Trabalho de Séries temporais

Sobre os dados

Nosso conjunto de dados é referente aos financiamentos feitos pelo Banco Nacional do Desenvolvimento (BNDES) de forma indireta, ou seja, o cliente solicita o financiamento por meio de uma instituição financeira credenciada e não diretamente ao BNDES

No financiamento indireto pois quem assue o risco do não pagamento passa a ser a instituição financeira credenciada em questão, que também negocia os prazos e condições do financiamento, assim como cabe a ela aceitar ou não o financiamento.

O financiamento indireto pode ser feita de forma automática, ou seja, que não precisa passar por homologação do BNDES ou não automática, aonde o banco avalia cada solicitação individualmente. Para o financiamento não automático há um valor mínimo de 20 milhões de reais.

Nossa base de dados consiste de informações de financiamentos indiretos de forma não automática informados pelo BNDES entre janeiro de 2017 a janeiro de 2024 2024. Nossa planilha originalmente constava com detalhes de 321272 operações. Para esse trabalho agregamos o número de operações e analisaremos elas em caráter mensal.

```
#pacotes
install.packages("remotes")
install.packages("haven")
install.packages("forecast")
```

Mostrar saída oculta

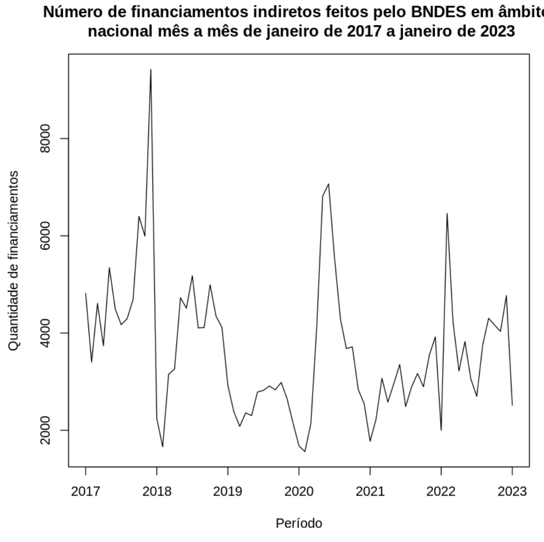
```
#pacotes
library(remotes)
library(haven)
library(readr)
library(forecast)
library(lmtest)
```

Mostrar saída oculta

```
dados <-  read_csv("https://raw.githubusercontent.com/antoniaxavier/trabalho_series_temporais/main/dados_series.csv")
dados <-  dados$V1
serie <-  ts(dados, start = c(2017,1), end = c(2023, 1), frequency = 12 )
```

Mostrar saída oculta

```
plot(serie, main = "Número de financiamentos indiretos feitos pelo BNDES em âmbito \n nacional mês a mês de janeiro de 2017 a janeiro de
```



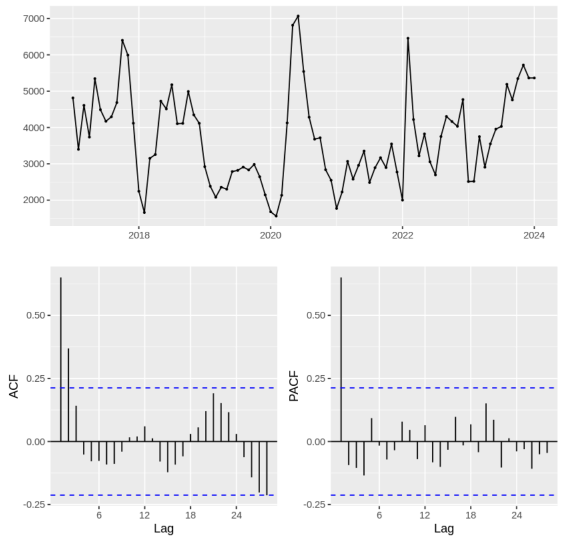
Imputamos para retirar valores extremos, marcados pelas eleições de 2022 e 2018

```
#usando a media
dados[12] <-  (dados[11] + dados[13])/2
dados[60] <-  (dados[59] + dados[61])/2
serie <-  ts(dados, start = c(2017,1), end = c(2024, 1), frequency = 12 )
```

FAC E FACP

Utilizaremos a FAC e FACP para identificar qual o modelo mais apropriado

```
ggtsdisplay(serie)
```



Identificamos AR(1)

```
modelo1 <- Arima(serie, order = c(1,0,0))
modelo1
```

```
Series: serie
ARIMA(1,0,0) with non-zero mean

Coefficients:
         ar1       mean
      0.6628  3765.0273
s.e.  0.0814   295.7528

sigma^2 = 901260:  log likelihood = -702.63
AIC=1411.26   AICc=1411.55   BIC=1418.58
```

Teste de significância dos coeficientes

```
coeftest(modelo1)
```

```
z test of coefficients:

            Estimate Std. Error z value  Pr(>|z|)
ar1       6.6277e-01 8.1401e-02   8.142 3.887e-16 ***
intercept 3.7650e+03 2.9575e+02  12.730 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Os coeficientes deram significativos

Testando com a série derivada

```
modelo1_derivada =  Arima(serie, order = c(1,1,0))
modelo1_derivada
coeftest(modelo1_derivada)
```

```
Series: serie
ARIMA(1,1,0)

Coefficients:
          ar1
      -0.0854
s.e.   0.1093

sigma^2 = 1062205:  log likelihood = -701.48
AIC=1406.95   AICc=1407.1   BIC=1411.82

z test of coefficients:

     Estimate Std. Error z value Pr(>|z|)
ar1 -0.085443   0.109294 -0.7818   0.4344
```

Não deu um bom ajuste

Sobrefixando

```
#AR(2)
modelo3 <- Arima(serie, order = c(2,0,0))
modelo3
coeftest(modelo3)
```

```
Series: serie
ARIMA(2,0,0) with non-zero mean

Coefficients:
         ar1      ar2       mean
      0.7282  -0.0992  3753.1383
s.e.  0.1085   0.1093   269.2495

sigma^2 = 903299:  log likelihood = -702.22
AIC=1412.44   AICc=1412.94   BIC=1422.21

z test of coefficients:

             Estimate  Std. Error z value  Pr(>|z|)
ar1          0.728171    0.108494  6.7116 1.925e-11 ***
ar2         -0.099188    0.109306 -0.9074    0.3642
intercept 3753.138291  269.249452 13.9393 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

O segundo termo não deu significativo

```
#ARMA(1,1)
modelo4 = Arima(serie, order = c(1,0,1))
modelo4
coeftest(modelo4)
```

```
Series: serie
ARIMA(1,0,1) with non-zero mean

Coefficients:
         ar1     ma1       mean
      0.5921  0.1279  3756.7299
s.e.  0.1293  0.1579   275.8591

sigma^2 = 904772:  log likelihood = -702.29
AIC=1412.57   AICc=1413.07   BIC=1422.34

z test of coefficients:

            Estimate Std. Error z value  Pr(>|z|)
ar1          0.59207    0.12929  4.5794 4.663e-06 ***
ma1          0.12790    0.15787  0.8101    0.4179
intercept 3756.72986  275.85911 13.6183 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

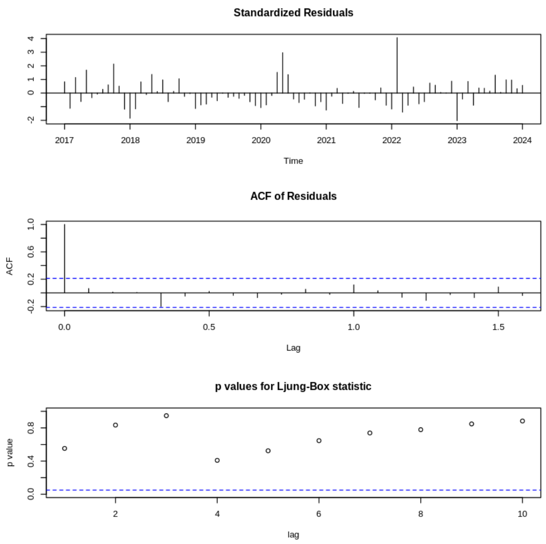
MA não significativo.

Assim, nosso modelo escolhido foi o AR(1)

$$\hat{Y}_t = 0.662\hat{Y}_{t-1}$$

Teste de LJung-Box

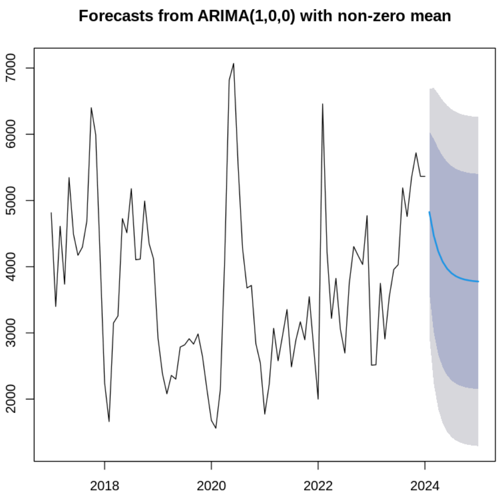
```
tsdiag(modelo1)
```



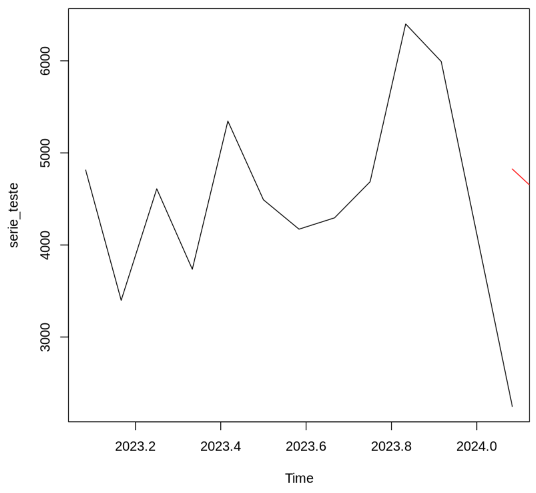
Todos os p-valores são maiores que 0.05 o que nos leva a não rejeitar a hipótese de erros serem um ruído branco a este nível de significância.

Fazendo previsões

```
serie_teste <-  ts(dados, start = c(2023,2), end = c(2024, 2), frequency = 12 )
Y_prev = forecast(modelo1, h = 12)
plot(Y_prev)
```




```
#plot previsão junto do periodo de validacao
plot(serie_teste, type='l')
lines(Y_prev$mean, col = 'red')
```




```
accuracy(serie_teste, as.numeric(Y_prev$mean))
```

A matrix: 1 × 5 of type dbl

	ME	RMSE	MAE	MPE	MAPE
Test set	-646.9835	1174.7	882.9337	-17.205	22.60567

22% de erro medio percentual (alto)