'urca' was built under R version 4.2.3
```

```
#Modelo com constante

ind.df1 =ur.df(ind)
summary(ind.df1)
```

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -20.5375  -0.7817   0.2226   1.0153  13.3841
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
```

```
## z.lag.1    -0.0001581  0.0015805  -0.100    0.920
## z.diff.lag  0.0053767  0.0633037   0.085    0.932
##
## Residual standard error: 2.665 on 249 degrees of freedom
## Multiple R-squared:  6.762e-05,  Adjusted R-squared:  -0.007964
## F-statistic: 0.00842 on 2 and 249 DF,  p-value: 0.9916
##
##
## Value of test-statistic is: -0.1
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

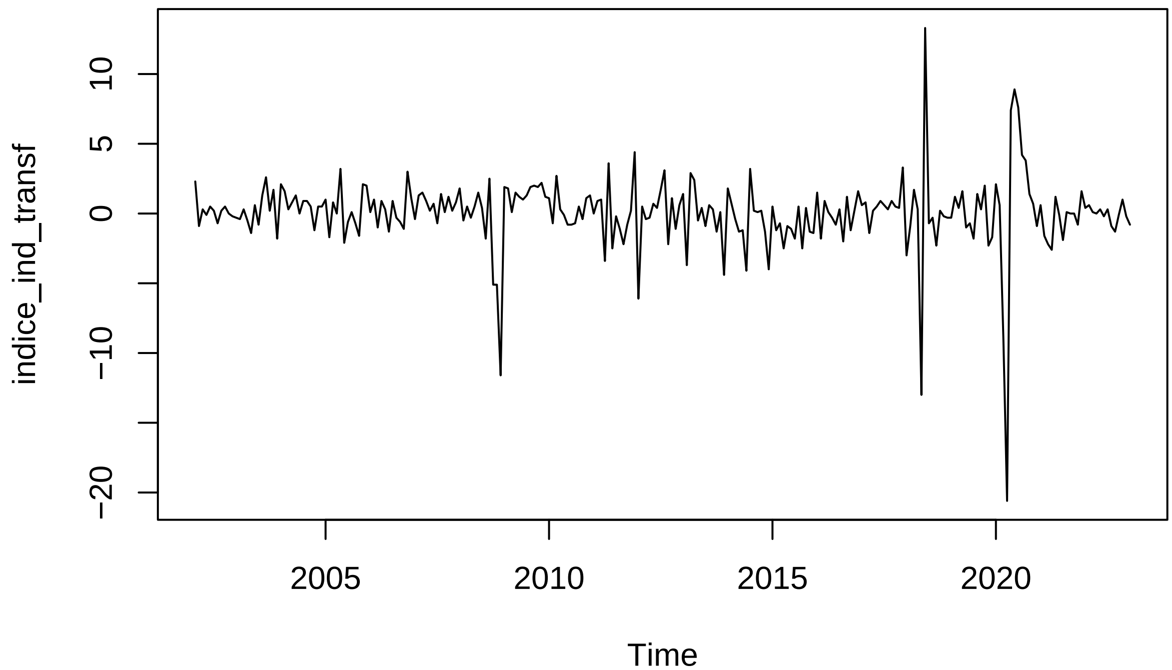
```
ind.df2 =ur.df(ind,type='drift',lags=12, selectlags = 'BIC')
summary(ind.df2)
```

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.0353  -0.8492   0.1275   1.0717  12.8587
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.19084    2.06829   2.510   0.0128 *
## z.lag.1     -0.04849    0.01933  -2.508   0.0128 *
## z.diff.lag   0.02929    0.06476   0.452   0.6515
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.695 on 237 degrees of freedom
## Multiple R-squared:  0.0259, Adjusted R-squared:  0.01768
## F-statistic: 3.151 on 2 and 237 DF,  p-value: 0.04459
##
##
## Value of test-statistic is: -2.5084 3.1534
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau2 -3.44 -2.87 -2.57
## phi1  6.47  4.61  3.79
```

Aplicando o teste ADF e analisando o resultando para a estatística T, tem-se que não pode-se rejeitar a H0, logo a série possui pelo menos uma raíz unitária.

tomando a primeira diferença e aplicando o teste ADF, tem-se:

```r
dind=diff(ind)
plot(dind)
```




```r
dind.df=ur.df(dind,type='none',lags=0)
summary(dind.df)
```

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -20.5534  -0.7987   0.2063   0.9984  13.3680
##
## Coefficients:
```

```
##         Estimate Std. Error t value Pr(>|t|)
## z.lag.1 -0.99477    0.06316  -15.75   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.66 on 250 degrees of freedom
## Multiple R-squared:  0.498,  Adjusted R-squared:  0.496
## F-statistic:   248 on 1 and 250 DF,  p-value: < 2.2e-16
##
##
## Value of test-statistic is: -15.7495
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

```r
dind.df1=ur.df(dind,type='none',lags=24,selectlags='BIC')
summary(dind.df1)
```

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -20.4995  -0.8164   0.2487   1.0452  13.4169
##
## Coefficients:
##            Estimate Std. Error t value Pr(>|t|)
## z.lag.1    -1.04406    0.09361  -11.15   <2e-16 ***
## z.diff.lag  0.05186    0.06646    0.78    0.436
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.779 on 225 degrees of freedom
## Multiple R-squared:  0.498,  Adjusted R-squared:  0.4935
## F-statistic: 111.6 on 2 and 225 DF,  p-value: < 2.2e-16
##
##
## Value of test-statistic is: -11.1529
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Para a série diferenciada uma vez L(1), nota-se que o teste de estatística T está a esquerda de tau1 5%, assim rejeita-se a hipótese nula, logo a série diferenciada uma vez não possui raíz unitária e é estacionária.

Teste de PP

Aplicando o teste de Phillips-Perron, que corrige as estatísticas convencionais de DF a partir de Z-tau-mu, tem-se:

```
#Série índice da industria de transformação
#modelo com constante
#library(urca)

ind.pp = ur.pp(ind,type='Z-tau',model='constant',lags='short')
#plot(ind.pp)
summary(ind.pp)
```

```
##
## ###################################
## # Phillips-Perron Unit Root Test #
## ###################################
##
## Test regression with intercept
##
##
## Call:
## lm(formula = y ~ y.l1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.2151  -0.8128   0.1043   1.0784  12.5701
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.70897    1.89820   2.481   0.0138 *
## y.l1         0.95584    0.01784  53.585   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.632 on 250 degrees of freedom
## Multiple R-squared:  0.9199, Adjusted R-squared:  0.9196
## F-statistic:  2871 on 1 and 250 DF,  p-value: < 2.2e-16
##
##
## Value of test-statistic, type: Z-tau  is: -2.4113
##
##          aux. Z statistics
## Z-tau-mu             2.417
##
## Critical values for Z statistics:
##                       1pct      5pct     10pct
## critical values -3.457766 -2.873097 -2.572877
```

```
ind.pp = ur.pp(ind,type='Z-alpha',model='trend',lags='short')
#plot(y.pp)
summary(ind.pp)
```

```
##
## ###################################
## # Phillips-Perron Unit Root Test #
## ###################################
##
## Test regression with intercept and trend
##
##
## Call:
## lm(formula = y ~ y.l1 + trend)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -20.9948  -0.8299   0.1365   1.0239  12.7230
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.986121   1.910496   2.610  0.00961 **
## y.l1         0.953236   0.017953  53.097  < 2e-16 ***
## trend       -0.002756   0.002294  -1.202  0.23067
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.63 on 249 degrees of freedom
## Multiple R-squared:  0.9204, Adjusted R-squared:  0.9197
## F-statistic:  1439 on 2 and 249 DF,  p-value: < 2.2e-16
##
##
## Value of test-statistic, type: Z-alpha  is: -10.9763
##
##            aux. Z statistics
## Z-tau-mu              2.9051
## Z-tau-beta           -1.2412
```

```r
ind.pp = ur.pp(ind,type='Z-tau',model='constant',lags='long')
#plot(y.pp)
summary(ind.pp)
```

```
##
## ###################################
## # Phillips-Perron Unit Root Test #
## ###################################
##
## Test regression with intercept
##
##
## Call:
## lm(formula = y ~ y.l1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.2151  -0.8128   0.1043   1.0784  12.5701
##
## Coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.70897    1.89820   2.481   0.0138 *
## y.l1         0.95584    0.01784  53.585   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.632 on 250 degrees of freedom
## Multiple R-squared:  0.9199, Adjusted R-squared:  0.9196
## F-statistic:  2871 on 1 and 250 DF,  p-value: < 2.2e-16
##
##
## Value of test-statistic, type: Z-tau  is: -2.2479
##
##          aux. Z statistics
## Z-tau-mu           2.2556
##
## Critical values for Z statistics:
##                      1pct      5pct     10pct
## critical values -3.457766 -2.873097 -2.572877
```

Nota-se que pelo teste PP, a estatística Z não está dentro do intervalo de confiança menor que 5%. Assim, não rejeita-se a hipótese nula. A série possui raiz unitária e é não-estacionária.

Teste KPSS

Já para o teste de KPSS, assume-se estacionariedade como hipótese nula.

```r
ind.kpss=ur.kpss(ind,type='mu',lags='short')
#plot(ind.kpss)
summary(ind.kpss)
```

```
##
## #######################
## # KPSS Unit Root Test #
## #######################
##
## Test is of type: mu with 5 lags.
##
## Value of test-statistic is: 0.8883
##
## Critical value for a significance level of:
##                 10pct  5pct 2.5pct  1pct
## critical values 0.347 0.463  0.574 0.739
```

```r
ind.kpss2=ur.kpss(ind,type='tau',lags='short')
#plot(ind.kpss2)
summary(ind.kpss2)
```

```
##
## #######################
```

```
## # KPSS Unit Root Test #
## #######################
##
## Test is of type: tau with 5 lags.
##
## Value of test-statistic is: 0.7832
##
## Critical value for a significance level of:
##                 10pct  5pct 2.5pct  1pct
## critical values 0.119 0.146  0.176 0.216
```

Como a estatística do teste KPSS é maior que o valor crítico, rejeita-se a hipótese nula. Ou seja, a série possui raíz unitária e é não-estacionária.

Teste de Dickey e Pantula

Utilizado para séries com maius de uma raiz unitária.

```
#install.packages('dynlm')
library(dynlm)
```

```
## Warning: package 'dynlm' was built under R version 4.2.3
```

```
## Carregando pacotes exigidos: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
#Passo 1: Teste para 2 RU
etapa1=dynlm(diff(log(ind),1,2)~trend(ind,scale=F)+diff(L(log(ind)),1,1))
summary(etapa1)
```

```
##
## Time series regression with "ts" data:
## Start = 2002(3), End = 2022(13)
##
## Call:
## dynlm(formula = diff(log(ind), 1, 2) ~ trend(ind, scale = F) +
##     diff(L(log(ind)), 1, 1))
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.249319 -0.007380  0.001795  0.008787  0.139570
##
## Coefficients:
```

```
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             2.417e-03  3.625e-03   0.667    0.505
## trend(ind, scale = F)  -1.742e-05  2.465e-05  -0.707    0.480
## diff(L(log(ind)), 1, 1) -9.937e-01  6.343e-02 -15.666   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02826 on 248 degrees of freedom
## Multiple R-squared:  0.4974, Adjusted R-squared:  0.4934
## F-statistic: 122.7 on 2 and 248 DF,  p-value: < 2.2e-16
```

```
#acf(etapa1$residuals,type='correlation')
#acf(etapa1$residuals,type='partial')

etapa2=dynlm(diff(log(ind),1,2)~diff(L(log(ind)),1,1))
summary(etapa2)
```

```
##
## Time series regression with "ts" data:
## Start = 2002(3), End = 2022(13)
##
## Call:
## dynlm(formula = diff(log(ind), 1, 2) ~ diff(L(log(ind)), 1, 1))
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.250719 -0.007914  0.002003  0.009153  0.138648
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.0001874  0.0017820   0.105    0.916
## diff(L(log(ind)), 1, 1) -0.9914594  0.0632856 -15.666   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02823 on 249 degrees of freedom
## Multiple R-squared:  0.4964, Adjusted R-squared:  0.4944
## F-statistic: 245.4 on 1 and 249 DF,  p-value: < 2.2e-16
```

```
#acf(etapa2$residuals,type='correlation')
#acf(etapa2$residuals,type='partial')

etapa3=dynlm(diff(log(ind),1,2)~diff(L(log(ind)),1,1)-1)
summary(etapa3)
```

```
##
## Time series regression with "ts" data:
## Start = 2002(3), End = 2022(13)
##
## Call:
## dynlm(formula = diff(log(ind), 1, 2) ~ diff(L(log(ind)), 1, 1) -
##     1)
##
```

```
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.250525 -0.007732  0.002192  0.009341  0.138846
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## diff(L(log(ind)), 1, 1) -0.99138    0.06316   -15.7   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02817 on 250 degrees of freedom
## Multiple R-squared:  0.4964, Adjusted R-squared:  0.4944
## F-statistic: 246.4 on 1 and 250 DF,  p-value: < 2.2e-16
```

```r
#acf(etapa3$residuals,type='correlation')
#acf(etapa3$residuals,type='partial')


#Etapa 2: Teste para 1 RU
etapa4=dynlm(diff(log(ind),1,2)~diff(L(log(ind)),1,1)+L(log(ind)))
summary(etapa4)
```

```
##
## Time series regression with "ts" data:
## Start = 2002(3), End = 2022(13)
##
## Call:
## dynlm(formula = diff(log(ind), 1, 2) ~ diff(L(log(ind)), 1, 1) +
##     L(log(ind)))
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.255571 -0.007958  0.000781  0.010317  0.133323
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)              0.25551    0.09412   2.715  0.00710 **
## diff(L(log(ind)), 1, 1) -0.96629    0.06318 -15.295  < 2e-16 ***
## L(log(ind))             -0.05479    0.02019  -2.713  0.00713 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02788 on 248 degrees of freedom
## Multiple R-squared:  0.5109, Adjusted R-squared:  0.507
## F-statistic: 129.5 on 2 and 248 DF,  p-value: < 2.2e-16
```

```r
#acf(etapa4$residuals,type='correlation')
#acf(etapa4$residuals,type='partial')
```

Aplicando o teste para duas raízes unitárias (que possuem significância econométrica) rejeita-se a hipótese nula. Mas para o passo dois não rejeita-se a hipótese nula. Assim, nota-se que há uma raíz unitária.

Teste de raiz unitária sazonal

Mesmo em séries sazonalmente ajustadas em alguns casos há presença de componente sazonal. Assim, será aplicada a essa série o teste de raiz unitária sazonal.

```r
#install.packages('uroot')
library(uroot)

#deterministic = a vector of length three containing zeros or ones to indicate, respectively, whether a
sazonal.ur=hegy.test(ind, deterministic = c(1,1,1), lag.method = "fixed", maxlag = 1)

sazonal.ur=hegy.test(ind,deterministic = c(1,1,1), lag.method = "BIC", maxlag = 12)
sazonal.ur$fitted.model
```

```
##
## Call:
## lm(formula = dx ~ 0 + ypi + xreg)
##
## Coefficients:
##       ypiYpi1         ypiYpi2         ypiYpi3         ypiYpi4         ypiYpi5
##     -0.003679       -0.132829       -0.030563       -0.068288       -0.089135
##       ypiYpi6         ypiYpi7         ypiYpi8         ypiYpi9        ypiYpi10
##     -0.137556       -0.136155       -0.153680       -0.251925        0.188607
##      ypiYpi11        ypiYpi12     xregxreg.c  xregxreg.trend      xregSD.SD2
##     -0.318281        0.073613        5.529686       -0.004296       -0.327284
##    xregSD.SD3      xregSD.SD4      xregSD.SD5      xregSD.SD6      xregSD.SD7
##     -0.413544       -1.388056       -0.258513        0.307683       -0.057947
##    xregSD.SD8      xregSD.SD9     xregSD.SD10     xregSD.SD11     xregSD.SD12
##      0.219653       -0.159699       -0.470497       -0.178071       -0.588727
```

```r
sazonal.ur$statistics
```

```
##        t_1        t_2      F_3:4      F_5:6      F_7:8     F_9:10    F_11:12
##  -2.069832  -3.897781  18.904758  23.102145  19.838505  24.027873  20.667193
##     F_2:12     F_1:12
## 128.146864 119.765446
```

```r
sazonal.ur$pvalues
```

```
##          t_1          t_2        F_3:4        F_5:6        F_7:8       F_9:10
## 0.4521054885 0.0004059189 0.0000000000 0.0000000000 0.0000000000 0.0000000000
##      F_11:12       F_2:12       F_1:12
## 0.0000000000 0.0000000000 0.0000000000
```

```r
#Teste DHF
#Instalando e carregando pacote de Dummies sazonais
#install.packages("gets")
library(gets)
```

```
## Warning: package 'gets' was built under R version 4.2.3
```

```
## Carregando pacotes exigidos: parallel
```

```r
teste.sazonal=dynlm(diff(ind,1,4)~trend(ind,scale=F)+ season(ind) + L(ind,4)+diff(L(ind,1:4),1,12))
summary(teste.sazonal)
```

```
##
## Time series regression with "ts" data:
## Start = 2003(5), End = 2023(1)
##
## Call:
## dynlm(formula = diff(ind, 1, 4) ~ trend(ind, scale = F) + season(ind) +
##     L(ind, 4) + diff(L(ind, 1:4), 1, 12))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -78.989  -5.323   0.282   4.456  52.060
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)               3.6307740 10.4859116   0.346   0.7295
## trend(ind, scale = F)    -0.0006495  0.0118637  -0.055   0.9564
## season(ind)Feb           -3.7576638  3.9039553  -0.963   0.3368
## season(ind)Mar           -0.4790273  3.8915773  -0.123   0.9021
## season(ind)Apr           -2.9027457  3.8805650  -0.748   0.4552
## season(ind)May            1.2296237  3.8431776   0.320   0.7493
## season(ind)Jun           -4.2872875  3.8329277  -1.119   0.2646
## season(ind)Jul           -2.5006849  3.8382511  -0.652   0.5154
## season(ind)Aug            0.1561993  3.8439690   0.041   0.9676
## season(ind)Sep           -3.2196321  3.8374598  -0.839   0.4024
## season(ind)Oct           -1.1678699  3.8392402  -0.304   0.7613
## season(ind)Nov           -0.8832010  3.8415515  -0.230   0.8184
## season(ind)Dec           -3.3332885  3.8478904  -0.866   0.3873
## L(ind, 4)                -0.0166759  0.0894249  -0.186   0.8522
## diff(L(ind, 1:4), 1, 12)1 -0.0068458  0.0027366  -2.502   0.0131 *
## diff(L(ind, 1:4), 1, 12)2 -0.0132739  0.0066908  -1.984   0.0485 *
## diff(L(ind, 1:4), 1, 12)3 -0.0102290  0.0066892  -1.529   0.1277
## diff(L(ind, 1:4), 1, 12)4 -0.0028539  0.0027345  -1.044   0.2978
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.09 on 219 degrees of freedom
## Multiple R-squared:  0.06602,    Adjusted R-squared:  -0.006479
## F-statistic: 0.9106 on 17 and 219 DF,  p-value: 0.5621
```

```r
#acf(teste.sazonal$residuals)
#pacf(teste.sazonal$residuals)
```

Para o teste de raíz unitária sazonal, conclui-se que existe uma raíz unitária não sazonal.