

Desagregação de Séries Temporais

Uma abordagem introdutória

Rodrigo Hermont Ozon



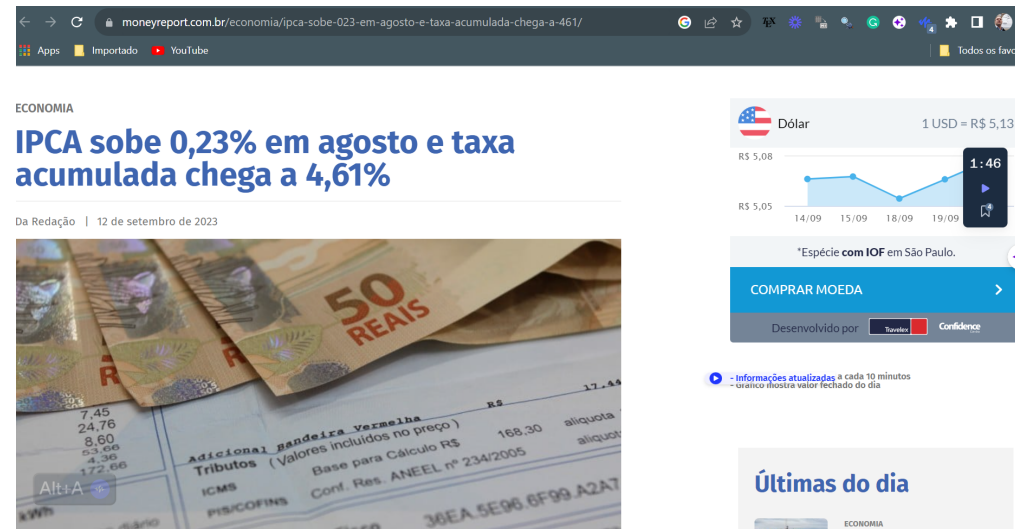
Conteúdo

O uso de desagregação de séries temporais com o pacote tempdisagg em modelos econométricos

- Quando podemos e usamos agregação de dados de séries temporais ?
- Distinções entre dividir e desagregar dados de séries temporais
- Como e por quê desagregar uma série temporal ?
- Quais critérios de desagregação devo empregar ?
- Alguns exemplos de aplicação
- Esclarecimento de dúvidas
- Agradecimentos

Agregando dados de séries temporais

Alguns exemplos de séries temporais usualmente agregadas...



Fonte: moneyreport.com.br, 12/09/23

Aqui os economistas fazem um produtório pra denotar a taxa acumulada até o período n :

$$i_{\text{acumulada}} = \prod_{t=1}^n (1 + i) \cdot (1 + i)$$

Agregando dados de séries temporais



Fonte: [Agencia de Noticias IBGE](#)

Aqui geralmente se faz uma soma da quantidade produzida em kgs ou toneladas ao longo de um determinado período de tempo:

$$\widehat{\text{Safra}}_t = \sum_{t=1}^n \text{volume}_t$$

Dividir x Desagregar dados de séries temporais

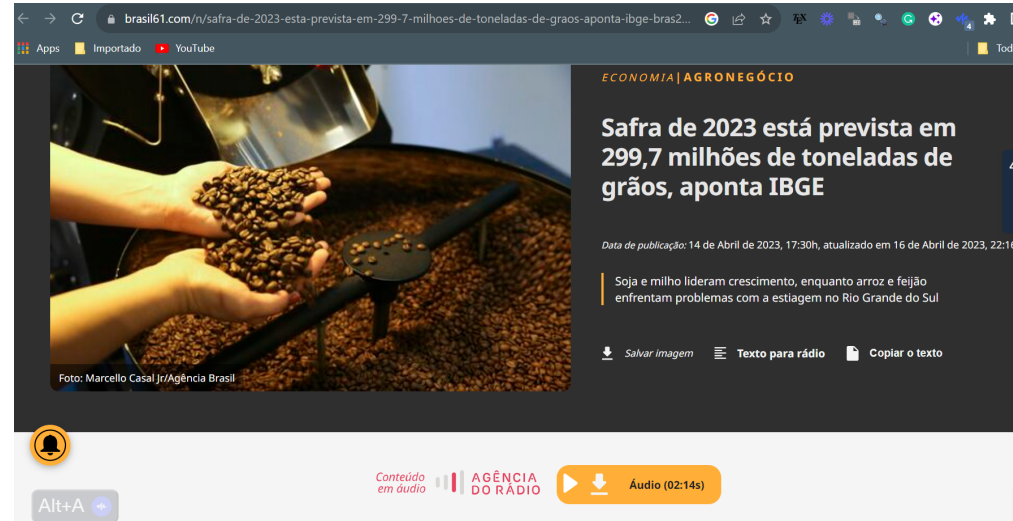
Imagine que você necessita calcular as parcelas de um financiamento mensalmente. Você sabe de antemão que a taxa de juros básica (referência) da economia brasileira é a taxa SELIC que é unicamente de contabilização anual...



Fonte: monitormercado.com.br

Aqui a taxa SELIC mensal é dada por $SELIC_{mês} = (1 + i)^{(q/t)} = (1 + 13,25\%)^{(1/12)} \Rightarrow 1,042290887\% a. m.$

Dividir x Desagregar dados de séries temporais



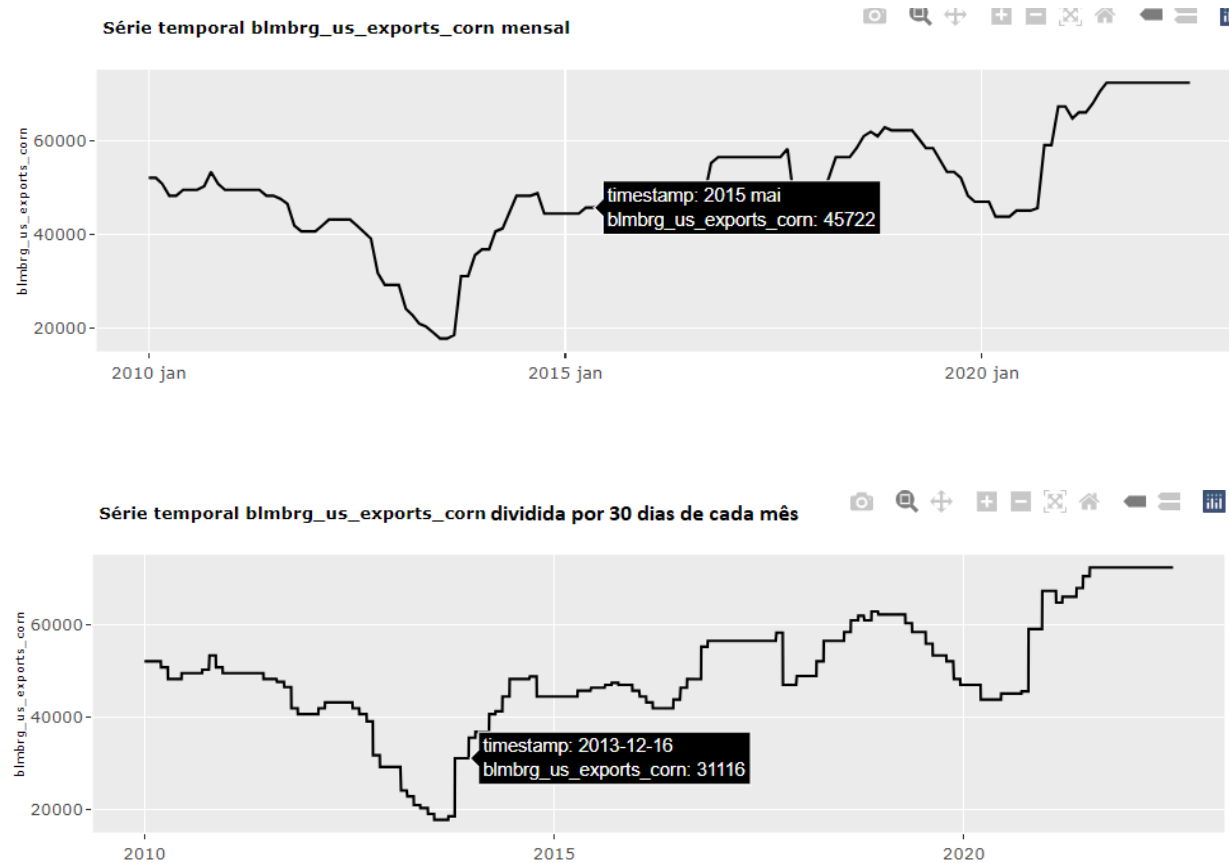
Fonte: [Brasil61](https://brasil61.com)

- Aqui simplesmente dividiríamos os valores da produção prevista pro ano de 2023, se quiséssemos dados mensais por 12...

Dividir x Desagregar dados de séries temporais

- Mas, qual seria o problema em se fazer essas operações simples de dividir, somar ou multiplicar ?
- E se quiséssemos construir um modelo econométrico de séries temporais, digamos pra prever os preços diários do milho na CBOT em função de alguns dados de demanda e de oferta, por exemplo ?

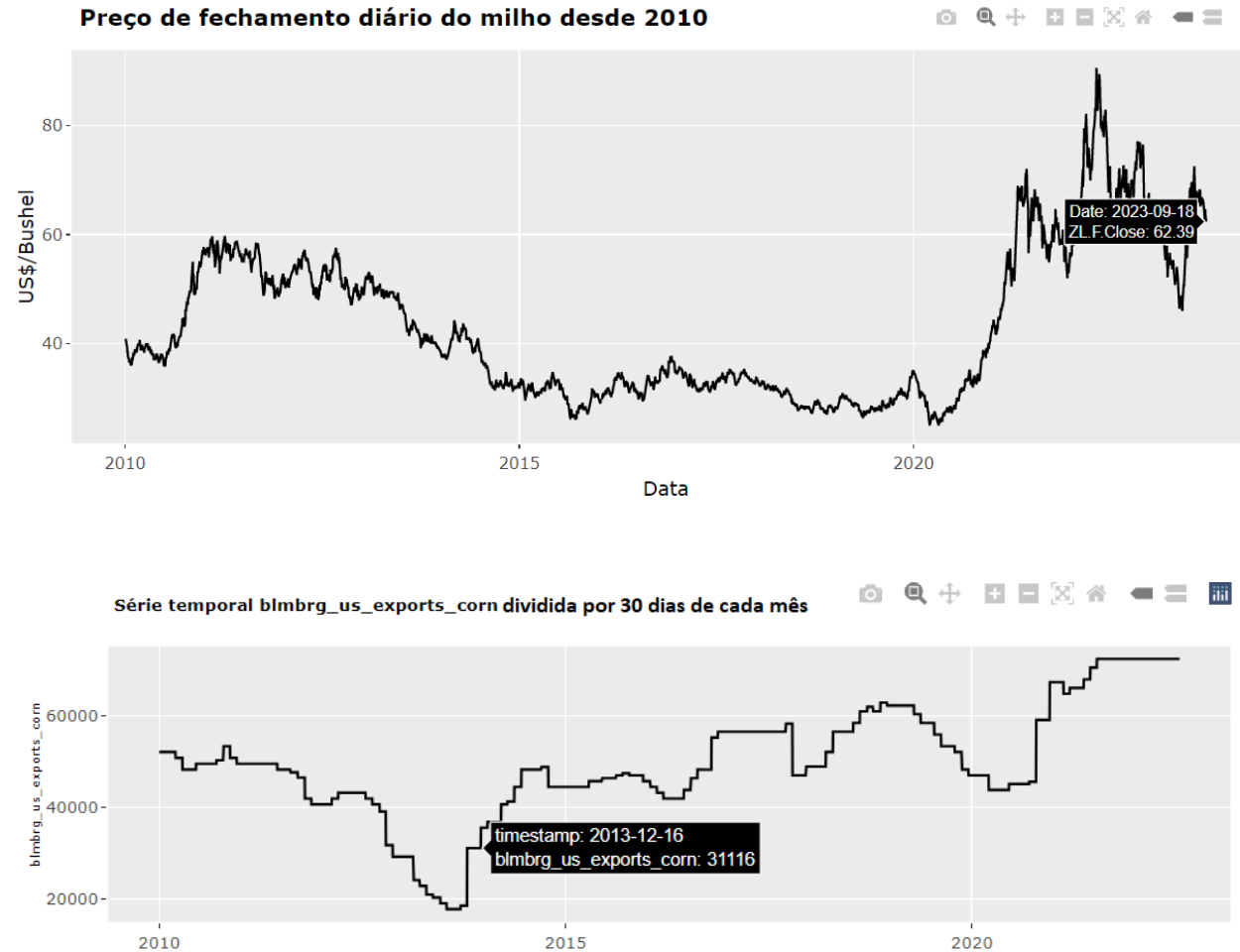
Dividir x Desagregar dados de séries temporais



Fonte: Bloomberg a partir de dados de FAS/USDA, dados de exportações mensais e “diarizadas” de milho norte-americano

Dividir x Desagregar dados de séries temporais

Será que terá um bom fit estatístico ?



Fonte: Bloomberg a partir de dados de FAS/USDA, dados de preços de fechamento diário de milho norte-americano x exportações americanas

Dividir x Desagregar dados de séries temporais

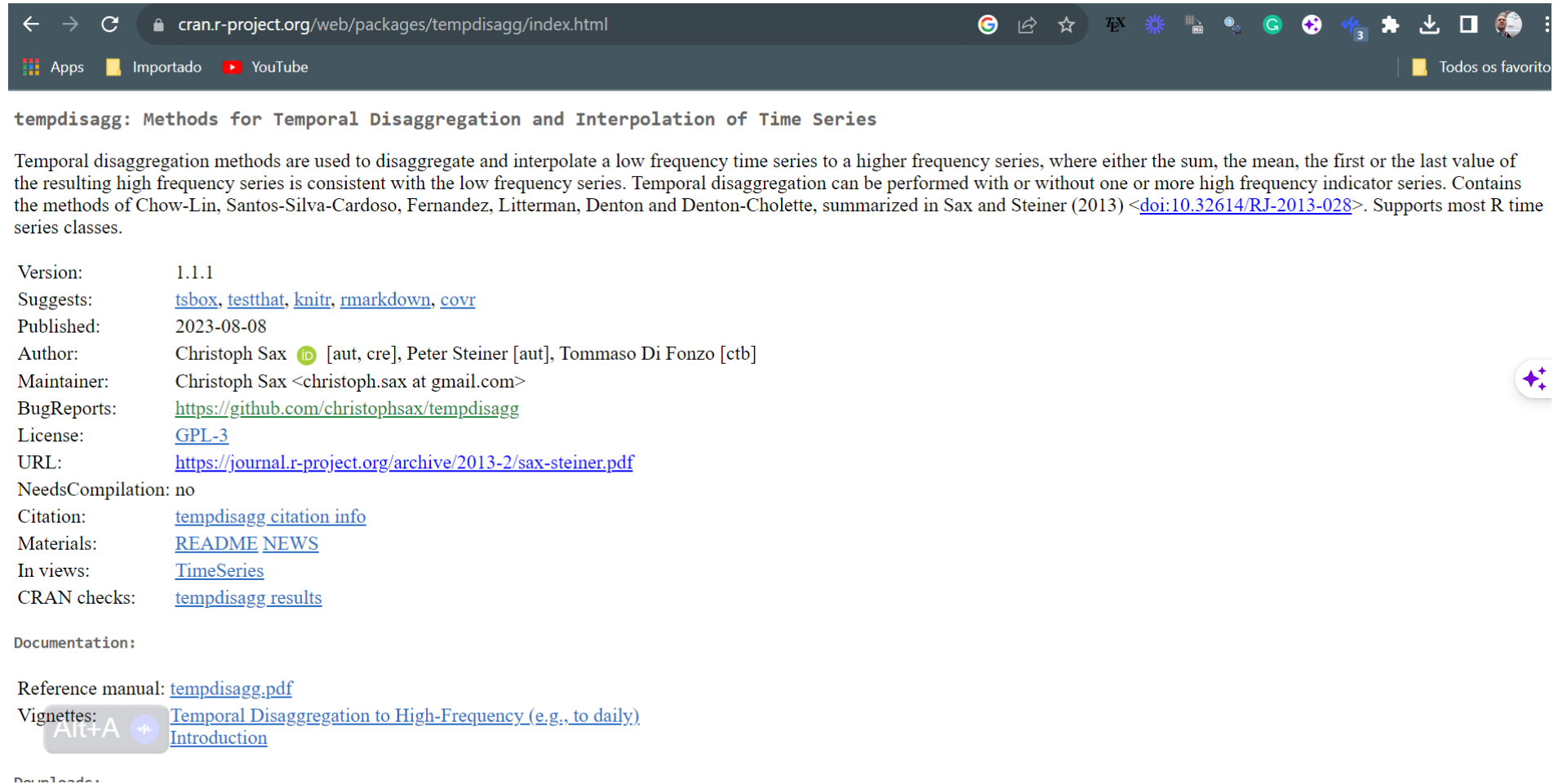
"Uma variável aleatória é nada mais que uma função que mapeia os resultados de um processo estocástico (aleatório) em valores numéricos."

(Zibetti, 2023)

- Relações determinísticas x estocásticas (vide Gujarati & Porter, 2011 p. 28)
- O sentido “econômico” das relações com os dados disponíveis precisa existir !
- Um bom modelo sempre será aquele que mais se adequa à realidade de acordo com as possibilidades ...
- Mas então, como construir um modelo que seja “mais realista” com os dados de exportação, produção, taxas e etc. ?
- Como preservar a verdadeira variabilidade nas exportações dia-a-dia e construí-la de maneira mais realista ?
- Afinal, quem exporta, exporta todo dia, quem produz, produz todo dia e não somente mensalmente p. ex. ...

Desagregação de séries temporais

O glorioso trabalho e contribuição de [Sax & Steiner \(2013\)](#) !!!



The screenshot shows a web browser window displaying the CRAN page for the `tempdisagg` package. The browser's address bar shows the URL `cran.r-project.org/web/packages/tempdisagg/index.html`. The page title is `tempdisagg: Methods for Temporal Disaggregation and Interpolation of Time Series`. The main text describes the package's purpose: "Temporal disaggregation methods are used to disaggregate and interpolate a low frequency time series to a higher frequency series, where either the sum, the mean, the first or the last value of the resulting high frequency series is consistent with the low frequency series. Temporal disaggregation can be performed with or without one or more high frequency indicator series. Contains the methods of Chow-Lin, Santos-Silva-Cardoso, Fernandez, Litterman, Denton and Denton-Cholette, summarized in Sax and Steiner (2013) <[doi:10.32614/RJ-2013-028](https://doi.org/10.32614/RJ-2013-028)>. Supports most R time series classes."

Metadata fields include:

- Version: 1.1.1
- Suggests: [tsbox](#), [testthat](#), [knitr](#), [rmarkdown](#), [covr](#)
- Published: 2023-08-08
- Author: Christoph Sax [aut, cre], Peter Steiner [aut], Tommaso Di Fonzo [ctb]
- Maintainer: Christoph Sax <christoph.sax@gmail.com>
- BugReports: <https://github.com/christophsax/tempdisagg>
- License: [GPL-3](#)
- URL: <https://journal.r-project.org/archive/2013-2/sax-steiner.pdf>
- NeedsCompilation: no
- Citation: [tempdisagg_citation_info](#)
- Materials: [README](#) [NEWS](#)
- In views: [TimeSeries](#)
- CRAN checks: [tempdisagg_results](#)

Documentation:

- Reference manual: [tempdisagg.pdf](#)
- Vignettes: [Temporal Disaggregation to High-Frequency \(e.g., to daily\)](#), [Introduction](#)

Download:

Fonte: [CRAN, 2023](#)

Desagregação de séries temporais

- Os autores uparam alguns “tutoriais de uso” [Vignettes – “Disaggregating annual data to quarterly”](#) para usuários do R;
- Ps* No Python existe [uma iniciativa de replicação](#) da tempdisagg, mas me pareceu inacabada ([veja aqui o repo no GitHub](#) da biblioteca [timedisagg](#))
- Como que eu utilizei e adaptei o tratamento da desagregação para o meu caso ?

Desagregação blmbrg_us_exports_corn

E partimos para desagregá-la corretamente para uma série diária (vide Sax & Stein, 2013)

Paralelização

Pacotes

Master Table

Tratamento blmbrg_us_exports_corn

Desagregação

blmbrg_us_exports_corn

```
# O pacote exige que se crie um dataframe contendo a coluna com as datas contra a serie sozinha

blmbrg_us_exports_corn_df <- monthly_dataset %>%
  select(

    timestamp,

    blmbrg_us_exports_corn

  )
```

```
modelo <- predict( td(blmbrg_us_exports_corn_df~ 1,
                     to = "daily",
                     conversion = "sum", # type of conversion: "sum", "mean" (or: "average"), "first" or "last".
                     method = "fast" ,
                     truncated.rho = 0 # Lower bound for the autoregressive parameter  $\rho$ . If set to 0 (default),
no negative values are allowed. If set to -1, truncation is disabled.
                     )) # %>% # Transforma mês pra dia
# mutate(
#   br_domestic_consumption_diaria = value + 0.9*(noise( value, "white", 1 ) )
# )

# method of temporal disaggregation: -----
# "chow-lin-maxLog",
# "chow-lin-minrss-ecotrim",
# "chow-lin-minrss-quilis",
# "chow-lin-fixed",
# "dynamic-maxLog" (experimental),
# "dynamic-minrss" (experimental),
# "dynamic-fixed" (experimental),
# "fernandez",
# "litterman-maxLog",
# "litterman-minrss",
# "litterman-fixed",
# "denton-cholette",
# "denton",
# "fast",
# "uniform" or "ols".
# -----

modelo <- modelo %>%
  rename(
    timestamp = "time",
    blmbrg_us_exports_corn = "value"
  )

modelo <- modelo %>%
  mutate( # Como o modelo faz a variavel <0 insiro esse Limitador
    blmbrg_us_exports_corn = case_when(
      blmbrg_us_exports_corn < 0 ~ 0,
      blmbrg_us_exports_corn >=0 ~ blmbrg_us_exports_corn
    )
  )

#modelo
```

Critérios de Desagregação

- Mas, como sei qual critério devo utilizar para a minha série temporal, em específico ?
- Afinal, não posso desagregar uma série temporal de um dado de produção do mesmo modo que desagrego uma taxa, por exemplo ...
- Como desagrego, pra diário, uma série temporal de preços com disponibilidade de publicação unicamente semanal numa fonte oficial ?
- Como desagrego, pra construir uma série diária, um dado que é um número-índice, [como o índice de anomalia do El Niño, por exemplo ?](#)

Critérios de Desagregação

Paralelização

Pacotes

Master Table

Tratamento blmbrg_us_exports_corn

Desagregação

blmbrg_us_exports_corn

```
blmbrg_us_exports_corn

)

modelo <- predict( td(blmbrg_us_exports_corn_df ~ 1,
  to = "daily",
  conversion = "sum", # type of conversion: "sum", "mean" (or: "average"), "first" or "last".
  method = "fast",
  truncated.rho = 0 # lower bound for the autoregressive parameter  $\rho$ . If set to 0 (default),
no negative values are allowed. If set to -1, truncation is disabled
))
# mut
# b
# )

# method of temporal disaggregation: -----
# "chow-lin-maxlog",
# "chow-lin-minrss-ecotrim",
# "chow-lin-minrss-quilis",
# "chow-lin-fixed",
# "dynamic-maxlog" (experimental),
# "dynamic-minrss" (experimental),
# "dynamic-fixed" (experimental),
# "fernandez",
# "litterman-maxlog",
# "litterman-minrss",
# "litterman-fixed",
# "denton-cholette",
# "denton",
# "fast",
# "uniform" or "ols".
# -----

modelo <- modelo %>%
  rename(
    timestamp = "time",
    blmbrg_us_exports_corn = "value"
  )

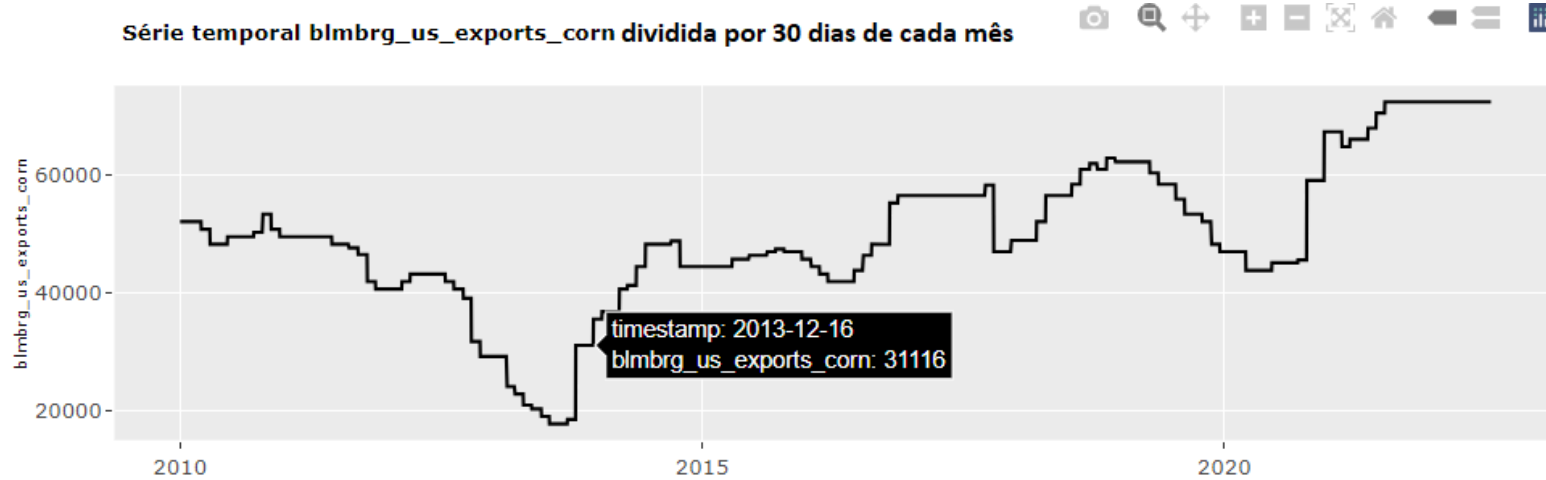
modelo <- modelo %>%
```

**Critérios possíveis de empregar com a
função de desagregação...**

Critérios de Desagregação

- Como no meu caso, os dados que eu possuía se referiam as quantidades de milho exportadas mensalmente, eu utilizei o critério de desagregação pela soma.
- Afinal os órgãos de estatística oficiais dos governos contabilizam uma soma do volume total exportado em toneladas mensalmente e as divulgam dessa maneira...
- Mas se a série for de preços, posso utilizar o critério de desagregação por “similaridade” como o last ou o first por exemplo;
- Para variáveis que seriam taxas acumuladas, posso utilizar o critério de desagregação pelas médias (geométricas) por exemplo...
- Para números-índices como o El Niño por exemplo, posso utilizar de técnicas mais elaboradas de desagregação combinando p. ex. similaridade com os desvios em relação as oscilações médias... Ou criar uma feature utilizando um simulador de ruído branco...

Um exemplo de séries desagregadas



Mostrando mais exemplos...

- Desagregação da série temporal da Produção de Milho no Brasil de mês pra dia;
- Desagregação da série do El Niño
- Desagregação e preenchimento de missing data com técnica de backcasting e bootstrapping
- Desagregação de dados temporais de área colhida de milho na Ucrânia
- Desagregação a séries de preços do óleo de palma negociado na Malásia/Mundo

Referências

Angrist, Joshua D., and Jörn-Steffen Pischke. *Mastering 'Metrics: The Path from Cause to Effect*. Princeton University Press, 2014.

Angrist, Joshua D., and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 2009.

FAO. *Price volatility in food and agricultural markets: Policy Responses*. 2011. Disponível em: <http://www.oecd.org/tad/agricultural-trade/48152638.pdf>. Acessado em Abr 2022

Pearl, Judea, Madelyn Glymour, and Nicholas P. Jewell. *Causal Inference in Statistics: A Primer*. Wiley, 2016.

Granger, C. W. J. *Empirical modeling in economics*. Reino Unido: Cambridge University Press, 1999. p. 58.

Gujarati, D.,N. *Basic Econometrics*, fourth edition, McGraw-Hill/Irwin, 2011.

Hyndman, R.J., & Athanasopoulos, G. (2021) *Forecasting: principles and practice*, 3rd edition, OTexts: Melbourne, Australia. [OTexts.com/fpp3](https://otexts.com/fpp3). Accessed on <may 2023>.

Maddala, G.,S. *Introdução à Econometria*, terceira edição, LTC, Rio de Janeiro, 2001.

Pedace, R. *Econometrics for Dummies*. John Wiley & Sons, 2013.

Sax, C. & Steiner, P. *Temporal Disaggregation of Time Series*. In The R Journal Vol. 5/2, December 2013 ISSN 2073-4859.

Obrigado!

Rodrigo Hermont Ozon

⇒ A todo o comitê organizador do 4º R Day e do [1th SouthStat Meeting](#)

⇒ Ao professor [Matheus Dal Molin Ribeiro](#) da UTFPR;

⇒ Agradecimentos aos pesquisadores do [PPGEPS/PUCPR](#), aos ouvintes e em especial ao meu orientador, prof. Dr. Gilberto Reynoso Meza

- Meu perfil no [Google Scholar](#)
- Meu [CV Lattes](#)
- Meu [site com posts, tutoriais e artigos](#)
- Meu [perfil no LinkedIn](#)

