

O GRANDE TÍTULO

Alisson Rosa e Vítor Pereira

Resumo

One Piece > Naruto

Sumário

1	Introdução	1
2	Informações	1
3	Testes	5
4	Modelagem	6
4.1	Médias Móveis Simples	7
4.2	Média Móvel Simples com Previsão	8
4.3	Regressão LOESS	9
	Referências	9

1 Introdução

Naruto é uma animação japonesa (anime) que adapta a série de mangá escrita e ilustrada por Masashi Kishimoto, que conta a história de Naruto Uzumaki, um jovem ninja que constantemente procura por reconhecimento e sonha em se tornar Hokage, o ninja líder de sua vila.

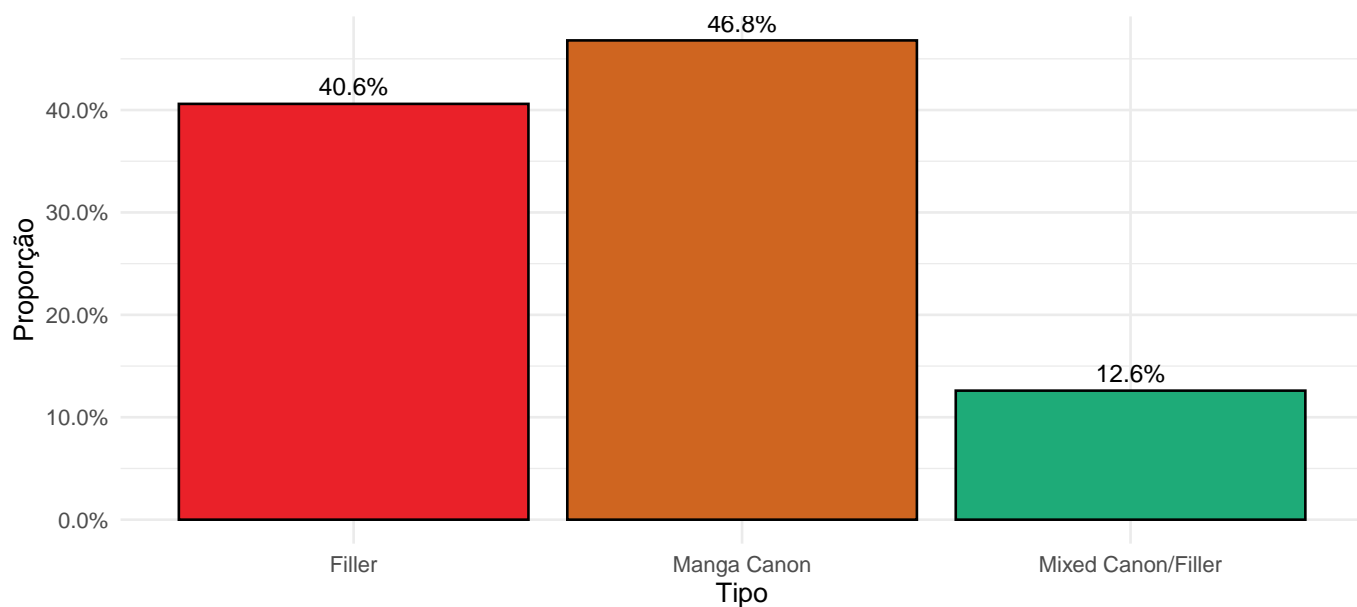
A história é dividida em duas partes, a primeira parte se passa nos anos da pré-adolescência de Naruto (clássico), e a segunda parte se passa em sua adolescência (Shippuden). Nesse trabalho trataremos sobre o Naruto Shippuden, desenvolvendo-o em uma perspectiva de séries temporais, vamos adotar como variável de interesse a avaliação dos episódios, assim veremos pontos importantes da saga ao longo do tempo.

O Banco de dados utilizado é totalmente original, foi criado fazendo *Web scraping* de dois sites diferentes, para aqueles interessados o banco foi disponibilizado aqui.

2 Informações

Como o anime é uma adaptação, existem episódios fiéis ao mangá (Manga Canon) e episódios originais do próprio anime, em outras palavras não seguem o material original, esses episódios são chamados de Fillers, existem também episódios que seguem a trama do mangá mas além disso possuem elementos novos, esses chamados de Mixed Canon.

Vejamos pelo gráfico a seguir a porcentagem do tipo dos episódios



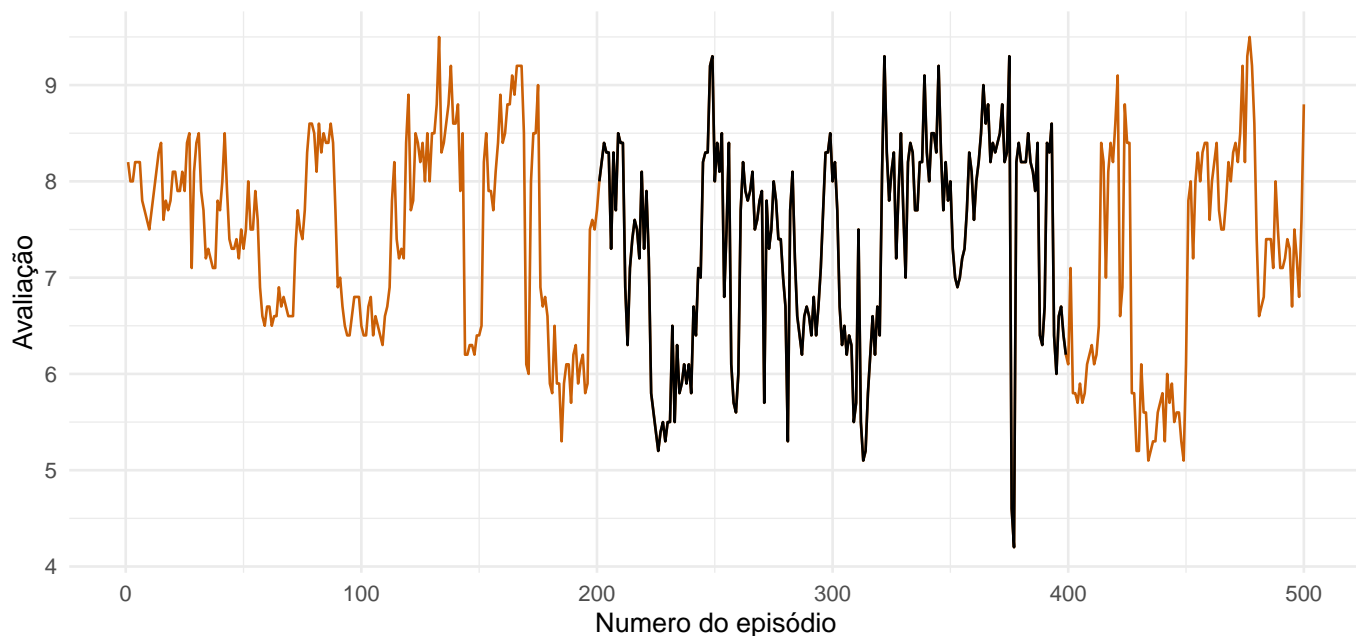
Fonte: Elaborado pelos autores

É um fato bastante curioso, Naruto Shippuden possui uma quantidade altíssima de episódios fillers, aproximadamente 41%, note pela tabela que são somente 31 episódios a mais canônicos.

Tabela 1: Valores Absolutos do Tipo de Episódio

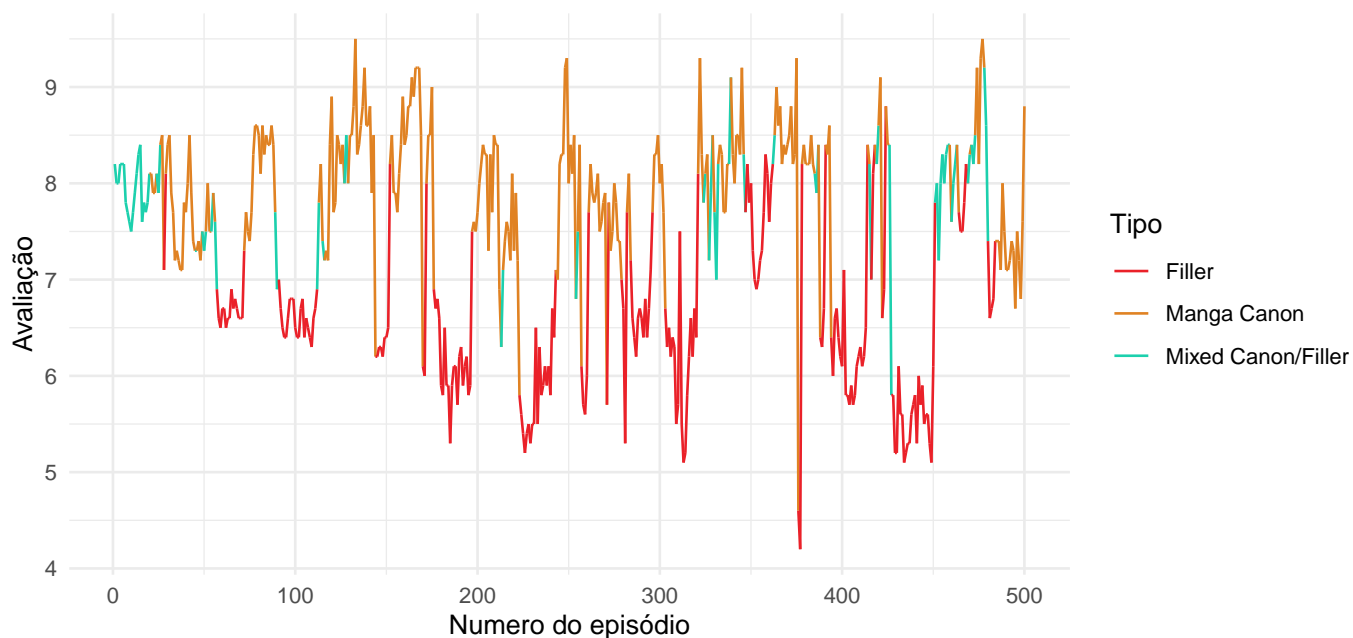
Tipo	Quantidade
Filler	203
Manga Canon	234
Mixed Canon/Filler	63

Em termos de avaliação ao longo do tempo, podemos utilizar como índice ordinal o próprio número do episódio já que que é uma função injetora no tempo.



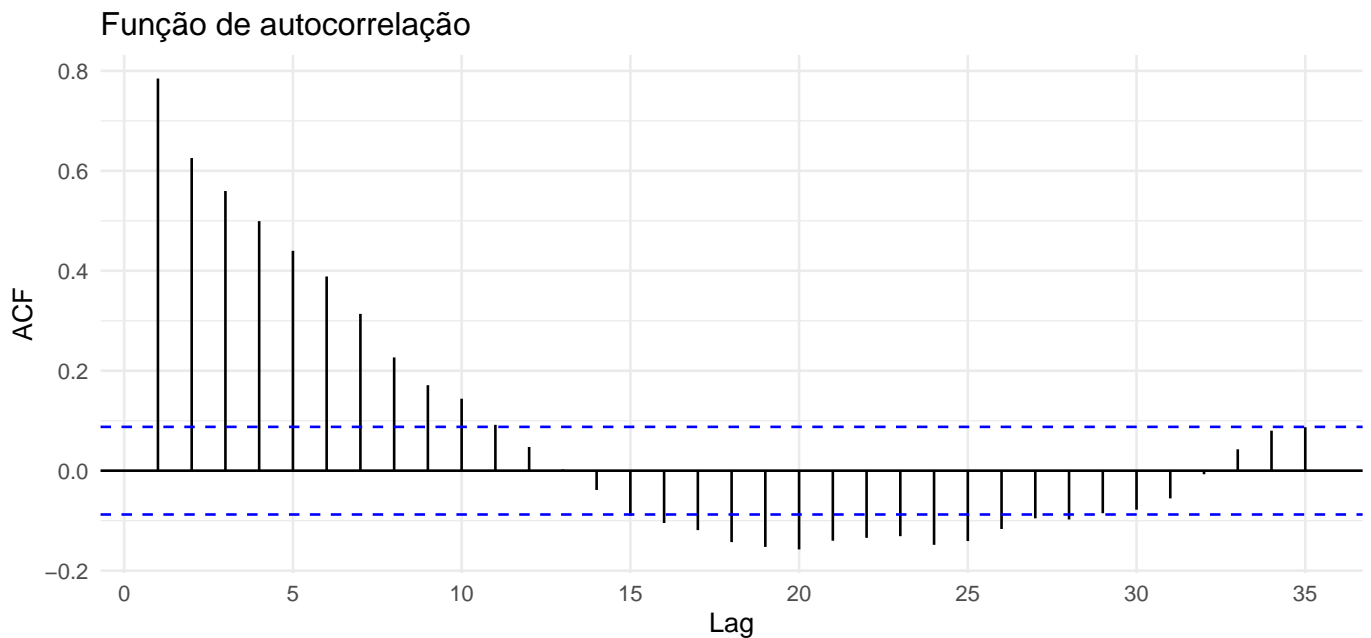
Há bastante coisa a se notar nesse gráfico, primeiramente pelos episódios 200 e 400, existe um indício de alteração na média das variáveis aleatórias, portanto furando o pressuposto de média constante ao longo do tempo, perto dos episódios finais da saga nota-se também inúmeras quedas bruscas na avaliação sem um contra-peso de avaliações com notas altas, outro indicio de não estacionariedade.

É importante tentarmos entender um pouco sobre esses episódios com notas baixas, assim vamos ver esse gráfico pelo tipo de episódio.

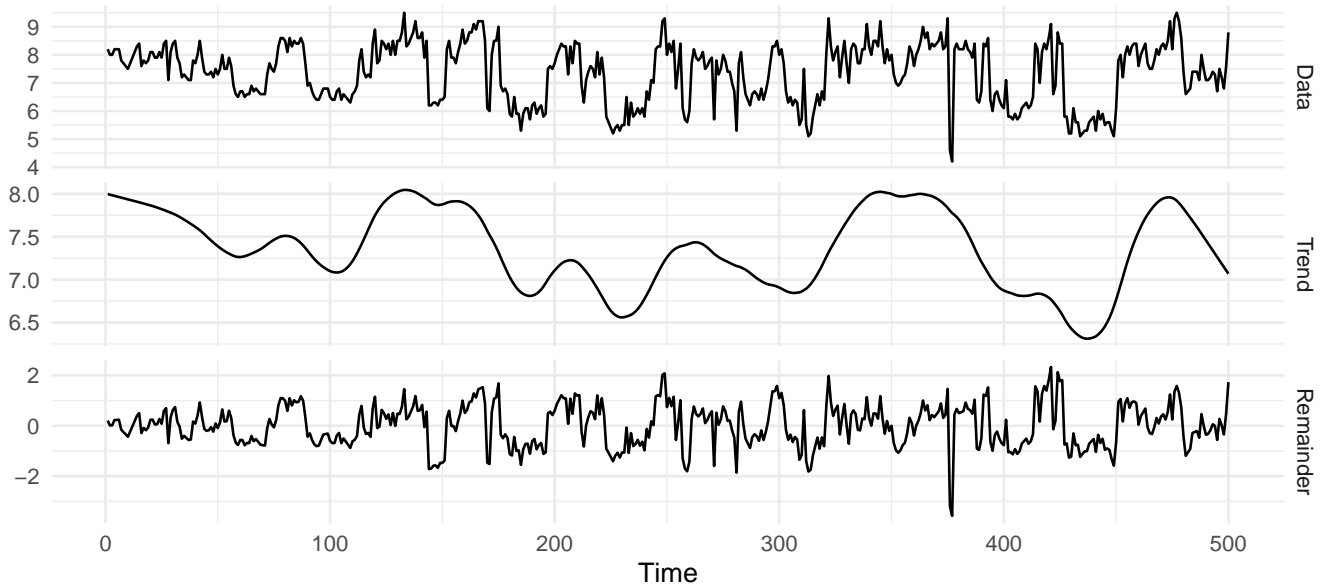


Assim fica fácil ver que os episódios com menor avaliação são em sua maior parte fillers, toda queda brusca de avaliação tem um episódio filler envolvido.

É necessário também avaliar a correlação nas avaliações:



Uma abordagem importante também é decompor a série temporal, apriori supõe-se a existência de três elementos, a saber: Tendência, sazonalidade e resíduo. Todos esses serão explorados cuidadosamente nas seções posteriores, assim vamos decompor a série para ter um vislumbre de tais elementos:



```
##
## Shapiro-Wilk normality test
##
## data:  data.frame(decompose)$Remainder
## W = 1, p-value = 0.006
```

3 Testes

Os gráficos de seção anterior evidenciaram a possibilidade da série não ser estacionária, assim faz-se necessário verificar se vale de fato para o processo estocástico, com isso vamos precisar aplicar testes de hipóteses para averiguar algumas propriedades, como existência de tendência e sazonalidade.

3.0.1 Tendência

Tendência refere-se a um algum comportamento não - estocástico da série em algum momento do tempo, se tal comportamento só acontece em alguns momentos específicos do tempo, chamamos de tendência estocástica, do contrário é dita determinística. Vamos começar pelos testes de tendência determinística, que em termos sumarizados possuem como hipótese:

$$H_0 : \text{A série possui tendência determinística}$$

Tabela 2: Testes de Tendência Determinística

Testes	P-valor
Cox-Stuart	0.0497*
Wald-Wolfowitz	<0,0001*
Mann-Kendall	0.0386*

Conforme mostrado pela Tabela anterior podemos ver que pelos testes de Cox-Stuart, Wald-Wolfowitz, Mann-Kendall, com índice de significância de 5%, temos que todos os p-valores são menores que 0.05, então concluímos que existe tendência determinística na série temporal.

3.0.2 Sazonalidade

Sazonalidade acontece quando a série possui um comportamento que se repete frequencialmente, vamos nessa subseção testar se existe sazonalidade nas avaliação dos episódios, assim temos como hipótese:

$$H_0 : \text{A série possui sazonalidade}$$

Os testes aplicados foram Kruskal-Wallis, Friedman e Autocorrelação em lags Sazonais, assim gerando a seguinte tabela:

Tabela 3: Testes de Sazonalidade

Testes	P-valor
Kruskal-Wallis	0.745
Friedman	0.878
Autocorrelação em lags Sazonais	0.701

Assim podemos ver que pelos testes aplicados com índice de significância de 5%, concluímos que não existe sazonalidade.

3.0.3 Testes de Raiz unitária

Para começar tal seção, primeiros vamos definir um passeio aleatório em sua forma simplificada:

$$Y_t = Y_{t-1} + \epsilon_t$$

Onde t refere-se aos índices de ordenação, e ϵ um termo aleatório, para facilidade vamos assumir que $E(\epsilon_t) = \mu$ e $var(\epsilon_t) = \sigma^2 \forall t$, assim subtraindo-se Y_{t-1} em ambos os lados tem-se

$$Y_t - Y_{t-1} = \epsilon_t$$

que é um processo estacionário, o processo anterior é dito processo estacionário em diferença, em outros casos também chamado de processo integrado. O exemplo anterior trata-se de um caso mais geral

$$Y_t = \phi Y_{t-1} + \epsilon_t$$

onde evidentemente $\phi = 1$, é fácil mostrar que se $|\phi| < 1$ tem-se um processo estocástico, aqui portanto, estamos interessados nesse caso, para isso utilizaremos testes de hipótese.

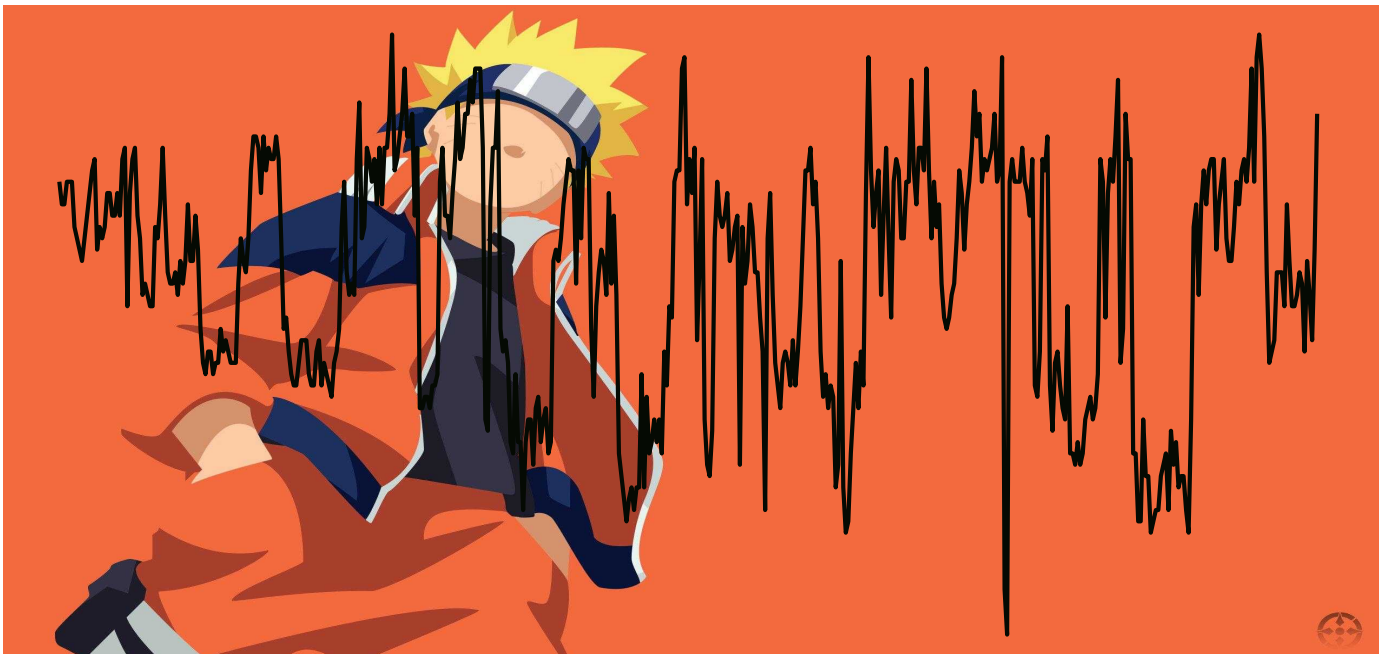
Fala-se também que o caso anterior na forma simplificada possui raiz unitária, por causa do operador Lag, aqui não definido, porém pode-se ler sobre em [1]

Tabela 4: Testes de Tendência Estocástica

Testes	P-valor
Kwiatkowski-Phillips-Schmidt-Shin (KPSS)	0.1
Augmented Dickey-Fuller (ADF)	0.01*
Phillips-Perron (PP)	0.01*

Conforme mostrado pela Tabela anterior podemos ver que pelos testes de Dickey-Fuller Aumentado, Phillips-Perron, KPSS, com índice de significância de 5%, temos que os p-valores dos testes ADF e PP são menores que 0.05 e do teste KPSS é maior que 0.05, então concluímos que não existe tendência estocástica na série temporal. Pois, nos testes ADF e PP, a hipótese nula é a existência de raiz unitária e no teste de KPSS, a hipótese nula é não existência de raiz unitária.

4 Modelagem



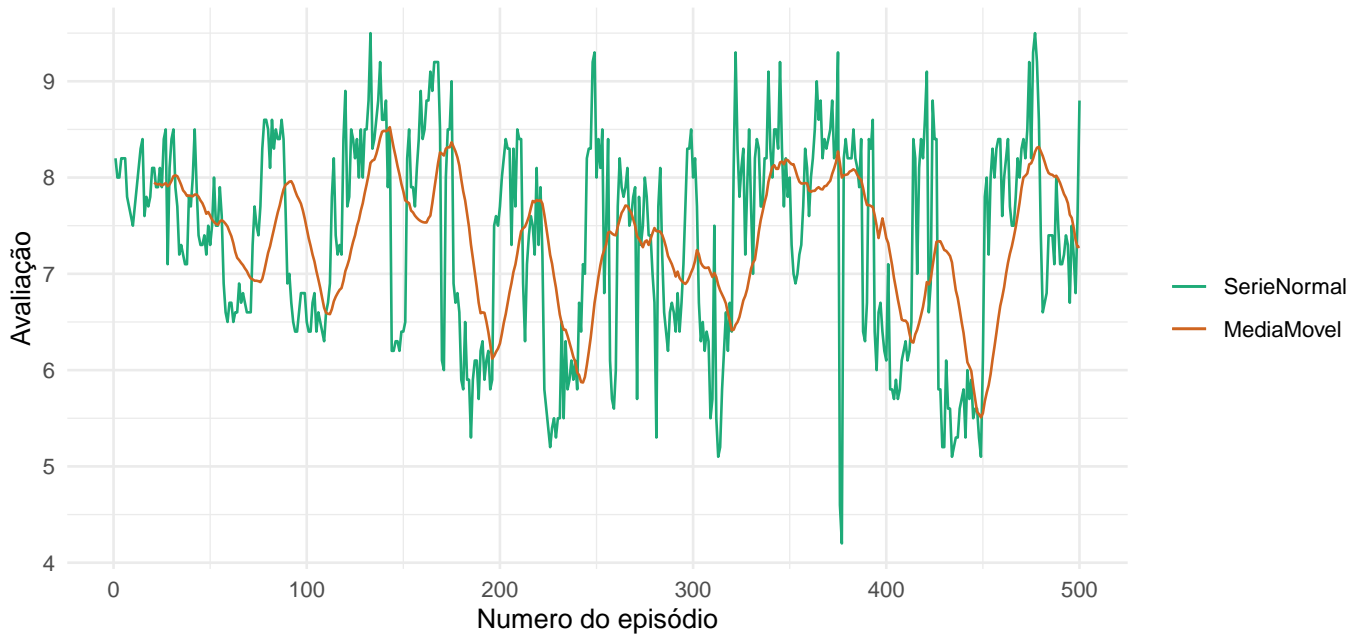
```
## # A tibble: 21 x 2
##   Saga                                mean
##   <chr>                                <dbl>
## 1 s17 The Fourth Great Ninja War - The Return Of Squad Seven 8.52
## 2 s8 Two Saviors                                           8.35
## 3 s6 Prophecy and Vengeance                               8.29
## 4 s15 The Fourth Great Ninja War - Sasuke And Itachi       8.18
## 5 s4 Immortal Devastators                                  8.17
```

```
## 6 s1 Kazekage Rescue 7.99
## 7 s10 The Gathering of the Five Kage 7.74
## 8 s18 The Fourth Great Ninja War - Obito Uchiha Against The Shinobi Forc~ 7.70
## 9 s12 Nine-Tailed Fox Taming And Karmic Encounters 7.61
## 10 s16 Kakashi: Shadow Of The Anbu Black Ops 7.55
## # ... with 11 more rows
```

4.1 Médias Móveis Simples

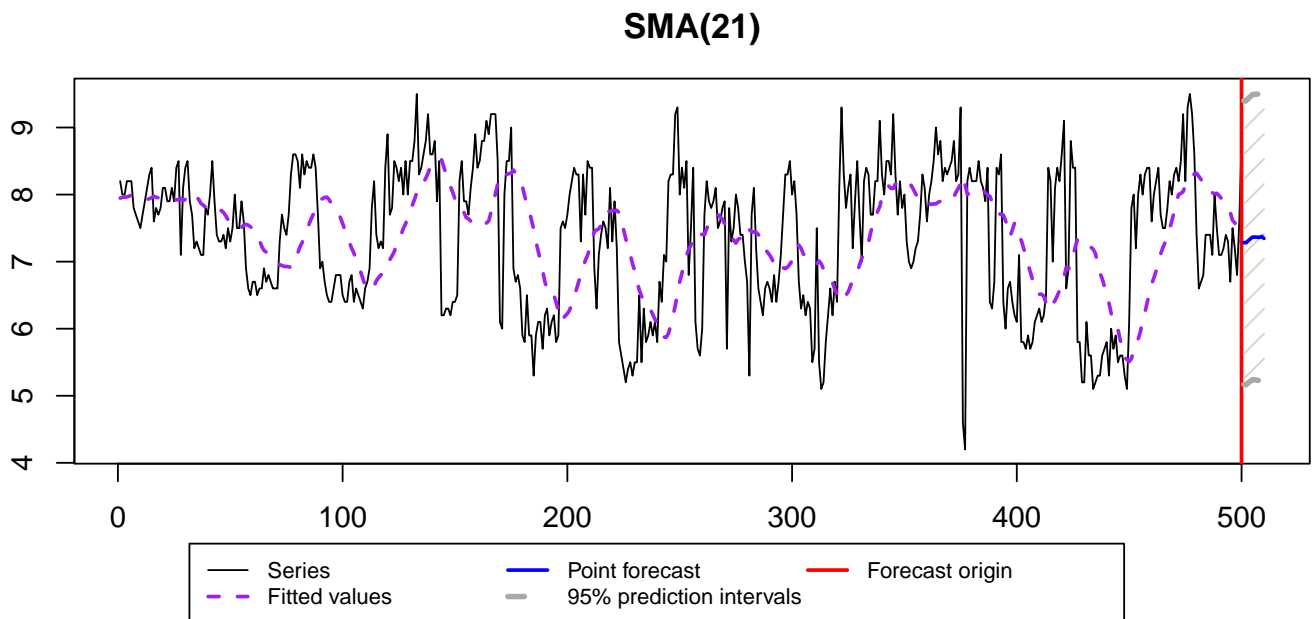
##	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
## 501	8.8	7.92	9.68	7.4489	10.2
## 502	8.8	7.55	10.05	6.8892	10.7
## 503	8.8	7.27	10.33	6.4598	11.1
## 504	8.8	7.03	10.57	6.0977	11.5
## 505	8.8	6.82	10.78	5.7788	11.8
## 506	8.8	6.64	10.96	5.4904	12.1
## 507	8.8	6.46	11.14	5.2252	12.4
## 508	8.8	6.30	11.30	4.9784	12.6
## 509	8.8	6.15	11.45	4.7466	12.9
## 510	8.8	6.01	11.59	4.5273	13.1
## 511	8.8	5.87	11.73	4.3188	13.3
## 512	8.8	5.74	11.86	4.1195	13.5
## 513	8.8	5.61	11.99	3.9284	13.7
## 514	8.8	5.49	12.11	3.7445	13.9
## 515	8.8	5.38	12.22	3.5671	14.0
## 516	8.8	5.27	12.33	3.3954	14.2
## 517	8.8	5.16	12.44	3.2291	14.4
## 518	8.8	5.05	12.55	3.0676	14.5
## 519	8.8	4.95	12.65	2.9105	14.7
## 520	8.8	4.85	12.75	2.7575	14.8
## 521	8.8	4.75	12.85	2.6083	15.0
## 522	8.8	4.66	12.94	2.4626	15.1
## 523	8.8	4.56	13.04	2.3202	15.3
## 524	8.8	4.47	13.13	2.1808	15.4
## 525	8.8	4.38	13.22	2.0443	15.6
## 526	8.8	4.30	13.30	1.9105	15.7
## 527	8.8	4.21	13.39	1.7793	15.8
## 528	8.8	4.13	13.47	1.6504	15.9
## 529	8.8	4.04	13.56	1.5239	16.1
## 530	8.8	3.96	13.64	1.3995	16.2
## 531	8.8	3.88	13.72	1.2772	16.3
## 532	8.8	3.80	13.80	1.1568	16.4
## 533	8.8	3.72	13.88	1.0383	16.6
## 534	8.8	3.65	13.95	0.9216	16.7
## 535	8.8	3.57	14.03	0.8066	16.8
## 536	8.8	3.50	14.10	0.6932	16.9
## 537	8.8	3.43	14.17	0.5813	17.0
## 538	8.8	3.35	14.25	0.4710	17.1
## 539	8.8	3.28	14.32	0.3621	17.2
## 540	8.8	3.21	14.39	0.2546	17.3
## 541	8.8	3.14	14.46	0.1485	17.5
## 542	8.8	3.07	14.53	0.0436	17.6
## 543	8.8	3.01	14.59	-0.0600	17.7
## 544	8.8	2.94	14.66	-0.1624	17.8
## 545	8.8	2.87	14.73	-0.2637	17.9

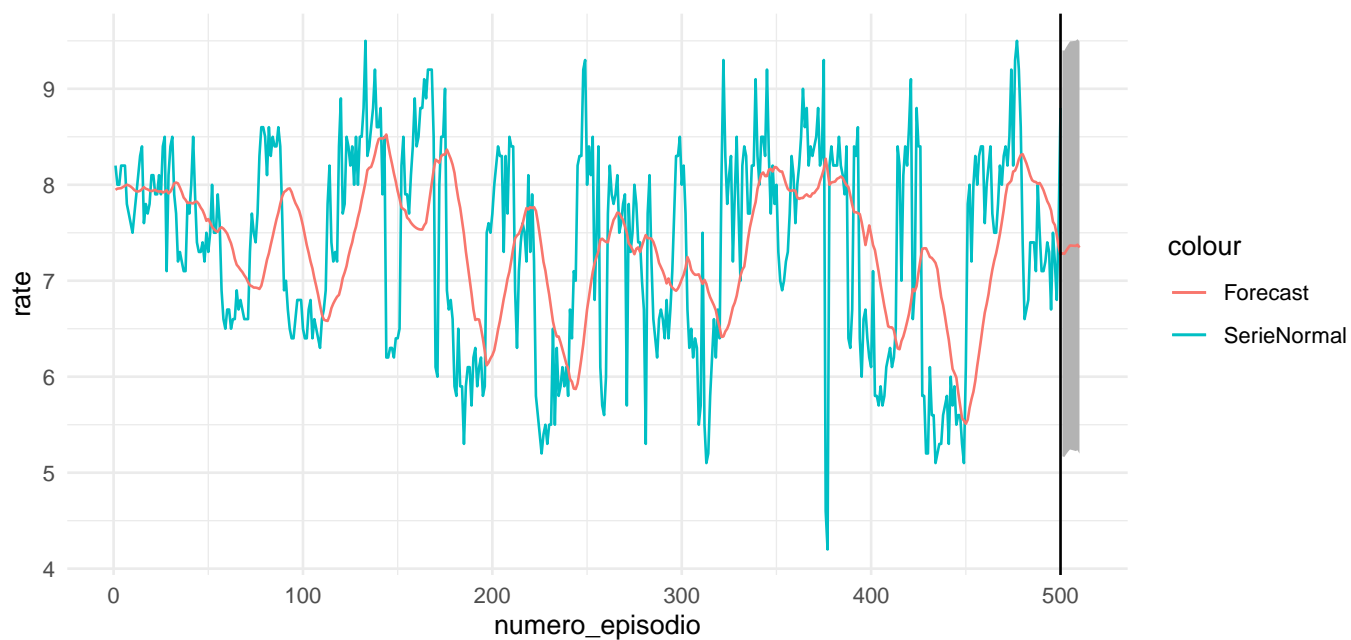
## 546	8.8	2.81	14.79	-0.3639	18.0
## 547	8.8	2.74	14.86	-0.4629	18.1
## 548	8.8	2.68	14.92	-0.5610	18.2
## 549	8.8	2.62	14.98	-0.6580	18.3
## 550	8.8	2.55	15.05	-0.7540	18.4



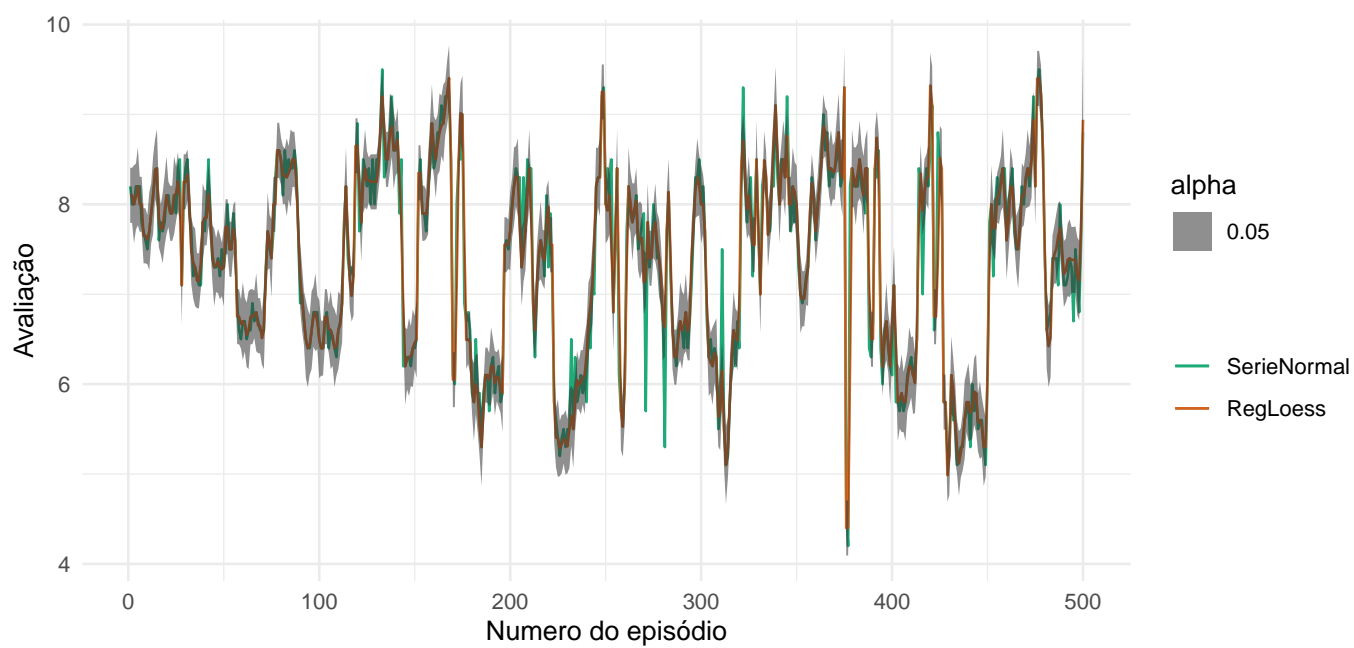
4.2 Média Móvel Simples com Previsão

Utilizando o pacote smooth.





4.3 Regressão LOESS



Referências

- [1] Morettin PA, Toloi CM. Análise de séries temporais: Modelos lineares univariados. Editora Blucher; 2018.