

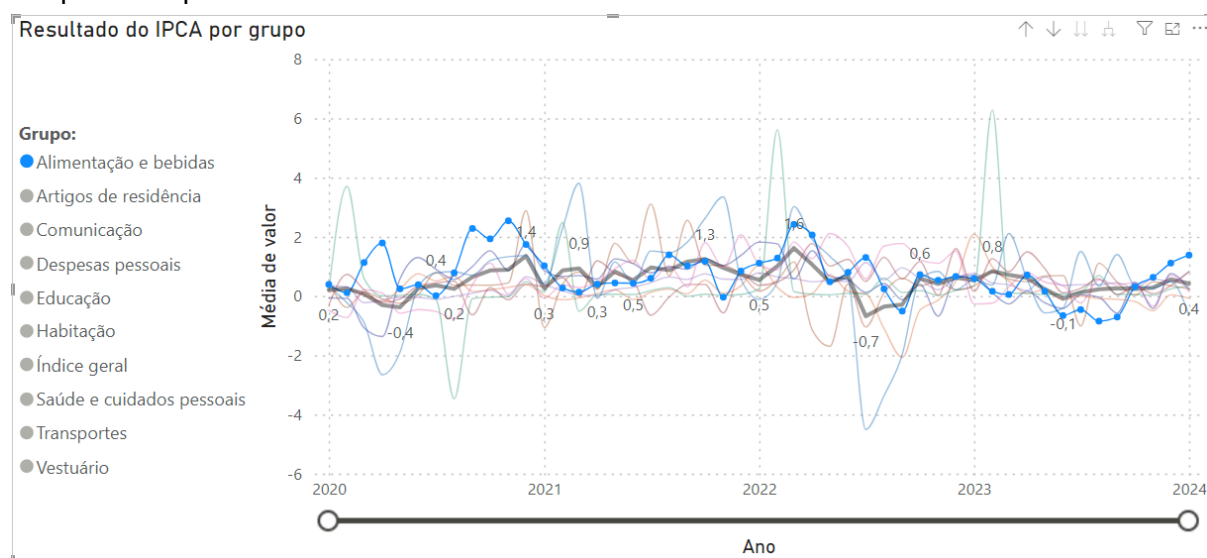
# Observações relevantes

Análise dos dados de variação mensal dos subítemos, itens, grupos, subgrupos e índice geral que compõem o IPCA

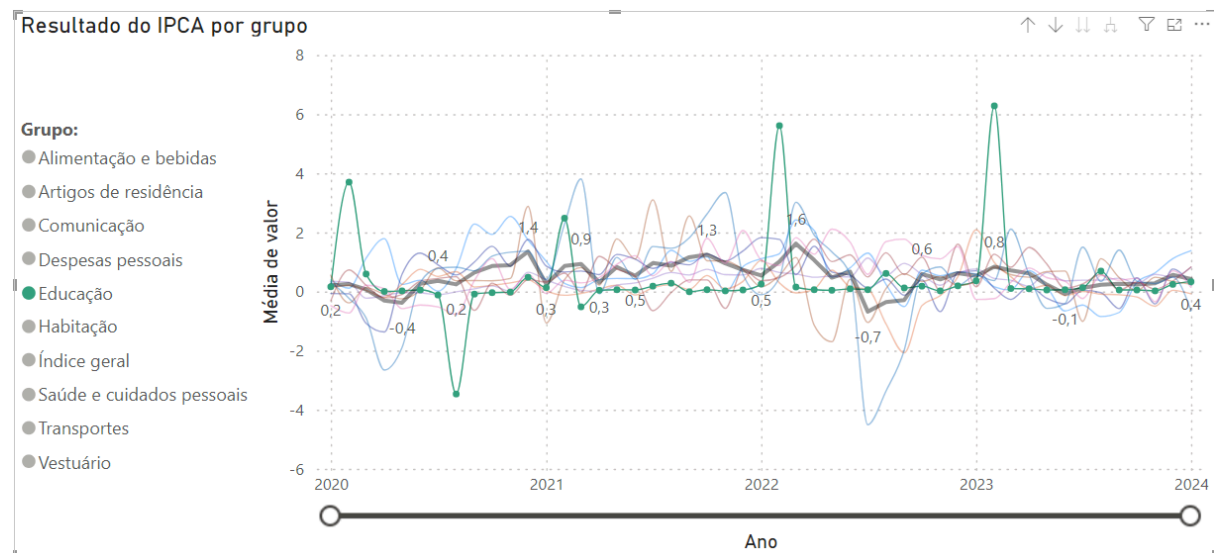
por Pedro H. R. Caramuru Crotti

## Análise exploratória

1. Das estatísticas descritivas para a variação mensal do índice oficial de inflação do País, o Índice de Preços ao Consumidor - Amplo (IPCA), podemos destacar que o **maior resultado desde 2006 para o índice geral foi registrado em março de 2022, (+1,62%)** e a **menor variação em julho do mesmo ano (-0,68%)**. Extremos tão próximos evidenciam a volatilidade econômica causada pela pandemia da covid-19, pela instabilidade geopolítica, e suas consequências, como a interrupção de cadeias de suprimentos e a quebra de expectativas produtivas.
2. Graficamente nota-se que certos grupos de consumo possuem uma **variação de preços mais sazonal que outros**.
  - a. Um desses exemplos é o grupo de Alimentação, cujo comportamento acompanha variações no ciclo produtivo agrícola e apresenta altas de preço nos períodos próximos ao fim de ano.

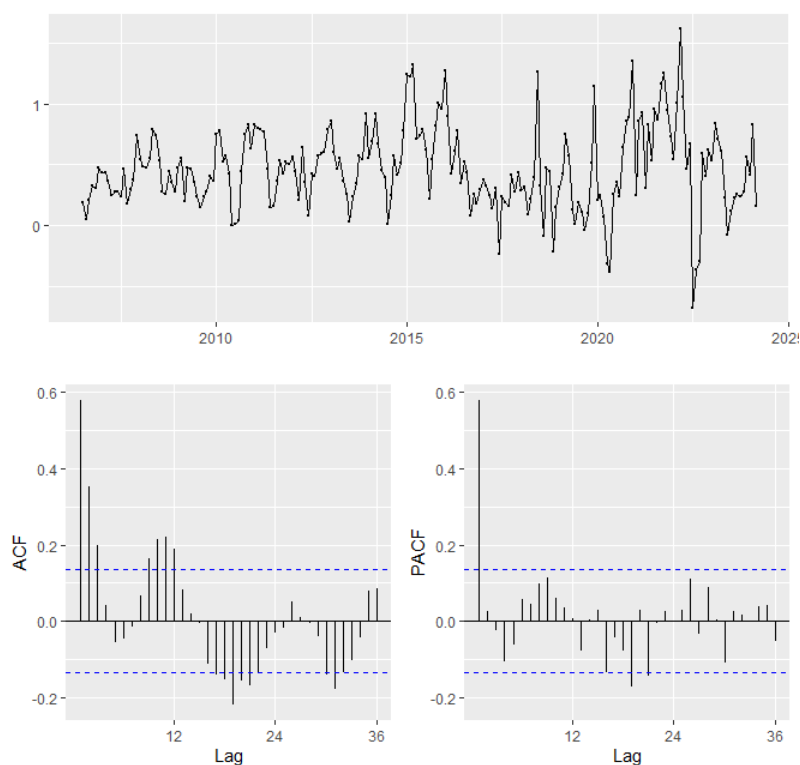


- b. Outro exemplo, ainda mais evidente, é o **comportamento do grupo de Educação, que registra altas explosivas todo mês de fevereiro**, quando são realizados os reajustes de mensalidade. O único mês em que houve contribuição negativa expressiva do grupo de Educação foi no ano de 2020, quando o isolamento social fez com que diversas famílias optassem por retirar seus filhos ou deixar de frequentar as escolas.



Reprodução: PowerBI. Fonte IBGE/ API Sidra. Elaboração: Pedro Caramuru.

3. A média da variação mensal da inflação no período foi de 0,4624% e a mediana, de 0,4400%, evidenciando uma concentração dos resultados nos dois quartis inferiores da distribuição. Nota-se também um aumento da volatilidade (e da variância) da distribuição do índice geral mais acentuada a partir do ano de 2014. Esse resultado é respaldado pela realidade, uma vez que a partir deste ano o Brasil tem atravessado sucessivas crises econômicas.



Reprodução: RStudio. Fonte IBGE/ API Sidra. Elaboração: Pedro Caramuru.

- a. Da série temporal, nota-se que a função de autocorrelação nos aponta para uma grande correlação positiva entre os instantes  $t$  e  $t - 1$ . E uma correlação negativa entre os instantes  $t$  e os múltiplos de  $t - 6$ , em linha

com a literatura econômica de defasagem contracíclica da política monetária.

## Inferência estatística

4. A fim de se aprofundar nos modelos autorregressivos, é preciso garantir que a série não tenha raízes unitárias. Uma vez que se trata da variação do índice de preços (proxy de um resultado de primeira diferença), espera-se estacionariedade da série temporal. E, de fato, pelos testes Augmented Dickey-Fuller, temos que:

```
> adf.test(ts_IPCA, nlag = 4, output = TRUE) #redundância
Augmented Dickey-Fuller Test
alternative: stationary

Type 1: no drift no trend
      lag   ADF p.value
[1,]    0  -4.06    0.01
[2,]    1  -3.27    0.01
[3,]    2  -2.86    0.01
[4,]    3  -2.75    0.01
Type 2: with drift no trend
      lag   ADF p.value
[1,]    0  -7.44    0.01
[2,]    1  -6.50    0.01
[3,]    2  -6.03    0.01
[4,]    3  -6.16    0.01
Type 3: with drift and trend
      lag   ADF p.value
[1,]    0  -7.42    0.01
[2,]    1  -6.48    0.01
[3,]    2  -6.01    0.01
[4,]    3  -6.14    0.01
----
Note: in fact, p.value = 0.01 means p.value <= 0.01

> summary(adf_IPCA0)

#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-1.26157 -0.14024 -0.01253  0.16202  0.84427

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.933e-01  4.952e-02   3.903 0.000129 ***
z.lag.1      -4.118e-01  6.431e-02  -6.403 1.03e-09 ***
tt           -7.704e-06  3.224e-04  -0.024 0.980961
z.diff.lag   -2.168e-02  7.075e-02  -0.306 0.759595
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2791 on 204 degrees of freedom
Multiple R-squared:  0.2101,    Adjusted R-squared:  0.1984
F-statistic: 18.08 on 3 and 204 DF,  p-value: 1.914e-10

value of test-statistic is: -6.4032 13.6778 20.5166

Critical values for test statistics:
      1pct  5pct 10pct
tau3  -3.99 -3.43 -3.13
phi2   6.22  4.75  4.07
phi3   8.43  6.49  5.47
```

Reprodução: RStudio. Fonte IBGE/ API Sidra. Elaboração: Pedro Caramuru.

Logo, é possível rejeitar a hipótese nula em favor da estacionariedade da série temporal.

5. Analisando os critérios de informação dos processos autorregressivos quanto a dispersão das funções de autocorrelação e de autocorrelação parcial, temos que o modelo que melhor explica o desempenho da série temporal de variação mensal da inflação é ARMA(1,0), com coeficientes estatisticamente significantes.

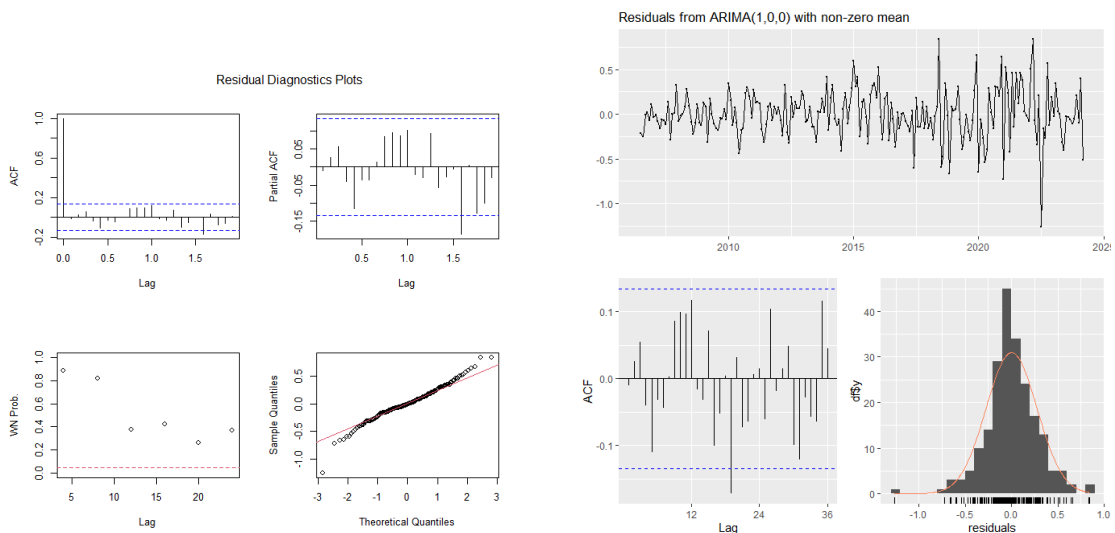
```
> coeftest(ar1)

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1      0.581723   0.055593  10.464 < 2.2e-16 ***
intercept 0.458687   0.044642  10.275 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Reprodução: RStudio. Fonte IBGE/ API Sidra. Elaboração: Pedro Caramuru.

6. Cabe destacar que apesar de o modelo ser a melhor aproximação do comportamento da variação mensal dos índices de inflação, nota-se que os resíduos ainda apresentam certo desvio do comportamento de uma distribuição normal.



Reprodução: RStudio. Fonte IBGE/ API Sidra. Elaboração: Pedro Caramuru.

7. Por último, a retirada de outliers pouco contribuiu para a melhora dos modelos de regressão. O único ponto apontado como outliers é o ponto de máximo, observado em março/2022 e o valor é estatisticamente significativo a um nível de 99,9%.

```
Call:
lm(formula = ts_IPCA ~ dummy_mar22)

Residuals:
    Min       1Q   Median       3Q      Max
-1.13693 -0.20693 -0.01693  0.17307  0.89307

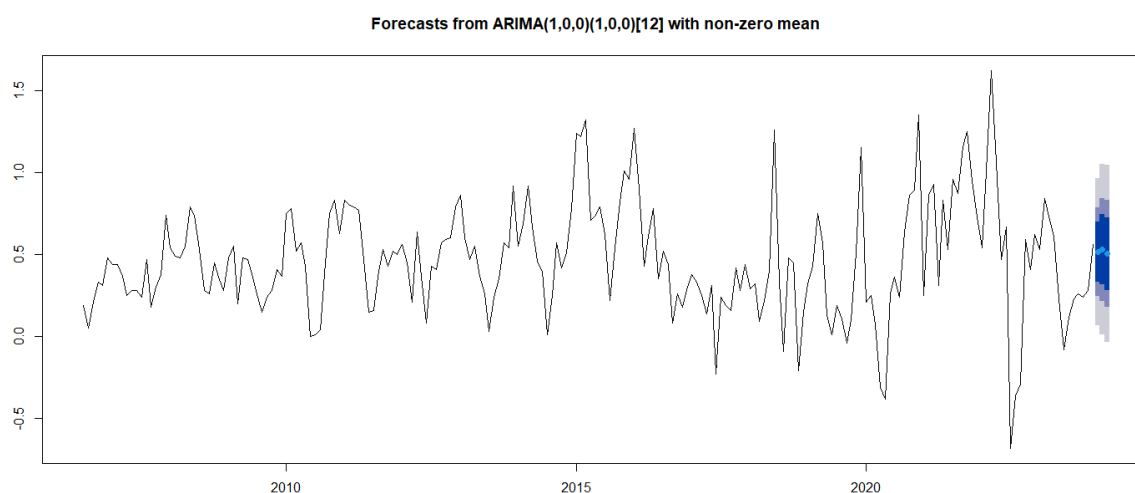
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.45693    0.02263   20.196 < 2e-16 ***
dummy_mar22  1.16307    0.33021    3.522 0.000524 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3294 on 211 degrees of freedom
Multiple R-squared:  0.05553,    Adjusted R-squared:  0.05106
F-statistic: 12.41 on 1 and 211 DF, p-value: 0.0005244
```

## Forecasting

8. Por último, testou-se a habilidade do modelo de se prever os próximos resultados. Para isso, a série temporal foi restrita ao último mês de 2023 e foi demandado do modelo que fossem projetados os meses de janeiro, fevereiro e março de 2024. Os dados estão resumidos na tabela e imagem a seguir.

	realizado	Point Fcst.	50% level	68% level	90% level
jan/24	<b>0,42%</b>	0,51%	0,33% ↔ 0,70%	0,24% ↔ 0,78%	0,06% ↔ 0,96%
fev/24	<b>0,83%</b>	0,52%	0,32% ↔ 0,74%	0,21% ↔ 0,84%	0,01% ↔ 1,04%
mar/24	<b>0,16%</b>	0,50%	0,28% ↔ 0,72%	0,17% ↔ 0,82%	-0,03% ↔ 1,04%



Reprodução: RStudio. Fonte IBGE/ API Sidra. Elaboração: Pedro Caramuru.

9. Nota-se que quanto mais distante o resultado, maior o bracket de projeção no qual o resultado realizado se encontrou. Para próximas iterações, espera-se melhores resultados de uma análise que leve em consideração os comportamentos individuais e sazonais de cada um dos grupos - ou subgrupos - que compõem a amostra do índice oficial de inflação.