



UNIVERSIDADE ESTADUAL DA PARAÍBA
CAMPUS I - CAMPINA GRANDE
CENTRO DE CIÊNCIAS E TECNOLOGIA
DEPARTAMENTO DE ESTATÍSTICA
CURSO DE GRADUAÇÃO EM BACHARELADO EM ESTATÍSTICA

ANA BEATRIZ RAMOS DA SILVA

ANÁLISE TEMPORAL DA EVOLUÇÃO DE BUSCAS NO GOOGLE
SOBRE AUTISMO: TENDÊNCIAS DE INCLUSÃO E COMPREENSÃO.

CAMPINA GRANDE - PB
2025

ANA BEATRIZ RAMOS DA SILVA

**ANÁLISE TEMPORAL DA EVOLUÇÃO DE BUSCAS NO GOOGLE
SOBRE AUTISMO: TENDÊNCIAS DE INCLUSÃO E COMPREENSÃO.**

Trabalho de Conclusão de Curso apresentado à Coordenação do Curso de Estatística da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de Bacharela em Estatística.

Orientador(a): Prof. Dr. Silvio Fernando Alves Xavier Júnior

Coorientador(a): Prof. Pedro Augusto Ferreira da Silva

CAMPINA GRANDE - PB

2025

É expressamente proibida a comercialização deste documento, tanto em versão impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que, na reprodução, figure a identificação do autor, título, instituição e ano do trabalho.

S586a Silva, Ana Beatriz Ramos da.
Análise temporal da evolução de buscas no google sobre autismo [manuscrito] : tendências de inclusão e compreensão / Ana Beatriz Ramos da Silva. - 2025.
44 f. : il. color.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Estatística) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2025.

"Orientação : Prof. Dr. Silvio Fernando Alves Xavier Junior, Departamento de Estatística - CCT".

"Coorientação: Prof. Grad. Pedro Augusto Ferreira da Silva, DE".

1. Transtornos do espectro autista. 2. Análise temporal. 3. Modelos de previsão. I. Título

21. ed. CDD 310

ANA BEATRIZ RAMOS DA SILVA

ANÁLISE TEMPORAL DA EVOLUÇÃO DE BUSCAS NO GOOGLE SOBRE
AUTISMO: TENDÊNCIAS DE INCLUSÃO E COMPREENSÃO

Trabalho de Conclusão de Curso
apresentado à Coordenação do Curso
de Estatística da Universidade Estadual
da Paraíba, como requisito parcial à
obtenção do título de Bacharela em
Estatística

Aprovada em: 14/05/2025.

BANCA EXAMINADORA

Documento assinado eletronicamente por:

- **Ricardo Alves de Olinda** (***.035.004-**), em **22/05/2025 06:12:26** com chave **e156cda836ec11f086e606adb0a3afce**.
- **Débora de Sousa Cordeiro** (***.592.134-**), em **21/05/2025 20:58:22** com chave **7a2aa2b4369f11f0ac1806adb0a3afce**.
- **Silvio Fernando Alves Xavier Junior** (***.025.684-**), em **21/05/2025 21:45:23** com chave **0b94f98836a611f08e9506adb0a3afce**.

Documento emitido pelo SUAP. Para comprovar sua autenticidade, faça a leitura do QRCode ao lado ou acesse https://suap.uepb.edu.br/comum/autenticar_documento/ e informe os dados a seguir.

Tipo de Documento: Folha de Aprovação do Projeto Final

Data da Emissão: 22/05/2025

Código de Autenticação: 4a97de



“À minha mãe e ao meu esposo, por todo o apoio incondicional no decorrer da minha jornada, DEDICO.”

AGRADECIMENTOS

Agradeço, primeiramente, a Deus, pois sem Ele nada me seria possível. À minha mãe, Ana Paula, que nunca mediu esforços para que eu pudesse estudar e sempre lutou por mim com garra e amor. Ao meu avô, Paulo, que não pôde estudar, mas, ainda assim, fazia questão de me levar à escola todos os dias, ensinando-me, com seu exemplo, o verdadeiro valor da dedicação e do amor. Ao meu esposo, Nathan, que, em todos os momentos, foi meu abrigo, meu porto seguro, meu melhor amigo e nunca mediu esforços para me ver realizar este e tantos outros sonhos. Sem vocês, nada seria possível.

Também à minha banca por aceitar o convite e por todo o apoio oferecido neste projeto e em outros que já realizamos juntos. Agradeço ao meu coorientador por toda a ajuda no desenvolvimento deste trabalho. Ademais, expresso meus agradecimentos as minhas amigas de caminhada Luana e Giovana.

Por fim, sou grata ao meu orientador por ter aceitado me orientar, pelos ensinamentos que vão além das fronteiras deste trabalho e por sempre ser tão solícito. Ademais, a todo o corpo docente da UEPB sou grata por todo aprendizado compartilhado no decorrer deste curso.

“O coração do homem planeja o seu caminho,
mas o Senhor lhe dirige os passos.”
(Provérbios 16:9)

RESUMO

Este estudo teve como principal objetivo avaliar o comportamento das buscas pelo termo “autismo” no buscador Google, considerando as pesquisas realizadas no Brasil, para, em seguida, prever a popularidade futura do termo. Isso se justifica pelo crescente interesse na condição de saúde, suas comorbidades e possíveis causas associadas a essa neurodivergência. Observou-se que, ao longo dos anos analisados, há uma forte tendência de aumento no interesse pelo termo. Dessa forma, o comportamento da série temporal resultante dessas buscas não pode ser considerado estacionário. Assim, a aplicação se deu a partir de modelos capazes de lidar adequadamente com séries temporais não estacionárias. Para essa análise, foram utilizados os modelos *ARIMA*, *Holt-Winters (HW)* nas formas aditiva e multiplicativa, além do *Prophet* também com sazonalidade aditiva e multiplicativa. Para comparar os modelos que apresentaram bons resultados e selecionar o mais adequado, foi avaliado o erro quadrático médio (*RMSE*). Os resultados indicaram que o modelo *Holt-Winters* na forma aditiva apresentou o melhor ajuste, capturando de maneira mais eficiente os componentes de sazonalidade e tendência dessa série temporal. O estudo evidencia o aumento do interesse pelo tema ao longo dos anos, ressaltando a importância do monitoramento dessas buscas e o potencial avanço na inclusão de indivíduos que possuem esse espectro.

Palavras-chave: TEA; análise temporal; modelos de previsão.

ABSTRACT

This study aimed to evaluate the behavior of searches for the term "autism" on the Google browser, considering queries made in Brazil, in order to subsequently predict the future popularity of the term. This is justified by the growing interest in the health condition, its comorbidities, and possible causes associated with this neurodivergence. It was observed that, over the analyzed years, there is a strong upward trend in interest in the term. Therefore, the behavior of the time series resulting from these searches cannot be considered stationary. Thus, the applied models must be capable of adequately handling non-stationary time series. For this analysis, the ARIMA, Holt-Winters (HW) models in both additive and multiplicative forms, as well as the Prophet model, were used. The Holt-Winters model was applied with both additive and multiplicative seasonality. To compare the models that showed good results and select the most suitable one, the Root Mean Square Error (RMSE) was evaluated. The results indicated that the Holt-Winters model in its additive form provided the best fit, capturing the seasonality and trend components of this time series more effectively. The study highlights the increasing interest in the topic over the years, emphasizing the importance of monitoring these searches and the potential advancement in the inclusion of individuals with this condition.

Keywords: ASD; temporal analysis; predictive models.

LISTA DE ILUSTRAÇÕES

Figura 3.1 – Série temporal da busca pelo termo “autismo” no Google de janeiro de 2004 até janeiro de 2025.	25
Figura 3.2 – Decomposição da série de busca por “autismo” em pontuação na escala original em componetes sazonais, tendência e aleatoriedade dos resíduos.	26
Figura 3.3 – Série temporal da busca pelo termo “autismo” no Google de janeiro de 2004 até janeiro de 2025 em escala logarítmica.	27
Figura 3.4 – Funções de autocorrelação amostral e parcial da série original. . .	27
Figura 3.5 – Série temporal da busca pelo termo “autismo” no Google de janeiro de 2004 até janeiro de 2025 em escala logarítmica e diferenciada. .	28
Figura 3.6 – Funções de autocorrelação amostral e parcial da série logarítimizada e diferenciada.	29
Figura 3.7 – Decomposição da série de busca por “autismo” em pontuação na escala e diferenciada em componetes sazonais, tendência e aleatoriedade dos resíduos.	30
Figura 3.8 – Previsão do modelo $ARIMA(0,0,1)(0,1,1)_{12}$ 12 passos à frente para a pontuação de popularidade do termo “autismo”.	31
Figura 3.9 – Ajuste do modelo de <i>Holt-Winters</i> Aditivo um passo à frente (linha vermelha) para a série em escala log (linha preta).	33
Figura 3.10 – Ajuste do modelo de <i>Holt-Winters</i> Multiplicativo um passo à frente (linha vermelha) para a série em escala log (linha preta).	34
Figura 3.11 – Previsão de <i>Holt-Winters</i> 12 passos à frente no modelo Multiplicativo.	35
Figura 3.12 – Previsão de <i>Holt-Winters</i> 12 passos à frente no modelo Aditivo. .	35
Figura 3.13 – Ajuste e previsão do modelo <i>Prophet</i> Multiplicativo.	37
Figura 3.14 – Ajuste e previsão do modelo <i>Prophet</i> Aditivo	38

LISTA DE TABELAS

Tabela 2.1 – Avaliação de raiz unitária utilizando teste Dickey-Fuller	22
Tabela 3.1 – Resumo do Modelo <i>ARIMA</i>	31
Tabela 3.2 – Previsão do modelo <i>ARIMA</i> (0, 0, 1)(0, 1, 1) ₁₂ 12 passos à frente para a pontuação de popularidade do termo “autismo”.	32
Tabela 3.3 – Valores das constantes de suavização otimizados para os modelos Aditivo e Multiplicativo na suavização exponencial tripla para a série de popularidade do termo “autismo” no Google.	33
Tabela 3.4 – Previsões Mensais para os Modelos Aditivo e Multiplicativo de <i>Holt-Winters</i>	34
Tabela 3.5 – Métricas de erro dos Modelos Aditivo e Multiplicativo.	36
Tabela 3.6 – Critérios de Informação para os Modelos Aditivo e Multiplicativo	36
Tabela 3.7 – Previsões mensais para os Modelos Aditivo e Multiplicativo <i>Prophet</i>	38
Tabela 3.8 – Métricas de erro e coeficiente de determinação dos Modelos Aditivo e Multiplicativo para o algoritmo <i>Prophet</i>	39

SUMÁRIO

	Página
1	INTRODUÇÃO 11
2	METODOLOGIA 13
2.1	Conceitos básicos 13
2.2	Métodos Estatísticos de previsão de séries temporais 14
2.3	Métodos de Suavização Exponencial 14
2.3.1	Suavização Exponencial Simples 15
2.3.2	Suavização Exponencial Dupla ou Suavização Exponencial de <i>Holt</i> 15
2.3.3	Suavização Exponencial Tripla 16
2.4	<i>Prophet</i> 18
2.5	Modelagem <i>Box-Jenkins</i> 18
2.5.1	Modelos autoregressivos (<i>AR</i>) 19
2.5.2	Modelos de médias móveis (<i>MA</i>) 19
2.5.3	Modelo auto-regressivo de médias móveis (<i>ARMA</i>) 19
2.5.4	Modelo auto-regressivo integrado de médias móveis (<i>ARIMA</i>) 19
2.5.5	Modelo auto-regressivo de médias móveis sazonal (<i>SARIMA</i>) 20
2.5.6	Modelo auto-regressivo de médias móveis sazonal com variáveis exógenas (<i>SARIMAX</i>) 20
2.5.7	Função de autocorrelção (<i>FAC</i>) 20
2.5.8	Função de autocorrelção parcial (<i>FACP</i>) 21
2.6	Teste de estacionariedade de Dickey-Fuller 21
2.7	Validação do ajuste 22
2.8	Avaliação da precisão do ajuste 23
3	APLICAÇÃO 25
3.1	Análise Exploratória 25
3.2	Modelagem 30
3.2.1	Modelagem <i>Box-Jenkins</i> 30
3.2.2	Modelagem Holt-Winters 32
3.2.3	Algoritmo Prophet 37
4	CONCLUSÃO 40
	REFERÊNCIAS 41

1 INTRODUÇÃO

O Transtorno do Espectro Autista (TEA) para Ramos e Fujita (2017), é uma condição associada a diversos fatores que expõem diferentes manifestações e intensidades. De acordo com Robertson e Baron-Cohen (2017), indivíduos caracterizados com este transtorno podem apresentar interesses exclusivos, comportamento padronizado ou monótono, além de divergências no processamento sensorial. Para Rios et al. (2015), autismo não é uma doença contagiosa. Contudo, fala-se de uma epidemia ao autismo como forma de denotar o aumento repentino no número de casos em curto período de tempo.

Segundo Grinker (2008), o aumento exorbitante no número de casos de autismo pode ser atribuído à mudança na descrição fornecida pela psiquiatria a um conjunto de sintomas e comportamentos que já eram anteriormente apresentados na população. "Não foi a epidemia que fez o autismo visível, mas a visibilidade do autismo que fez a epidemia" (Eyal, 2010). Assim, os cuidados e busca por informação trouxe grandes impactos no diagnóstico do autismo e substituição da suposição de retardo mental.

O número de pesquisas sobre o TEA também cresceu consideravelmente. Entre os anos de 1940 e 1999, foram publicados 6.054 artigos que falam sobre esta condição. De acordo com Constanzo et al. (2015), posteriormente, no período de 2000 a 2012, 16.741 artigos foram encontrados, formando assim quase o triplo de buscas das décadas mencionadas anteriormente. Assim, é nítido o interesse e preocupação que cerca as particularidades do TEA.

Diante do expressivo aumento no número de casos nos últimos anos, é plausível supor que o interesse da população também tenha se intensificado, revelando diferentes padrões de comportamento ao longo do tempo. Para captar essas dinâmicas de forma mais aprofundada, faz-se necessário recorrer a uma abordagem estatística mais robusta do que uma análise exploratória inicial, trata-se da análise de séries temporais. Para Morettin e Tolo (2006) (2006), uma série temporal é composta por um conjunto de observações ordenadas no tempo sobre determinado fenômeno. Já para Martínez e Zamprogn (2003), essa análise permite compreender a evolução de uma variável ao longo do tempo, além de possibilitar previsões futuras com base nesse comportamento. Nesse cenário, o uso da análise de séries temporais mostra-se particularmente relevante diante do notável crescimento nos diagnósticos de Transtorno do Espectro Autista (TEA) nas últimas décadas.

Para Borges et al. (2024), esse aumento tem sido alarmante, podendo ser atribuído a diversas causas, dentre eles estão fatores que vão desde uma maior conscientização até fatores genéticos, climáticos e ambientais. Segundo Lord et al. (2018), o TEA, anteriormente considerado uma condição rara, atualmente afeta cerca de 1 a 1,15% da população mundial, embora essa taxa varie conforme fatores geográficos e sociodemográficos. Essas transformações reforçam a importância de acompanhar e compreender a evolução do in-

teresse social e científico sobre o tema ao longo do tempo.

Nesse sentido, este estudo tem como principal objetivo aplicar modelos de análise de séries temporais para entender e prever o comportamento futuro do interesse da população brasileira pelo TEA, considerando que a compreensão e a inclusão são passos fundamentais para a promoção de uma sociedade mais informada e acolhedora.

2 METODOLOGIA

Nesta seção serão expostos os conceitos que compõem a metodologia utilizada para a construção deste estudo. Sendo estes, os conceitos básicos que introduzem o que é uma série temporal e suas características, a definição das três modelagens abordadas, os testes que avaliam o comportamento da janela temporal, as métricas de avaliação e por fim, a validação dos ajustes aplicados.

2.1 Conceitos básicos

Segundo Morettin e Tolo (2006) uma série temporal é um conjunto de observações de uma variável aleatória ordenadas no tempo t , ou seja, o conjunto de observações em relação a variável de interesse são avaliados ao longo do tempo. As janelas temporais podem ser discretas ou contínuas, tal característica irá depender da natureza da variável de interesse. Caso a variável se comporte de forma discreta, tem-se uma série temporal discreta por outro lado, em situações que a variável resposta assuma valores contínuos em seu suporte tem-se uma série temporal contínua. Ademais, uma série temporal discreta pode ser extraída de uma série de natureza contínua. Para entender o comportamento de uma série temporal, suas suposições e interpretar o que ela demonstra da característica de interesse, faz-se necessário entender os conceitos básicos a seguir:

- **Tendência:** Refere-se ao comportamento da variável de interesse ao longo do tempo. A tendência pode mostrar comportamentos diferentes como: crescimento ou decrescimento, linear e exponencial. Para obter a tendência da série temporal observada tem-se 3 métodos distintos que são: mínimos quadrados, modelos de médias móveis e ajuste exponencial.
- **Sazonalidade:** Similarmente a tendência, a sazonalidade revela sobre o comportamento da série, entretanto, de maneira distinta, com a sazonalidade pode-se observar se o comportamento da série se repete ao longo do tempo.
- **Nível:** Valor ou faixa recorrente de valores que a variável pode assumir na oscilação do conjunto de observações em relação a tendência da série.
- **Erro:** Diferença entre o valor observado na série e a previsão realizada através da obtenção de seu comportamento.

Um processo estocástico pode ser definido como: $Z = \{Z(t), t \in T\}$. Dessa forma, um processo estocástico é um conjunto de variáveis aleatórias, em que todas precisam estar definidas no mesmo espaço de probabilidades (Ω, A, P) . O conjunto t é normalmente tomado como conjuntos dos inteiros ou o conjunto dos números reais (Morettin; Tolo,

2018). Sendo assim, $Z(t)$ é uma função de dois argumentos com parâmetro t e sua probabilidade (Ω).

Estacionariedade: Para uma série temporal ser considerada estacionária sua média é constante em todo processo, mesmo que ela se desenvolva aleatoriamente no decorrer do tempo. Dito isto, uma série temporal estacionária tem sua média e covariância totalmente “livres” de t . A ideia primordial do conceito de estacionariedade é que as leis probabilísticas que guiam a série temporal não se alterem durante toda a movimentação da série. Uma série estacionária pode ter estacionariedade forte ou fraca.

Considera-se que uma série temporal é fracamente estacionária, se sua média μ é livre de t , ou seja, é constante em todo processo da série para $\forall t$ bem como, sua autocovariância $\gamma_{t,s} = Cov(X_t, X_s)$ depende apenas da diferença do intervalo percorrido $t - s$. Em contrapartida, para uma série temporal ser estritamente estacionária, isto é, com estacionariedade forte, tem-se condições estatísticas mais robustas, nesse caso precisa-se que sua média, variância e autocovariância sejam constantes durante todo o processo e para $\forall t$.

2.2 Métodos Estatísticos de previsão de séries temporais

Segundo Morettin e Tolo (2006, p. 83), “A maioria dos métodos de previsão baseia-se na ideia de que observações passada contém informações sobre o padrão de comportamento da série temporal”. Logo, observando padrões anteriores pode-se prever comportamentos futuros com um menor risco de decisões inapropriadas. Para minimização do erro, também, faz-se necessário a utilização do método correto para implementação. Ou seja, os métodos de previsão de séries temporais trazem a possibilidade de projeção de valores futuros a serem assumidos pela variável de interesse, baseando-se nos valores já assumidos por tal variável no passado e no presente. Com uma análise de séries temporais, busca-se a construção de modelos, sejam eles estatísticos ou matemáticos, que possam descrever o comportamento da série e prever padrões que se manifestarão ao longo do tempo. No presente trabalho, serão abordados os seguintes métodos de previsão estatísticos: Suavização Exponencial, *Box-Jenkins* e *Prophet*.

2.3 Métodos de Suavização Exponencial

A previsão por Suavização ou Alisamento Exponencial utiliza somente a série de dados abordada, sem necessidade de variáveis externas para realização das previsões (Souza; Samohyl; Miranda, 2009). Segundo Morettin e Tolo (2006, p. 83), “a grande popularidade atribuída aos métodos de suavização é devida à simplicidade, à eficiência computacional e à sua razoável precisão”. Tal método, traz consigo de maneira entendível a sazonalidade, o nível, o erro e a tendência que os dados em estudo apresentam. Neste sentido, serão apresentadas três possibilidades mais usuais de alisamento exponencial.

2.3.1 Suavização Exponencial Simples

Pode-se utilizar o método de Suavização Exponencial Simples (SES) quando na série temporal abordada não existe indícios de sazonalidade nem tendências, em outras palavras, os valores correspondentes a variável de interesse oscilam de maneira aleatória em torno de uma constante, de tal forma que não seja possível avaliar tendências ou comportamento sazonal. De acordo com Morettin e Tolo (2006) o modelo matemático pode ser descrito como:

$$\bar{Z}_t = \alpha Z_t + (1 - \alpha)\bar{Z}_{t-1}, \bar{Z}_1 = Z_1, \quad t = 1, \dots, N. \quad (2.1)$$

Em que, \bar{Z}_t é o valor exponencialmente alisado e α é a constante presente no alisamento $0 \leq \alpha \leq 1$. A medida que o valor de α diminui, as previsões finais se tornam mais estáveis. Logo, quanto menor o valor α mais aleatória é a série estudada.

Vantagens do AES -

Ainda sobre os autores mencionados acima, existem algumas vantagens em utilizar o (AES), tais quais:

- i. Facilidade na compreensão;
- ii. Solução econômica;
- iii. Flexibilidade pela variação da constante α ;
- iv. Necessidade de armazenar somente Z_t , \bar{Z}_t e α ; e
- v. O valor de $\alpha = 2/(r - 1)$ traz estimativas similares ao método de médias móveis simples (MMS) com parâmetro r Montgomery, Jennings e Kulahci (2015).

Desvantagens do AES -

Como desvantagem tem-se, primordialmente, a dificuldade em determinar o melhor valor de α para minimizar o erro e construir um bom modelo de previsão.

2.3.2 Suavização Exponencial Dupla ou Suavização Exponencial de *Holt*

Para a SES, tem-se como critério que a série temporal estudada não possua tendência, dado a isso, Holt (1957) desenvolveu o Alisamento Exponencial Duplo, que é uma extensão da Suavização Exponencial Simples. Tal método pode ser usado quando se é encontrado tendência linear nos dados, seja ela crescente ou decrescente, mas, não seja encontrado indícios de sazonalidade. Ainda sobre Morettin e Tolo (2006), para os casos em que a série apresenta tendências lineares sendo ela positiva ou negativa, os resultados obtidos para previsão dos valores futuros a serem assumidos podem ser minimizados ou exagerados. Para a Suavização Exponencial de *Holt*, similarmente a SES, utiliza-se uma constante

para suavização, porém, ao invés de adotar apenas uma constante para o alisamento do nível, acrescenta-se uma nova invariável que descreva a tendência presente na série. Para a estimação dos valores do nível e tendência da série, tem-se as seguintes equações

$$\bar{Z}_t = \alpha Z_t + (1 - \alpha)(\bar{Z}_{t-1} + \hat{T}_{t-1}), \quad 0 < \alpha < 1 \text{ e } t = 2, \dots, N, \quad (2.2)$$

$$\hat{T}_t = \beta(\bar{Z}_t - \bar{Z}_{t-1}) + (1 - \beta)\hat{T}_{t-1}, \quad 0 < \beta < 1 \text{ e } t = 2, \dots, N, \quad (2.3)$$

α e β são as constantes de suavização. As equações 2.2 e 2.3 alteram as previsões feitas a priori quando uma nova observação é adicionada aos dados, como ocorre em todos os métodos de alisamento. \bar{Z}_t traz a estimativa do nível da série avaliada no tempo t e \hat{T}_t é a estimativa da tendência linear no tempo t . Para as constantes α e β busca-se encontrar valores que minimizam o erro quadrático médio.

Vantagens do Alisamento Exponencial Duplo - Similarmente ao (SES), tem-se as seguintes vantagens: (Morettin; Tolo, 2006).

- i. Simples de compreender;
- ii. Solução econômica;
- iii. Flexibilidade pela variação das constantes α e β .

Desvantagens do Alisamento Exponencial Duplo -

Como desvantagem, também similarmente ao modelo anterior, a principal limitação é a dificuldade em determinar os mais qualificados valores para as constantes de alisamento α e β .

2.3.3 Suavização Exponencial Tripla

O método de Suavização Exponencial Tripla ou método de *Holt-Winters* é uma extensão da suavização exponencial dupla, ou seja, na suavização dupla o método conta com tendência na série, mas não pode ser usado em casos que a série possua sazonalidade. Por isto, fez-se necessário a expansão do método de Holt (1957) e foi desenvolvido por Winter (1960). O modelo de *Holt-Winters* foi desenvolvido para ser aplicado em séries que possuam tendências e sazonalidades. Ademais, tem-se procedimentos diferentes para aplicação de *Holt-Winters*. O método a ser aplicado irá depender da natureza da série estudada (Morettin; Tolo, 2006). Os procedimentos baseiam-se em 3 equações com constantes de alisamento, similarmente à Suavização Simples e Dupla, com acréscimo de mais uma constante que se associe ao componente sazonal da série. Os procedimentos são de Série Sazonal Multiplicativa e Série Sazonal Aditiva.

- **Série Sazonal Multiplicativa**

Considere um série sazonal com período s . Para Morettin e Toloi (2006) a variante mais usual do método *HW* considera o fator sazonal F_t como sendo multiplicativo, enquanto a tendência permanece aditiva. Em suma, o método *HW* para efeitos sazonais multiplicativos é aplicado em modelos que a dimensão do ciclo de sazonalidade varia harmoniosamente com o nível da série, portanto, tem-se a seguinte expressão:

$$Z_t = \mu F_t + T_t + a_t, \quad t = 1, \dots, N, \quad (2.4)$$

As equações de alisamento, são:

$$\hat{F}_t = \gamma \left(\frac{Z_t}{\bar{Z}_t} \right) + (1 - \gamma) \hat{F}_{t-s}, \quad 0 < \gamma < 1, \quad t = s + 1, \dots, N, \quad (2.5)$$

$$\bar{Z}_t = \alpha \left(\frac{Z_t}{\hat{F}_{t-s}} \right) + (1 - \alpha) (\bar{Z}_{t-1} + \hat{T}_{t-1}), \quad 0 < \alpha < 1, \quad t = s + 1, \dots, N, \quad (2.6)$$

$$\hat{T}_t = \beta (\bar{Z}_t - \bar{Z}_{t-1}) + (1 - \beta) \hat{T}_{t-1}, \quad 0 < \beta < 1, \quad t = s + 1, \dots, N. \quad (2.7)$$

As Equações 2.5, 2.6 e 2.7 representam as estimativas de sazonalidade, nível e tendência e suas respectivas contantes de alisamento (Morettin; Toloi, 2006). Estas equações permitem o alisamento da série, sendo essenciais para os modelos que compõem a modelagem de *Holt-Winters*. Ademais, com a utilização destes parâmetros o modelo se torna flexível ao comportamento da série.

• Série Sazonal Aditiva

O método *HW* para sazonalidade aditiva é aplicado onde o nível da série é independente da dimensão do ciclo sazonal. Quando o fator sazonal é aditivo, tem-se a seguinte expressão:

$$Z = \mu_t + T_t + F_t + a_t. \quad (2.8)$$

Logo, as equações de alisamento tornam-se

$$\hat{F}_t = \gamma (Z_t - \bar{Z}_t) + (1 - \gamma) \hat{F}_{t-s}, \quad 0 < \gamma < 1, \quad (2.9)$$

$$\bar{Z}_t = \alpha (Z_t - \hat{F}_{t-s}) + (1 - \alpha) (\bar{Z}_{t-1} + \hat{T}_{t-1}), \quad 0 < \alpha < 1, \quad (2.10)$$

$$\hat{T}_t = \beta (\bar{Z}_t - \bar{Z}_{t-1}) + (1 - \beta) \hat{T}_{t-1}, \quad 0 < \beta < 1, \quad (2.11)$$

Dessa forma, α , β e γ correspondem aos parâmetros de alisamento utilizados no modelo de suavização exponencial tripla.

2.4 *Prophet*

O *Prophet* é um algoritmo desenvolvido pelo *Facebook* (grupo Meta) que possui grandes capacidades para visualizar e melhorar projeções futuras. Para Taylor e Letham (2018), a proposta desse modelo é ser facilmente aplicável e com uso focado em séries temporais que apresentam fortes padrões sazonais, a sazonalidade é considerada aditiva quando o comportamento é constante ao longo da janela temporal, e multiplicativa quando a variação sazonal se altera de acordo com o nível da série. e efeitos gerados por feriados ou datas específicas. O *Prophet* é composto por três componentes: tendência, sazonalidade e feriados. Logo, o modelo pode ser descrito pela seguinte expressão matemática:

$$y(t) = g(t) + s(t) + h(t) + e_t, \quad (2.12)$$

em que $g(t)$ representa as mudanças associadas à tendência, $s(t)$ é a variação que compõe a sazonalidade, $h(t)$ captura os feriados, se existentes, e e_t é o erro associado a esse ajuste.

Ademais, esse modelo permite a alteração de parâmetros para que seja o mais fiel possível ao comportamento dos dados. Neste estudo, serão abordadas as metodologias provenientes do *Prophet* que consideram a sazonalidade como aditiva e multiplicativa.

2.5 Modelagem *Box-Jenkins*

A metodologia *Box-Jenkins* foi desenvolvida por George Box e Gwilym Jenkins na década de 1970 (Box; Jenkins, 1976). Tal método é frequentemente utilizado para analisar modelos paramétricos sejam estacionários ou não. Segundo Morettin e Toloí (2006) esta metodologia ajusta modelos autoregressivos integrados de médias móveis, sendo estes, modelos *ARIMA*(p, d, q).

Ainda sobre Morettin e Toloí (2006), esta metodologia se dá com um ciclo iterativo, no qual os dados são responsáveis pela estrutura do modelo, sendo os ciclos seguintes:

- **Especificação:** um grupo geral de modelos é considerado para a análise;
- **Identificação:** análise de autocorrelações, autocorrelações parciais e outras métricas estatísticas;
- **Estimação:** os parâmetros do modelo são estimados;
- **Análise de Diagnóstico:** avaliação do modelo ajustado, para verificar sua adequação através dos resíduos.

Caso o modelo ajustado não seja adequado o ciclo é repetido, além disso, existe a possibilidade de mais de um modelo se ajustar bem aos dados. Para Morettin e Toloí (2006) se o objetivo deste modelo for previsão, preferencialmente deve-se escolher o que forneça o menor erro quadrático médio para a previsão.

2.5.1 Modelos autoregressivos (*AR*)

Considerando que ϵ_t seja um processo aleatório com média zero e variância σ_ϵ^2 o processo \bar{Z} , é chamado auto-regressivo. Os modelos auto-regressivos fazem regressões sobre sí próprio (Cryer; Chan, 2008). Um *AR* de ordem (p) é dado pela expressão apresentada a seguir:

$$\bar{Z}_t = \phi_1 \bar{Z}_{t-1} + \phi_2 \bar{Z}_{t-2} + \dots + \phi_p \bar{Z}_{t-p} + a_t. \quad (2.13)$$

O operador auto-regressivo de ordem p , pode ser escrito como:

$$\phi(B)\bar{Z}_t = a_t. \quad (2.14)$$

2.5.2 Modelos de médias móveis (*MA*)

Seja ϵ_t um processo aleatório com média zero e variância σ_ϵ^2 , o processo \bar{Z} é denotado como modelo de médias móveis. Para os modelos de médias móveis de ordem (q), a equação segundo Morettin e Tolo (2006) é denotada como:

$$\bar{Z}_t = (1 - \theta_1 B - \dots - \theta_q B^q)a_t = \theta(B)a_t, \quad (2.15)$$

de forma que,

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (2.16)$$

é operador de médias móveis com ordem q .

2.5.3 Modelo auto-regressivo de médias móveis (*ARMA*)

Ainda sobre Morettin e Tolo (2006) os modelos *ARMA* como o nome sugere é uma combinação dos modelos auto-regressivos (*AR*) e de médias móveis (*MA*). Logo, uma solução adequada em alguns casos para a diminuição do número excessivos de parâmetros é adicionar componentes auto-regressivos e de médias móveis tendo a equação definida da seguinte forma:

$$\bar{Z}_t = \phi_1 \bar{Z}_{t-1} + \dots + \phi_p \bar{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}. \quad (2.17)$$

Logo, ϕ e θ são os operadores auto-regressivos e de médias móveis respectivamente.

2.5.4 Modelo auto-regressivo integrado de médias móveis (*ARIMA*)

Quando as séries não são estacionárias uma alternativa é a generalização do modelo *ARMA*, daí surge o modelo *ARIMA*. Para tornar uma série estacionária podemos usar a diferenciação. Geralmente duas diferenciações são suficientes para que a série torne-se estacionária, em outros casos, outros tipos de transformações são necessárias para

obtenção da estacionariedade da série (Montgomery; Jennings; Kulahci, 2015). A equação do modelo *ARIMA* é dada por:

$$W_t = \phi_1 W_{t-1} + \dots + \phi_p W_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} \quad (2.18)$$

2.5.5 Modelo auto-regressivo de médias móveis sazonal (*SARIMA*)

Os modelos *SARIMA* podem ser pensados como uma extensão dos modelos *ARIMA*, levando em consideração a adição do componente sazonal. A notação deste modelo é *SARIMA* (p,d,q) x (P,D,Q), de forma que:

- **P** é a quantidade de componentes sazonais e auto-regressivos estacionários.
- **D** é a quantidade de diferenciações utilizadas para tornar a série estacionária.
- **Q** é a quantidade de componentes sazonais de médias móveis que referem-se a invertibilidade do modelo.

Logo, a equação de ajuste do modelo é:

$$\phi(B)\Phi(B)[(1-B)^d(1-B^s)^D - \alpha]Z_t = \theta(B)\Theta(B)\epsilon_t. \quad (2.19)$$

2.5.6 Modelo auto-regressivo de médias móveis sazonal com variáveis exógenas (*SARIMAX*)

O modelo *SARIMAX* é uma regressão linear no formato semelhante ao modelo *SARIMA*, de forma que, *SARIMAX*(p,d,q,s) x (P,D,Q). O surgimento deste modelo se deu a uma adaptação do modelo *SARIMA* que inclui variáveis exógenas (X), ou seja variáveis externas auxiliam no entendimento e previsão da variável de interesse. Sua equação é dada por:

$$\phi(B)\Phi(B)(1-B)^D Z_t = \alpha + \mathbf{X}\boldsymbol{\beta} + \theta(B)\Theta(B^S)\epsilon_t \quad (2.20)$$

de forma que, **X** é a matriz de variáveis regressoras e **β** o vetor de parâmetros de estimação do modelo.

2.5.7 Função de autocorrelação (*FAC*)

Para Montgomery, Jennings e Kulahci (2015), em situações que a série temporal alcança a estacionariedade, duas observações W_t e W_{t+k} possui distribuição de probabilidade conjunta para qualquer tempo t e $t+k$. O intervalo que separa estas observações é chamado de *lag* (k).

Assim, a função de autocorrelação amostral (*FAC*), no lag de ordem k é definida como:

$$r_k = \frac{\sum_{t=k+1}^n (W_t - \bar{W})(W_{t+k} - \bar{W})}{\sum_{t=1}^n (W_t - \bar{W})^2}, \quad k = 1, 2, \dots \quad (2.21)$$

em que n é o número de observações que compõem a janela temporal. Além disso, se o número de observações for suficientemente grande $r_k \sim N(0, \frac{1}{n})$. Assim, conhecendo a distribuição de r_k pode-se avaliar a presença da correlação entre os *lags*. Para testar as hipóteses de autocorrelação, utiliza-se o teste de Ljung-Box o qual é definido pela seguinte expressão:

$$Q_k = n(n+2) \sum_{k=1}^m \frac{r_k^2}{n-k}, \quad (2.22)$$

no qual a hipótese nula testada é de que os resíduos não sejam autocorrelacionados e a hipótese alternativa é de que hajam evidências de correlação nos resíduos. A estatística de teste Q_k converge para uma distribuição χ^2 com q graus de liberdade (Cryer; Chan, 2008; Bueno, 2011).

2.5.8 Função de autocorrelação parcial (FACP)

De acordo com Shumway e Stoffer (2017) a função que mensura a autocorrelação entre W_t e W_{t-k} posteriormente a remoção das variáveis mediadoras, sendo estas $W_{t-1}, W_{t-2}, \dots, W_{t-k+1}$ é a função de autocorrelação parcial. Esta função, bem como a função de autocorrelação é comumente utilizada para a verificação da estacionariedade da série. Assim, as variáveis mediadora podem ser descritas por:

$$\begin{cases} W_t = \phi_{11}W_{t-1} + \epsilon_t \\ W_t = \phi_{11}W_{t-1} + \phi_{22}W_{t-2} + \epsilon_t \\ \vdots \\ W_t = \phi_{k1}W_{t-1} + \phi_{k2}W_{t-2} + \dots + \phi_{kk}W_{t-k} + \epsilon_t \end{cases} \quad (2.23)$$

Ademais, bem como a função demonstrada acima, para casos em que o número de observações seja suficientemente grande, encontra-se que $\phi \sim N(0, \frac{1}{n})$, com $k > p$ em que p é a ordem do processo AR .

2.6 Teste de estacionariedade de Dickey-Fuller

Em um modelo $Z_t = \rho Z_t + \epsilon_t$, em que ϵ_t é um ruído branco $\rho = 1$ a série, Z_t possui raiz unitária e não alcança a estacionariedade. Todavia, se $|\rho| < 1$ a série é estacionária, ou seja, não possui raiz unitária. Para Dickey e Fuller (1979), o processo que gera a série Z_t é definido por algum dos modelos apresentados a seguir:

1. $\Delta Z_t = \alpha + \beta_t + \lambda_3 Z_{t-1} + \epsilon_t$
2. $\Delta Z_t = \alpha + \lambda_2 Z_{t-1} + \epsilon_t$

$$3. \Delta Z_t = \lambda_1 Z_{t-1} + \epsilon_t$$

Em que $\lambda_i = \rho - 1 \forall_i = 1, 2, 3$ e α e β são invariantes passíveis de estimativa.

Assim, para avaliar a estacionariedade de uma série pelo teste de Dickey-Fuller as hipóteses a serem testadas serão apresentadas na 2.1. Sabe-se que, para que a série seja estacionária todas as hipóteses devem ser rejeitada.

Tabela 2.1 – Avaliação de raiz unitária utilizando teste Dickey-Fuller

Modelos	H_0	Regras de decisão
1	$\lambda_3 = 0$ $(\alpha, \beta, \lambda_3) = (0, 0, 0)$ $(\alpha, \beta, \lambda_3) = (\alpha, 0, 0)$	$\tau_3 > \text{valor crítico} \Rightarrow H_0$ não é rejeitada $\delta_2 < \text{valor crítico} \Rightarrow H_0$ não é rejeitada $\delta_3 < \text{valor crítico} \Rightarrow H_0$ não é rejeitada
2	$\lambda_2 = 0$ $(\alpha, \lambda_2) = (0, 0)$	$\tau_2 > \text{valor crítico} \Rightarrow H_0$ não é rejeitada $\delta_1 < \text{valor crítico} \Rightarrow H_0$ não é rejeitada
3	$\lambda_1 = 0$	$\tau_1 > \text{valor crítico} \Rightarrow H_0$ não é rejeitada

Fonte: Araújo et al. (2023).

2.7 Validação do ajuste

Como mencionado anteriormente, a modelagem de séries temporais possui três etapas, sendo uma delas a verificação da adequação do modelo ajustado à série temporal estudada. Segundo Burnham e Anderson (2004), é de extrema importância selecionar o melhor ajuste com base em princípios cientificamente comprovados. Assim, serão apresentados três critérios de informação: Akaike (AIC), Akaike Corrigido ($AICc$) e Bayesiano (BIC). Esses critérios, de acordo com Emiliano et al. (2010), podem ser comparados tanto em séries temporais quanto em outras modelagens estatísticas mais comuns, compondo a fase de validação desta etapa.

Segundo Lima, Castro e Cartaxo (2019), a outra fase da validação do ajuste é a análise dos resíduos, que será realizada por meio do teste de normalidade de Shapiro-Wilk, o qual avalia o comportamento dos resíduos comparando seu comportamento à distribuição normal, e do teste de Ljung-Box, que verifica a autocorrelação residual, de forma que, a autocorrelação residual em termos estatísticos deve ser equivalente a 0, conforme definido na Seção 2.5.7

- **AIC** : Akaike (1974) é contruído de forma assintótica por p , sabendo que, p é a quantidade de parâmetros que o modelo estima. Assim, seu critério de informação é definido por:

$$AIC = -2 \log L(\hat{\theta}) + 2(p). \quad (2.24)$$

- **BIC:** Schwarz (1978) propôs o Critério de Informação Bayesiano que é dado pela seguinte expressão:

$$BIC = -2 \log f(x_n | \theta) + p \log n. \quad (2.25)$$

- **AICc:** Proposto por Bozdogan (1987), o AICc é uma correção para o AIC, denotado por:

$$AICc = -2 \log L(\hat{\theta}) + 2(p) + 2 \frac{p(p+1)}{n-p-1}. \quad (2.26)$$

- **Shapiro-Wilk:** O teste de normalidade é um dos testes estatísticos mais abordados e aplicados em diversos tipos de análise. Como o próprio nome sugere, esse teste observa se os dados estudados se aproximam da distribuição Normal. Na literatura existem alguns testes de normalidade, neste trabalho aborda-se o mais comum deles, o qual foi proposto por Shapiro e Wilk (1965). A interpretação dessa estatística é simples, a hipótese nula declara que os dados seguem uma distribuição normal e a hipótese alternativa que não seguem este comportamento. Assim, para que a normalidade seja encontrada o p-valor deve ser maior que 0,05 que é o nível de significância adotado neste trabalho. A equação que gera essa estatística é dada por:

$$W = \frac{\left(\sum_{i=1}^n a_i X_{(i)}\right)^2}{\sum_{i=1}^n (X_i - \bar{X})^2}. \quad (2.27)$$

2.8 Avaliação da precisão do ajuste

Ainda sobre as etapas que elegem um modelo como adequado, tem-se a validação da precisão do ajuste. Para isso, utilizam-se as métricas de erro e os coeficientes que determinam o quanto o modelo ajustado é acurado. Assim, para Araújo et al. (2023) em uma série temporal constituída por T observações, as n últimas são retiradas para que as previsões para um futuro não observado sejam geradas, a partir da seguinte equação:

$$e_j = Z_j - \hat{Z}_j. \quad (2.28)$$

Logo, \hat{Z}_j representa o valor estimado para Z_j . Os erros apresentados a seguir são úteis para avaliar o ajuste do modelo proposto para previsão, considerando $k = T - n$. Além disso, o índice *U de Theil* proposto por Theil (1966) e o coeficiente de determinação R^2 são métricas relevantes para validar o desempenho dos modelos aplicados.

- **RMSE:** Erro quadrático médio

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=k+1}^M \epsilon_j^2}. \quad (2.29)$$

- **MAE**: Erro absoluto médio

$$MAE = \frac{1}{n} \sum_{j=k+1}^M |e_j|. \quad (2.30)$$

- **MAPE**: Erro percentual absoluto médio

$$MAPE = \frac{100}{n} \sum_{j=k+1}^M \left| \frac{\epsilon_j}{Z_j} \right|. \quad (2.31)$$

- **U de Theil**: Estatística de Theil

$$U = \frac{\sqrt{\frac{1}{n} \sum_{j=k+1}^M (\hat{Z}_j - Z_j)^2}}{\sqrt{\frac{1}{n} \sum_{j=k+1}^M Z_j^2}}. \quad (2.32)$$

- **R²**: Coeficiente de determinação

$$R^2 = 1 - \frac{\sum_{j=k+1}^M (Z_j - \hat{Z}_j)^2}{\sum_{j=k+1}^M (Z_j - \bar{Z})^2}. \quad (2.33)$$

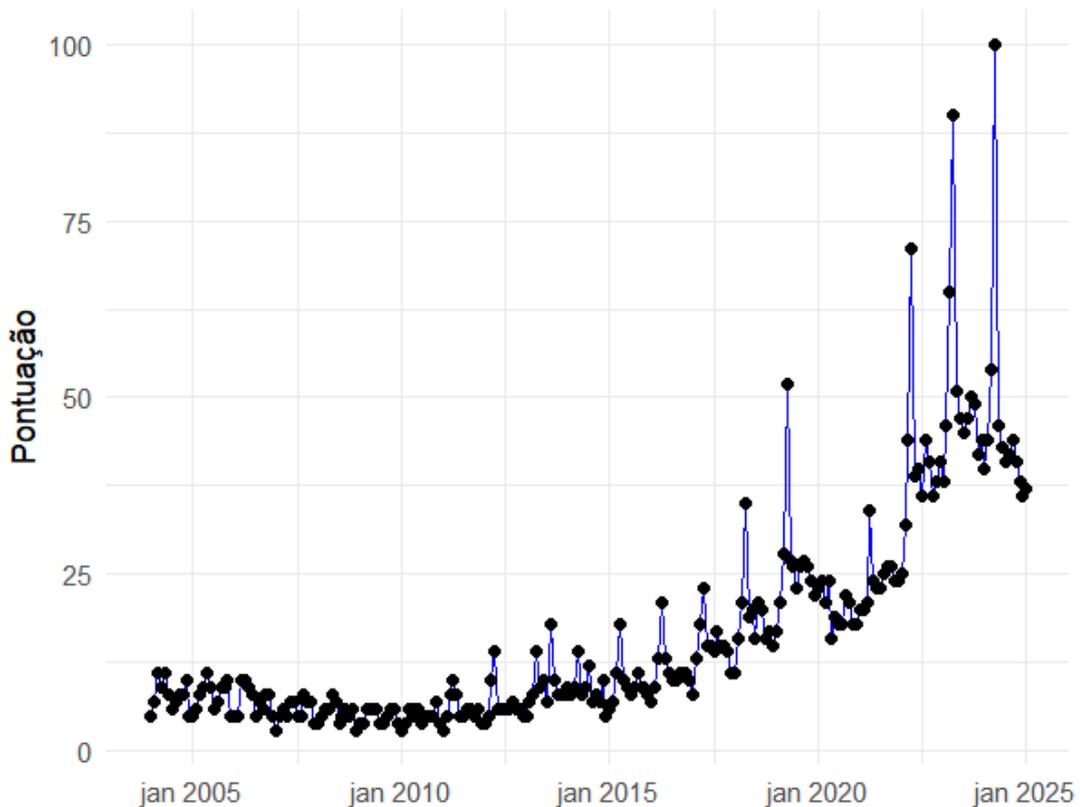
3 APLICAÇÃO

Nesta seção serão apresentados os resultados obtidos a partir de modelos de séries temporais aplicados a dados de buscas no Google pelo termo “autismo”, de janeiro de 2004 a janeiro de 2025 agrupados mensalmente no Brasil. Os dados estão disponíveis no *Google trends* no link <https://acesse.one/U2lxP> e toda a análise computacional foi realizada no R (Team, 2023) com auxílio dos pacotes *forecast*, *tseries*, *zoo*, *ggplot2*, *dplyr*, *prophet* e *lmtest*.

3.1 Análise Exploratória

Na Figura 3.1 é possível observar o comportamento das buscas pelo termo “autismo” no Google ao longo de mais de duas décadas. Nota-se um crescimento expressivo na pontuação de popularidade do termo. Além disso, comportamentos semelhantes são observados ao longo dos anos, atingindo escalas distintas

Figura 3.1 – Série temporal da busca pelo termo “autismo” no Google de janeiro de 2004 até janeiro de 2025.

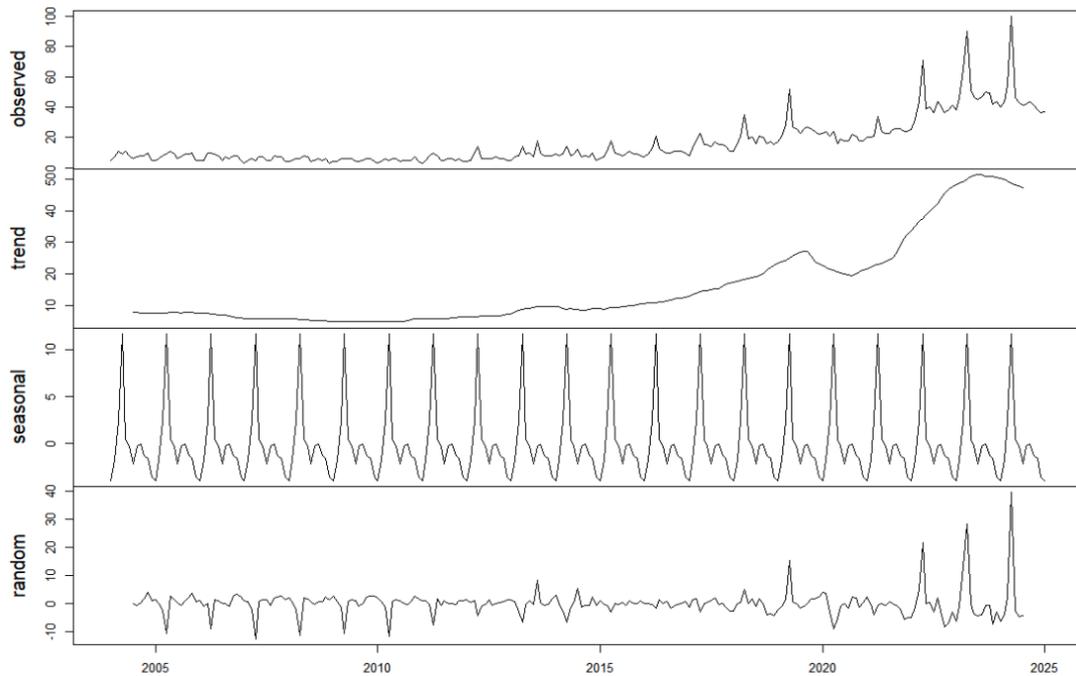


Fonte: Elaborada pela autora, (2025).

Na série apresentada acima, é nítida a grande variabilidade encontrada na janela de tempo investigada. Para contornar esse fenômeno utilizou-se a transformação logarítmica

na série em estudo. Este tipo de transformação é amplamente utilizada em séries que denotam heterocedasticidade, buscando que um comportamento homogêneo seja alcançado para um melhor ajuste.

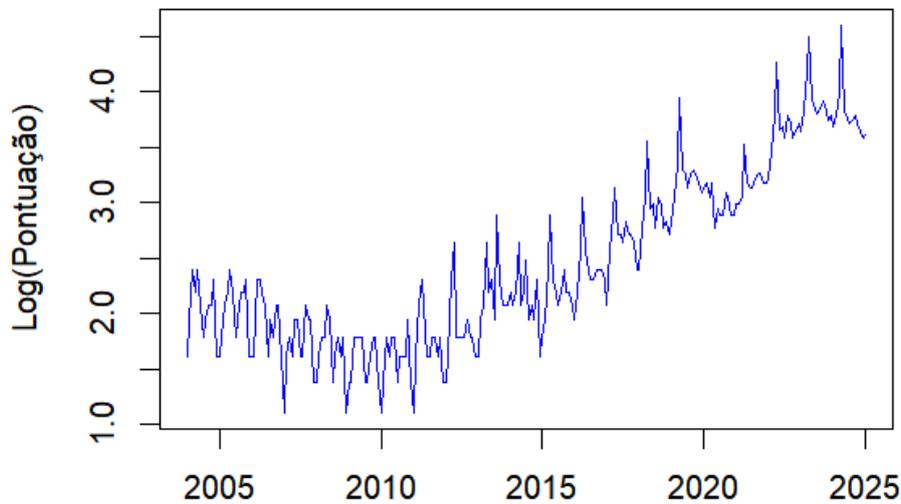
Figura 3.2 – Decomposição da série de busca por “autismo” em pontuação na escala original em componetes sazonais, tendência e aleatoriedade dos resíduos.



Fonte: Elaborada pela autora, (2025).

De acordo com a Figura 3.2 a série apresenta uma forte tendência que vem a ser característica de séries não estacionárias. Além disso a sazonalidade também é encontrada, sendo observada pelo comportamento repetitivo identificado no terceiro quadro. Para comprovação da não estacionariedade da série o teste de Dickey-Fuller foi aplicado o qual denotou a falta de estacionariedade com $p = 0,341$. A seguir será apresentada na Figura 3.3 a série temporal em escala logarítmica.

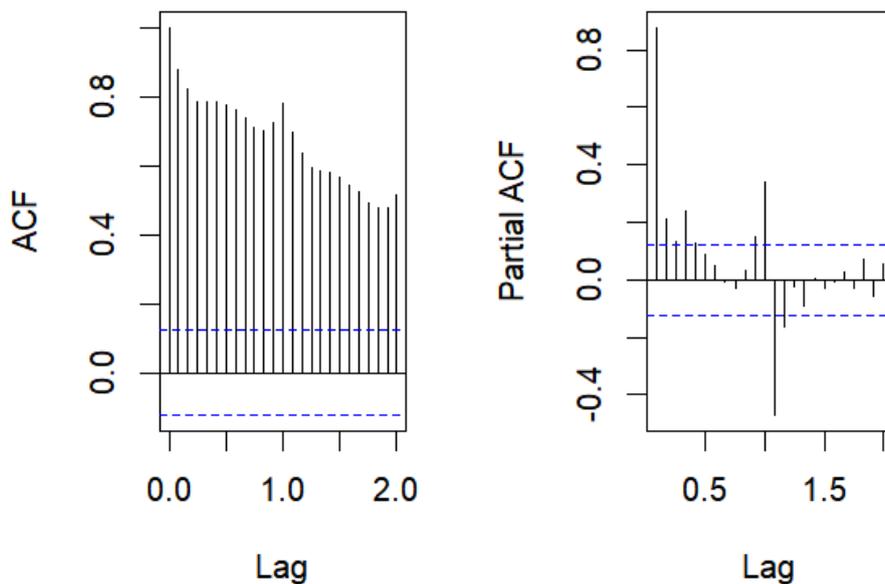
Figura 3.3 – Série temporal da busca pelo termo “autismo” no Google de janeiro de 2004 até janeiro de 2025 em escala logarítmica.



Fonte: Elaborada pela autora, (2024).

A série em escala logarítmica apresenta uma menor variabilidade, assim, ao longo da análise esta escala será adotada visando tornar a série mais proporcional, suavizar as oscilações existentes e melhorar performance dos modelos ajustados. As funções de autocorrelação amostral e autocorrelação parcial da série original são apresentadas na Figura 3.4. Esses resultados reiteram a hipótese de não estacionariedade da série, a qual foi comprovada pelo teste de Dickey-Fuller.

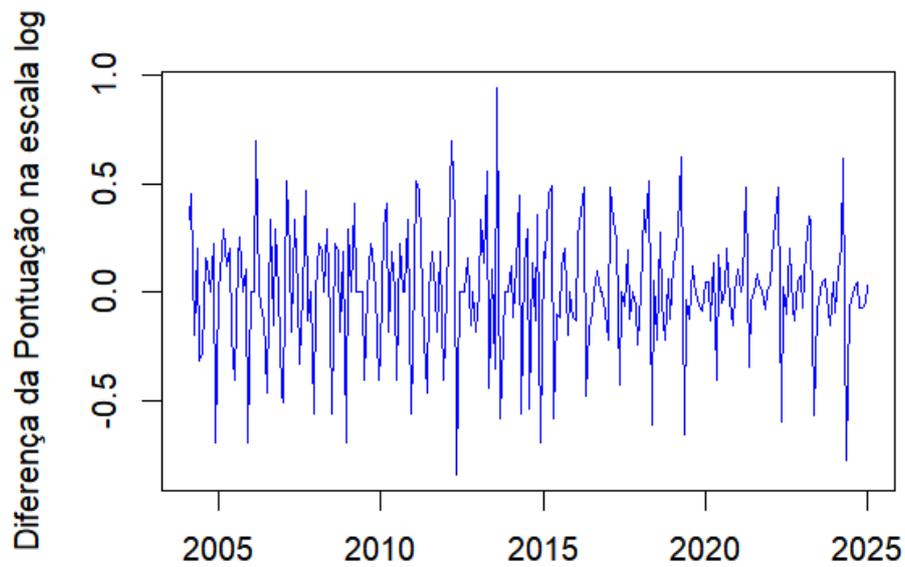
Figura 3.4 – Funções de autocorrelação amostral e parcial da série original.



Fonte: Elaborada pela autora, (2025).

Conforme a Figura 3.4 essa janela temporal apresenta uma forte tendência. Assim, mesmo após a aplicação da escala logarítmica, a série temporal não atingiu a estacionariedade com $p = 0,2539$. Portanto, foi necessária a aplicação de uma diferenciação. A série diferenciada é apresentada na Figura 3.5.

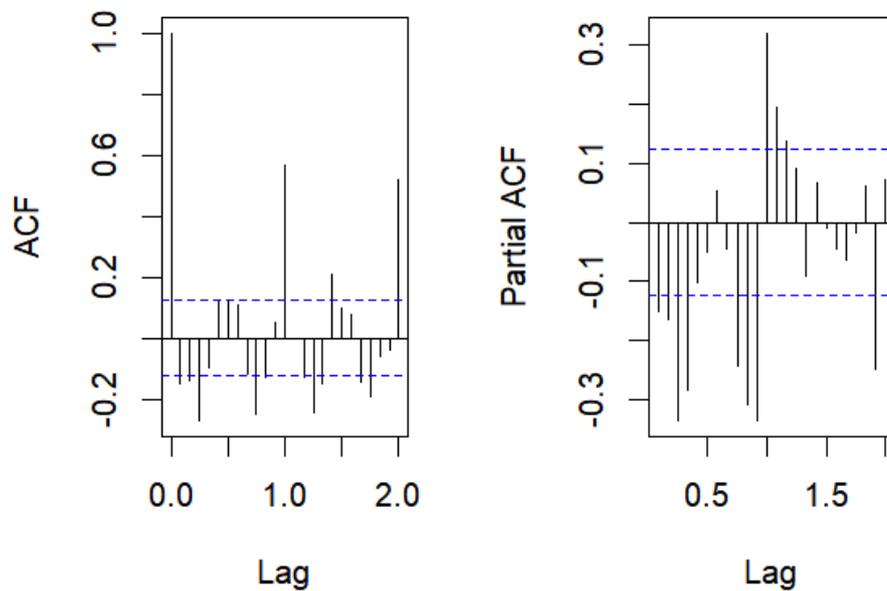
Figura 3.5 – Série temporal da busca pelo termo “autismo” no Google de janeiro de 2004 até janeiro de 2025 em escala logarítmica e diferenciada.



Fonte: Elaborada pela autora, (2025).

Após a diferenciação como apresentada na Figura 3.5 a série tornou-se estacionária com $p = 0,01$ no teste de Dickey-Fuller. Serão apresentadas as funções de função de autocorrelação (FAC) e função de autocorrelação parcial (FACP).

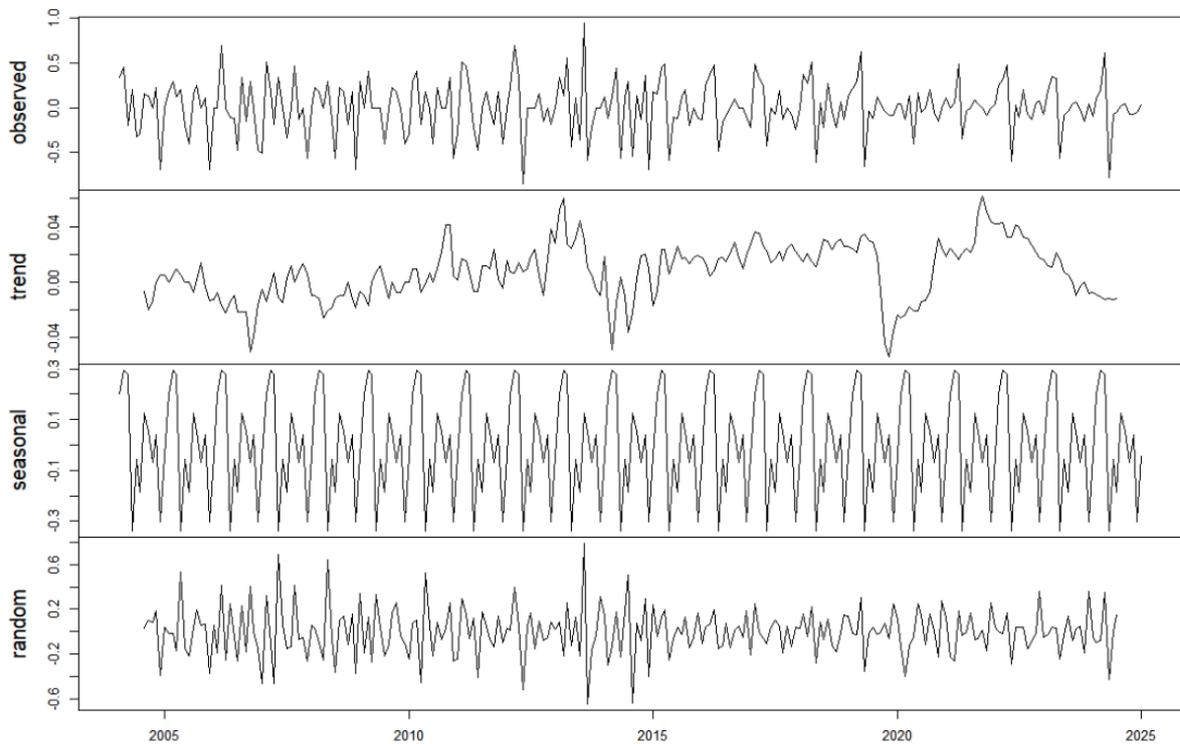
Figura 3.6 – Funções de autocorrelação amostral e parcial da série logarítimizada e diferenciada.



Fonte: Elaborada pela autora, (2025).

Os gráficos apresentados na Figura 3.6 denotam que existem picos de autocorrelação significativa os quais indicam sazonalidade. Além disso, foram avaliados os componentes de decomposição da série, resultados exibidos na Figura 3.7. A avaliação dos componentes que estruturam a série é de suma importância para a identificação do comportamento isolado de cada um deles, permitindo um maior foco em suas particularidades, além de contribuir para a escolha do modelo que melhor captura essas características.

Figura 3.7 – Decomposição da série de busca por “autismo” em pontuação na escala e diferenciada em componetes sazonais, tendência e aleatoriedade dos resíduos.



Fonte: Elaborada pela autora, (2025).

Logo, por meio da Figura 3.7 é possível observar que a tendência existente na série original foi contornada. Dessa forma, pode-se seguir com a aplicação dos modelos inerentes a análise de séries temporais.

3.2 Modelagem

As próximas aplicações referem-se à modelagem de séries temporais com os dados que representam a pontuação de popularidade do termo “autismo” no Brasil, no período de 2004 até janeiro de 2025. O foco está na compreensão dos padrões existentes e na previsão de valores futuros que irão compor a tendência de busca pelo termo nos próximos anos.

3.2.1 Modelagem *Box-Jenkins*

A modelagem de *Box-Jenkins* é amplamente utilizada para a análise de séries temporais que atendem à condição de estacionariedade. Segundo Box e Jenkins (1976), essa modelagem é conduzida em três etapas principais: identificação, estimação e verificação. Caso o modelo obtido não seja adequado, o processo retorna à fase de identificação (Werner; Ribeiro, 2003).

De acordo com Hyndman e Khandakar (2008), a melhor forma de realizar a identificação do modelo é por meio da função *auto.arima*, disponível no pacote *forecast* do

ambiente computacional R. Esse procedimento foi aplicado, e o modelo obtido foi um $ARIMA(0, 0, 1)(0, 1, 1)_{12}$.

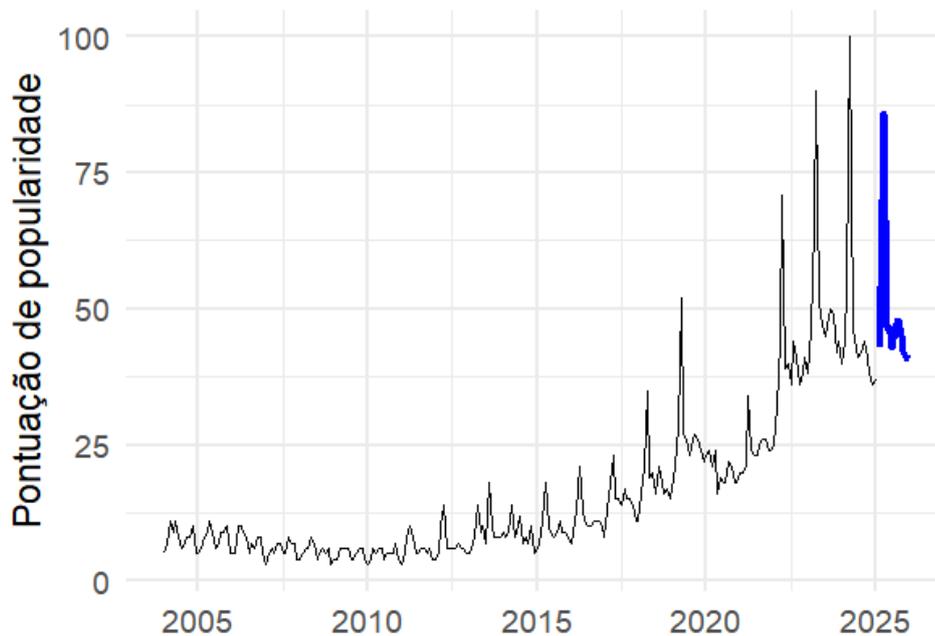
Tabela 3.1 – Resumo do Modelo $ARIMA$

Modelo	AIC	BIC	Erro de Previsão (RMSE)
$ARIMA(0,1,1)(0,1,1)_{12}$	-129,27	-118,83	0,1748

Fonte: Elaborada pela autora, (2025).

A partir do modelo exibido na Tabela 3.1, previsões para os 12 meses seguintes foram realizadas e serão apresentadas na Figura 3.14

Figura 3.8 – Previsão do modelo $ARIMA(0, 0, 1)(0, 1, 1)_{12}$ 12 passos à frente para a pontuação de popularidade do termo “autismo”.



Fonte: Elaborada pela autora, (2025).

A série apresentada na Figura 3.14 com adição das previsões permite a avaliação do comportamento futuro do interesse populacional pelo termo, os valores para a pontuação de popularidade de cada mês estão exibidos na Tabela 3.2.

Tabela 3.2 – Previsão do modelo $ARIMA(0, 0, 1)(0, 1, 1)_{12}$ 12 passos à frente para a pontuação de popularidade do termo “autismo”.

Mês	Previsão
Fevereiro 2025	42,9865
Março 2025	54,0329
Abril 2025	85,9171
Mai 2025	47,0956
Junho 2025	45,6510
Julho 2025	42,8650
Agosto 2025	46,3186
Setembro 2025	48,0008
Outubro 2025	45,0628
Novembro 2025	42,1270
Dezembro 2025	41,2216
Janeiro 2025	40,8455

Fonte: Elaborada pela autora, (2025).

Com base nessas previsões apresentadas na Tabela 3.2, observa-se que o termo continuará com alta popularidade, especialmente no mês de abril. Esse aumento pode ser atribuído ao Dia Mundial de Conscientização do Autismo, celebrado em 2 de abril.

O erro de previsão associado ao modelo ajustado é baixo, e o teste de Ljung-Box não evidencia padrões nos resíduos, com $p > 0,05$, o que indica um bom ajuste. No entanto, o modelo não atende ao pressuposto de normalidade, conforme indicado pelo teste de Shapiro-Wilk, com $p < 0,05$. Portanto, outras modelagens podem ser mais adequadas.

3.2.2 Modelagem Holt-Winters

No acompanhamento temporal da popularidade do termo “autismo” no Google, observa-se a presença de componentes de tendência e sazonalidade. Para lidar com essa característica, utilizam-se os algoritmos de suavização exponencial aditivo e multiplicativo de *Holt-Winters* que para Caiado (2006) é apropriado para séries que apresentam tendência linear e sazonalidade. Este método é amplamente aplicado em termos de séries temporais por se tratar de um método tradicional que de forma geral apresenta resultados satisfatórios (Hochleitner, 2021).

Em estudos como o de Raggi (2023), a suavização exponencial de *Holt-Winters* apresentou os melhores resultados para previsão, exibindo os menores erros em todas as métricas validadas. Dentre os outros modelos avaliados, estavam o *SARIMA*, *ARIMA*, *TBATS* e o algoritmo *Prophet*. Além disso, para preservar a escala original da série, a

modelagem foi realizada na escala logarítmica, e as previsões exponenciadas.

Outro aspecto a ser considerado é a escolha dos parâmetros α , β e γ , responsáveis pela suavização, os quais devem ser ajustados de forma que venha a minimizar o valor do Erro Quadrático Médio (*RMSE*) de um passo à frente. Na Tabela 3.3 serão apresentados os valores otimizados para cada constante de alisamento para os modelos multiplicativos e aditivos.

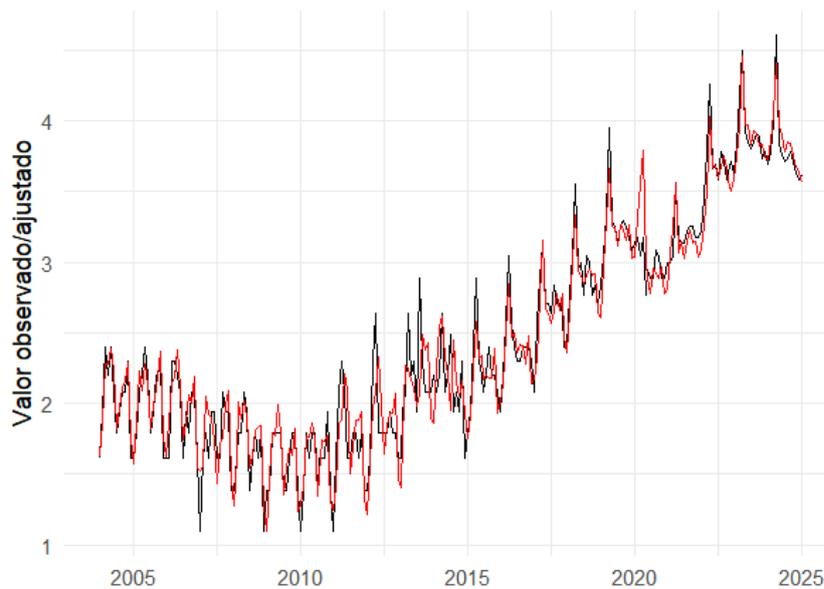
Tabela 3.3 – Valores das constantes de suavização otimizados para os modelos Aditivo e Multiplicativo na suavização exponencial tripla para a série de popularidade do termo “autismo” no Google.

Modelo	α (Nível)	β (Tendência)	γ (Sazonalidade)
HW (Aditivo)	0,3555	0,0001	0,2412
HW(Multiplicativo)	0,2862	0,0001	0,2697

Fonte: Elaborada pela autora, (2025).

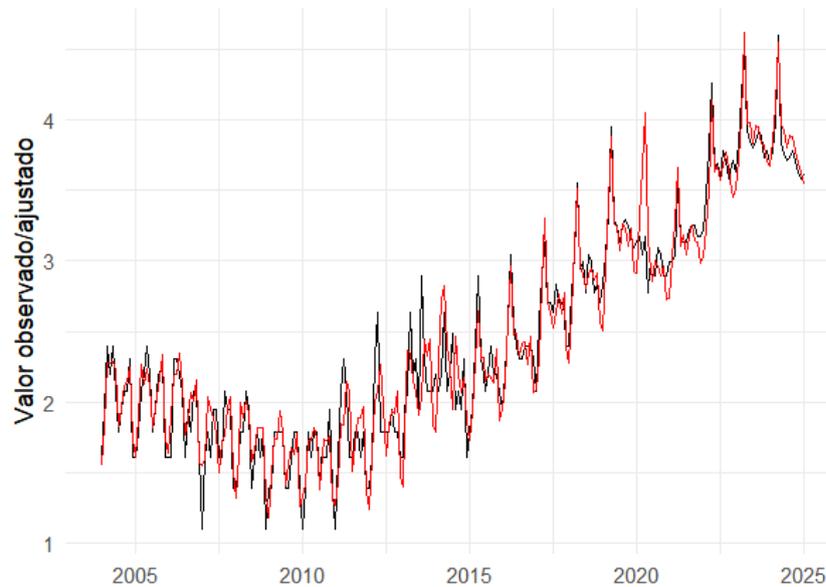
Dessa forma, foram ajustados os modelos de *Holt-Winters* nas versões Aditiva e Multiplicativa. As Figuras 3.9 e 3.10 apresentam os respectivos resultados dos ajustes, evidenciando o desempenho de cada abordagem na modelagem da série temporal para que posteriormente o melhor modelo dentro desta metodologia seja escolhido.

Figura 3.9 – Ajuste do modelo de *Holt-Winters* Aditivo um passo à frente (linha vermelha) para a série em escala log (linha preta).



Fonte: Elaborada pela autora, (2025).

Figura 3.10 – Ajuste do modelo de Holt-Winters Multiplicativo um passo à frente (linha vermelha) para a série em escala log (linha preta).



Fonte: Elaborada pela autora, (2025).

Ao observar a qualidade do ajuste nos modelos de *Holt-Winters* às previsões foram realizadas com 12 passos à frente. Os resultados serão apresentados na Tabela 3.4 a seguir.

Tabela 3.4 – Previsões Mensais para os Modelos Aditivo e Multiplicativo de *Holt-Winters*

Mês	Aditivo	Multiplicativo
Fevereiro 2025	41,4505	40,9540
Março 2025	51,7706	52,6662
Abril 2025	80,4884	88,3863
Mai 2025	44,1815	46,5189
Junho 2025	43,7934	46,5189
Julho 2025	41,5522	43,7586
Agosto 2025	45,2654	48,2729
Setembro 2025	47,1494	50,9082
Outubro 2025	44,1442	47,4810
Novembro 2025	41,5867	44,4758
Dezembro 2025	40,4306	42,4876
Janeiro 2026	39,3220	40,7368

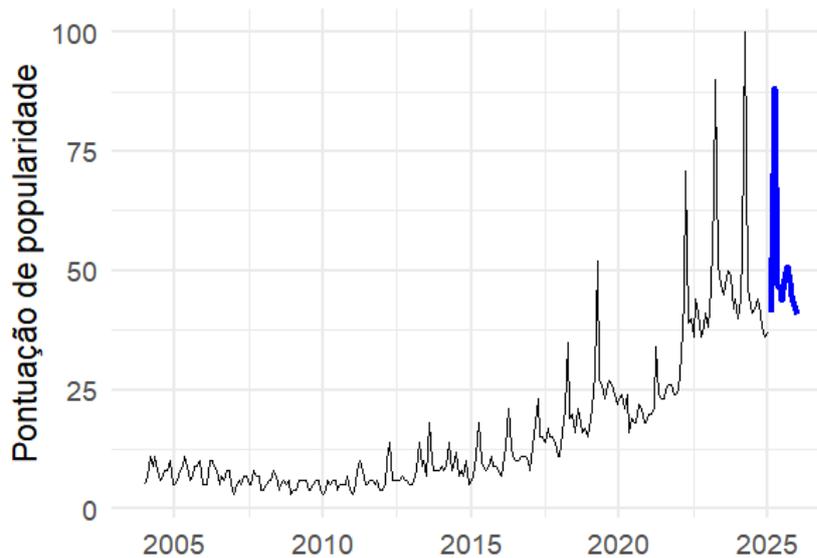
Fonte: Elaborada pela autora, (2025).

Observa-se que a previsão para abril de 2025 indica o maior nível de popularidade no período analisado. Esse pico no mês de abril teve início em 2010, possivelmente influenciado pela instituição do Dia Mundial de Conscientização sobre o Autismo pela ONU em 2007. Nos anos seguintes, campanhas de conscientização e mobilizações contribuíram

para o aumento do interesse. As previsões para os meses subsequentes situam-se, em sua maioria, entre 40 e 50 pontos em ambos os modelos, cujos valores são semelhantes.

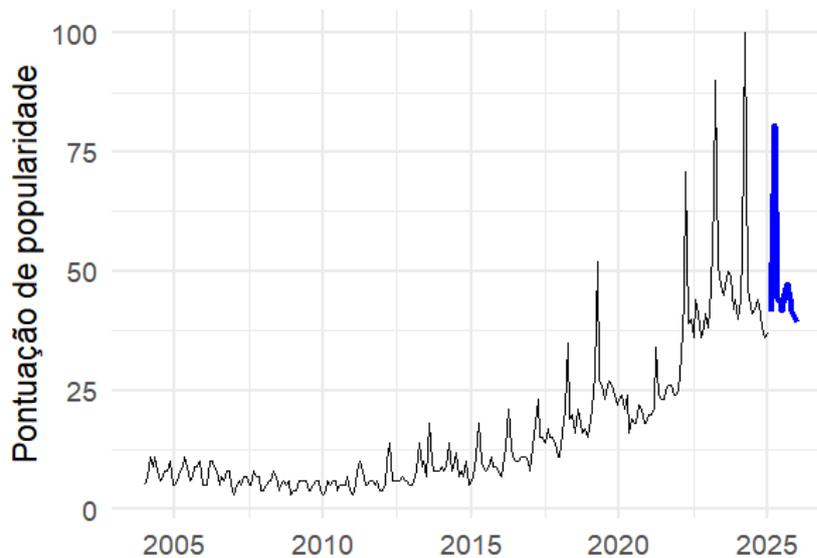
Na Figuras 3.11 e 3.12 pode-se observar a série registrada e suas previsões em 12 passos à frente para o modelo de alisamento exponencial Aditivo e Multiplicativo, assim as previsões se estendem até janeiro de 2026.

Figura 3.11 – Previsão de *Holt-Winters* 12 passos à frente no modelo Multiplicativo.



Fonte: Elaborada pela autora, (2025).

Figura 3.12 – Previsão de *Holt-Winters* 12 passos à frente no modelo Aditivo.



Fonte: Elaborada pela autora, (2025).

As previsões para os dois modelos são visualmente semelhantes. Para descobrir qual modelo adequa-se melhor aos dados, as métricas de erros e critérios de informações devem ser avaliadas. A seguir nas Tabelas 3.5 e 3.6 tem-se os valores das métricas e critérios de informação, respectivamente.

Tabela 3.5 – Métricas de erro dos Modelos Aditivo e Multiplicativo.

Métrica	Aditivo	Multiplicativo
<i>RMSE</i>	0,1739	0,1853
<i>MAE</i>	0,1275	0,1327
<i>MAPE</i>	5,8974	6,0692
<i>U de Theil</i>	0,3758	0,3459

Fonte: Elaborada pela autora, (2025).

Tabela 3.6 – Critérios de Informação para os Modelos Aditivo e Multiplicativo

Critério	Aditivo	Multiplicativo
<i>AIC</i>	548,9299	616,8725
<i>AICc</i>	551,5341	619,4767
<i>BIC</i>	608,9975	676,9401

Fonte: Elaborada pela autora, (2025).

Ambos os modelos foram combinados com constantes de suavização que atenuassem o erro, fornecendo valores relativamente baixos. Contudo, o modelo Aditivo é ligeiramente melhor ao observar as métricas de erro e os critérios de informação *AIC*, *BIC* e *AICc*. Para a maioria das métricas de erro, o modelo Aditivo de *Holt-Winters* revelou-se o mais acurado, utilizando as constantes de suavização otimizadas. Nos dois modelos, os erros são baixos, e a estatística que mensura o *U de Theil* apresenta valores próximos entre si e consideravelmente menores que 1. Assim, pode-se considerar que a modelagem de *Holt-Winters* foi adequada para este problema. Este resultado, corrobora com o estudo de Tibulo e Tibulo (2014), que revela o modelo aditivo de *Holt-Winters* como o melhor ajuste quando comparado aos métodos baseados na metodologia *Box-Jenkins*.

Em contrapartida, outros estudos indicam que essa metodologia pode não ser a melhor quando comparada a outros métodos ou avaliando suas próprias métricas. No estudo realizado por Samohyl, Rocha e Mattos (2001), o modelo de *Holt-Winters* apresentou uma estatística *U de Theil* > 1 , o que para Makridakis, Wheelwright e Hyndman (1998) indica que não há motivos para apostar nesta técnica. Dessa forma, a suavização exponencial tripla não foi capaz de detectar a estrutura dos dados corretamente. Além disso, a análise proposta por Angelo et al. (2011) aponta que a modelagem de *Holt-Winters* é menos

eficiente do que o modelo *ARIMA*, que faz parte da modelagem tradicional, e às Redes Neurais, que agregam aos modelos baseados em inteligência artificial.

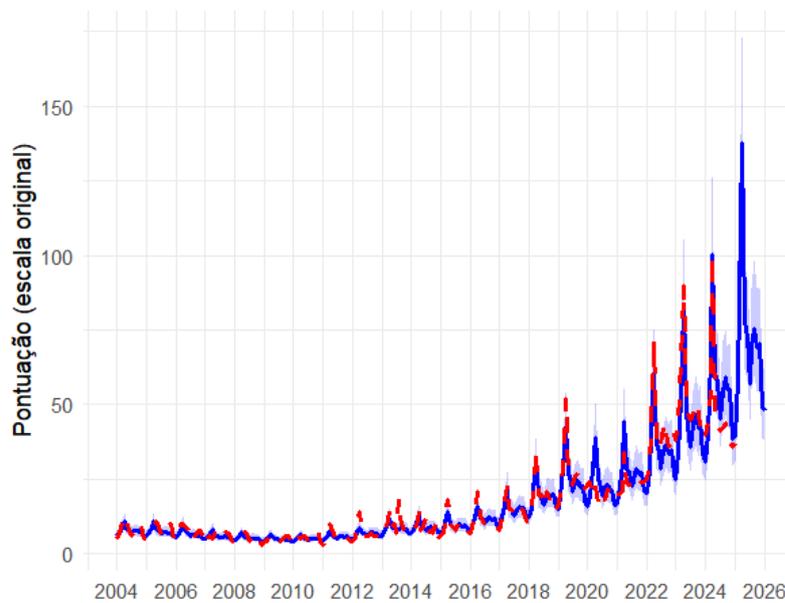
3.2.3 Algoritmo Prophet

Como mencionado anteriormente, a série histórica de popularidade do termo “autismo” apresenta uma tendência bem definida e sazonalidade forte e evidente. O *Prophet*, por sua vez, é um algoritmo simples que para Satrio et al. (2021), lida especialmente bem com sazonalidade e mudanças de tendência em longos períodos, sem perder a robustez. Dessa forma, o *Prophet* se mostra uma alternativa viável para modelar e prever o comportamento do interesse da população pelo autismo no Brasil nos próximos meses.

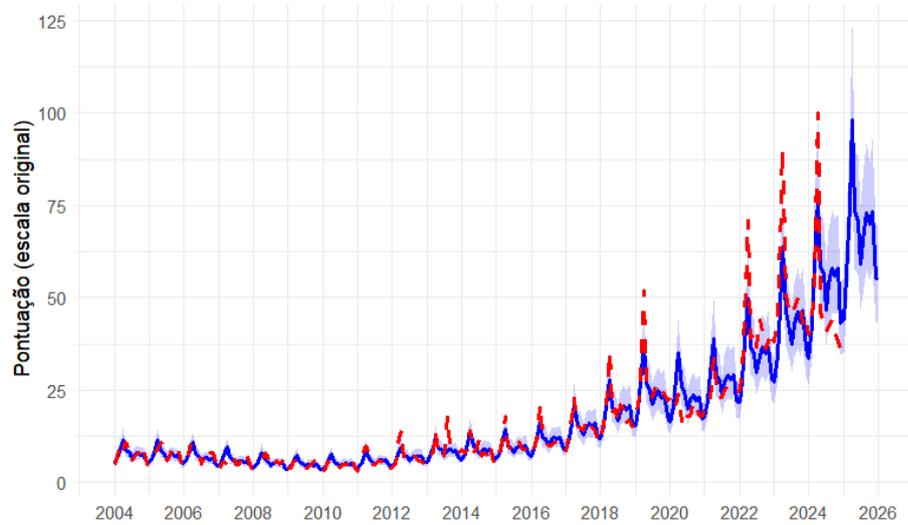
Sabendo que a série histórica possui alta variabilidade toda a modelagem foi realizada em escala logarítmica para que a flutuação seja reduzida.

O modelo *Prophet* foi ajustado de duas formas distintas, assim como o modelo de *Holt-Winters*, que foi aplicado nas versões aditiva e multiplicativa.

Figura 3.13 – Ajuste e previsão do modelo *Prophet* Multiplicativo.



Fonte: Elaborada pela autora, (2025).

Figura 3.14 – Ajuste e previsão do modelo *Prophet* Aditivo

Fonte: Elaborada pela autora, (2025).

É possível observar que as Figuras 3.13 e 3.14, apresentam indícios de que o modelo Multiplicativo captura melhor a estrutura dos dados. Contudo, para avaliar o melhor modelo serão utilizadas as métricas de erro, as quais serão exibidas na Tabela 3.8. Inicialmente as previsões para os meses de 2025 serão exibidas na Tabela 3.7.

Tabela 3.7 – Previsões mensais para os Modelos Aditivo e Multiplicativo *Prophet*

Mês	Aditivo	Multiplicativo
Fevereiro 2025	55,8207	58,3347
Março 2025	75,2149	84,8085
Abril 2025	98,1369	137,6036
Mai 2025	73,5354	77,5436
Junho 2025	70,3874	72,6159
Julho 2025	58,9401	56,9401
Agosto 2025	68,1183	69,5113
Setembro 2025	73,1145	75,6603
Outubro 2025	69,9167	68,9992
Novembro 2025	73,3441	70,5009
Dezembro 2025	55,4596	49,2304
Janeiro 2026	54,7492	48,0727

Fonte: Elaborada pela autora, (2025).

Ao observar os resultados de previsão e comparar com os modelos já ajustados observa-se que os valores são maiores que o previsto por outras modelagens apresentadas anteri-

amente, indicando possibilidade de superestimação. Contudo, o mês de Abril, bem como nas outras aborgadens apresentam uma maior estimação.

Tabela 3.8 – Métricas de erro e coeficiente de determinação dos Modelos Aditivo e Multiplicativo para o algoritmo *Prophet*.

Métrica	Aditivo	Multiplicativo
<i>RMSE</i>	0,1803	0,1891
<i>MAE</i>	0,1407	0,15
<i>MAPE</i>	6,1185	6,8339
R^2	0,9472	0,9419

Fonte: Elaborada pela autora, (2025).

As métricas de erro apresentadas, bem como o coeficiente de determinação, são próximas para os dois modelos. Todavia, o modelo aditivo é ligeiramente melhor que o modelo Multiplicativo, assim como na modelagem de *Holt-Winters*. Na análise de Almeida Corrêa e Vieira (2022), o modelo *Prophet* foi comparado ao modelo *ARIMA*, no qual o *Prophet* apresentou bons resultados, enquanto o modelo *ARIMA* não foi capaz de capturar o comportamento da série para realizar uma boa previsão.

Assim, o *Prophet* pode ser considerado uma boa abordagem para a análise de séries temporais. No estudo de Barreto, Bernardini e Oliveira (2024), foram comparadas as formas Aditiva e Multiplicativa do modelo *Prophet*, e os resultados, assim como nesse trabalho, foram próximos e satisfatórios para ambas as modelagens. Por outro lado, o modelo Multiplicativo se mostrou mais eficiente. De forma geral, o estudo mencionado acima destaca o *Prophet* como um algoritmo de fácil interpretação e usabilidade.

4 CONCLUSÃO

Neste estudo avaliou-se a popularidade do termo “autismo” no Brasil dentro da janela temporal de janeiro de 2004 a janeiro de 2025. Para a realização do ajuste e, posteriormente, da predição da pontuação de popularidade futura, foram utilizadas as modelagens *Box-Jenkins* e *Holt-Winters*, nas formas Aditiva e Multiplicativa, bem como o *Prophet*, também considerando as duas modalidades que compõem a modelagem *HW*. Em suma, o modelo escolhido via *Box-Jenkins* foi um $ARIMA(0, 1, 1), (0, 1, 1)_{12}$, o qual foi diagnosticado como o melhor dessa abordagem. Todavia, essa modelagem não atendeu a todos os pressupostos que garantem um melhor desempenho. Em contrapartida, as modelagens de *Holt-Winters* e *Prophet* foram satisfatórias nas modalidades abordadas. Os resultados apresentados foram semelhantes, mas o modelo de *Holt-Winters* aditivo foi o que demonstrou a melhor precisão ao comparar-se o *RMSE*.

Ademais, observa-se que o interesse pelo termo autismo é crescente no Brasil, e as previsões geradas sugerem que, nos próximos meses deste ano, a tendência de crescimento continuará presente. Assim, espera-se que a realidade de inclusão e compreensão também apresente uma tendência positiva, visto que, para compreender, é preciso entender; conhecer a comorbidade é essencial. Para estudos futuros, esta análise pode ser expandida para além das fronteiras do Brasil ou até mesmo deste continente, para que o comportamento mundial seja entendido a fim de identificar particularidades e distinções que possam existir entre as nações em relação ao interesse por esse termo, que cresce em entendimento de forma recorrente. Além disso, outras modelagens atreladas ao aprendizado de máquina podem ser abordadas, a fim de serem comparadas aos modelos que compõem a modelagem tradicional aqui discutidos.

REFERÊNCIAS

- AKAIKE, Hirotugu. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, IEEE, v. 19, n. 6, p. 716–723, 1974. DOI: <10.1109/TAC.1974.1100705>.
- ALMEIDA CORRÊA, Adriano de; VIEIRA, Sylvio André Garcia. Pré Processamento De Dados Para Análise de Séries Temporais na Previsão de Valores de Criptomoedas, Utilizando ARIMA e PROPHET, 2022. Acesso em: 25 mar. 2025. Disponível em: <https://www.tfgonline.lapinf.ufn.edu.br/media/midias/Adriano_Correa.pdf>.
- ANGELO, Claudio Felisoni de; ZWICKER, Ronaldo; FOUTO, Nuno Manoel Martins Dias; LUPPE, Marcos Roberto. Séries temporais e redes neurais: uma análise comparativa de técnicas na previsão de vendas do varejo brasileiro. **BBR-Brazilian Business Review**, FUCAPE Business School, v. 8, n. 2, p. 1–21, 2011.
- ARAÚJO, Eduardo Gomes de; JÚNIOR, Silvio Fernando Alves Xavier; BARBOSA, Nyedja Fialho Moraes; OLIVEIRA, Tiago Almeida de. Modelagem e previsão de modelos de séries temporais do consumo de energia elétrica na Região Nordeste do Brasil. **Sigmae**, v. 12, n. 1, p. 10–28, 2023. Disponível em: <<https://publicacoes.unifal-mg.edu.br/revistas/index.php/sigmae/article/view/2008>>.
- BARRETO, Tobias; BERNARDINI, Flavia; OLIVEIRA, Daniel de. Previsão de demanda de água potável em Cidades Inteligentes por meio do algoritmo de modelagem de séries temporais PROPHET. In: SBC. WORKSHOP de Computação Aplicada em Governo Eletrônico (WCGE). [S.l.: s.n.], 2024. P. 157–168.
- BORGES, Luciana Alves de Freitas Soares; YAMAMOTO, Isabel Rosa; LOPES, Ana Paula da Silva; MELO, André Alexandre; SIQUEIRA, Beatriz Otoni de; PEREIRA, Beatriz Leite; SOUZA, Carolina Miranda Tozetti de. Aumento nos casos de Transtorno do Espectro Autista em crianças: fatores e implicações. **Brazilian Journal of Implantology and Health Sciences**, v. 6, n. 11, p. 3697–3705, 2024. DOI: <10.36557/2674-8169.2024v6n11p3697-3705>. Disponível em: <<https://doi.org/10.36557/2674-8169.2024v6n11p3697-3705>>.
- BOX, George E. P.; JENKINS, Gwilym M. **Time Series Analysis: Forecasting and Control**. Revised. San Francisco: Holden-Day, 1976. P. 575. ISBN 978-0-8162-1104-3.
- BOZDOGAN, Hamparsum. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. **Psychometrika**, v. 52, p. 345–370, 1987.
- BUENO, Ronaldo L. S. **Econometria de Séries Temporais**. São Paulo: Cengage Learning, 2011.

- BURNHAM, Kenneth P.; ANDERSON, David R. Multimodel inference: understanding AIC and BIC in model selection. **Sociological Methods and Research**, Beverly Hills, v. 33, n. 2, p. 261–304, mai. 2004.
- CAIADO, R. G. G. **Modelagem de séries temporais para previsão da demanda de energia elétrica**. 2006. Dissertação de Mestrado – Universidade de São Paulo, São Paulo, SP.
- CONSTANZO, V.; CHERICONI, N.; AMENDOLA, F. A.; CASULA, L.; MURATORI, F.; SCATTONI, M. L. Early detection of autism spectrum disorders: from retrospective home video studies to prospective 'high risk' sibling studies. **Neuroscience Biobehavioral Reviews**, v. 55, p. 627–635, 2015.
- CRYER, Jonathan D.; CHAN, Kung-Sik. **Time Series Analysis: With Applications in R**. 2. ed. New York: Springer, 2008. DOI: {10.1007/978-0-387-75959-3}.
- DICKEY, David A.; FULLER, Wayne A. Distribution of the estimators for autoregressive time series with a unit root. **Journal of the American Statistical Association**, Taylor & Francis, v. 74, 366a, p. 427–431, 1979.
- EMILIANO, Paulo C; VEIGA, Elayne P; VIVANCO, MJ; MENEZES, Fortunato S. Critérios de informação de akaike versus bayesiano: análise comparativa. **19º Simpósio Nacional de Probabilidade e Estatística**, 2010.
- EYAL, G. **The autism matrix**. Cambridge: Polity Press, 2010.
- GRINKER, R. R. **Unstrange minds: remapping the world of autism**. Philadelphia: Basic Books, 2008.
- HOCHLEITNER, Augusto. **Análise comparativa entre métodos estatísticos tradicionais e baseados em inteligência artificial no problema de previsão de demanda**. Florianópolis, SC: [s.n.], 2021. Orientador: Ricardo Faria Giglio. Disponível em: {<https://repositorio.ufsc.br/handle/123456789/228199>}.
- HOLT, C. C. Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages. **International Journal of Forecasting**, v. 20, n. 1, p. 5–13, 1957.
- HYNDMAN, Rob J.; KHANDAKAR, Yeasmin. Automatic Time Series Forecasting: The forecast Package for R. **Journal of Statistical Software**, v. 27, n. 3, p. 1–22, 2008.
- LIMA, J. E. C.; CASTRO, L. F. de; CARTAXO, G. A. A. Aplicação do modelo SARIMA na previsão de demanda no setor calçadista / Application of the SARIMA model in the forecast for demand in the footwear sector. **ID on line Revista de Psicologia**, v. 13, n. 46, p. 892–913, 2019.
- LORD, Catherine; ELSABBAGH, Mayada; BAIRD, Gillian; VEENSTRA-VANDERWEELE VEENSTRA-VANDERWEELE, Jeremy. Autism spectrum disorder. **The Lancet**, v. 392, n. 10146, p. 508–520, 2018.

MAKRIDAKIS, Spyros; WHEELWRIGHT, Steven C.; HYNDMAN, Rob J. *Forecasting: Methods and Applications*. **Wiley**, John Wiley & Sons, Inc., 1998.

MARTÍNEZ, Raydonal Ospina; ZAMPROGNO, Bartolomeu. Comparação de algumas técnicas de previsão em análise de séries temporais. **Revista Colombiana de Estadística**, v. 26, n. 2, p. 129–157, 2003.

MONTGOMERY, Douglas C.; JENNINGS, Cheryl L.; KULAHCI, Murat. **Introduction to Time Series Analysis and Forecasting**. 2nd. Hoboken, NJ: John Wiley & Sons, 2015.

MORETTIN, P. A.; TOLOI, C. M. **Análise de Séries Temporais**. 2^a ed. revista e ampliada. [S.l.]: ABE–Projeto Fisher, Editora Edgar Blücher, 2006.

_____. _____. 2. ed. São Paulo: Editora Blücher, 2018.

RAGGI, Lucas Ribeiro. **Previsão de empregos na construção civil em Alagoas utilizando séries temporais**. Maceió, AL: [s.n.], 2023. Disponível em: [j{https://www.repositorio.ufal.br/handle/123456789/12933}i](https://www.repositorio.ufal.br/handle/123456789/12933).

RAMOS, T. C.; FUJITA, A. **Identificação de alterações em conectividades funcionais córtico-cerebelares no transtorno do espectro autista**. 2017. Dissertação (Mestrado em Ciência da Computação) – Universidade de São Paulo, São Paulo.

RIOS, Clarice; ORTEGA, Francisco; ZORZANELLI, Rafaela; NASCIMENTO, Leonardo Fernandes. Da invisibilidade à epidemia: a construção narrativa do autismo na mídia impressa brasileira. **Comunicação, Mídia e Consumo**, v. 19, n. 53, p. 325–335, 2015. DOI: {10.1590/1807-57622014.0146}.

ROBERTSON, C. E.; BARON-COHEN, S. Sensory perception in autism. **Nature Reviews Neuroscience**, v. 18, p. 671–684, 2017. DOI: {10.1038/nrn.2017.112}.

SAMOHYL, Robert Wayne; ROCHA, Rubson; MATTOS, Viviane Leite Dias de. Utilização do método de Holt-Winters para previsão do leite entregue às indústrias catarinenses, 2001.

SATRIO, Christophorus Benedicto Aditya; DARMAWAN, William; NADIA, Bellatasya Unrica; HANAFIAH, Novita. Time series analysis and forecasting of coronavirus disease in Indonesia using ARIMA model and PROPHET. **Procedia Computer Science**, v. 179, p. 524–532, 2021. 5th International Conference on Computer Science and Computational Intelligence 2020. ISSN 1877-0509. DOI: {10.1016/j.procs.2021.01.036}. Disponível em: [j{https://www.sciencedirect.com/science/article/pii/S1877050921000417}i](https://www.sciencedirect.com/science/article/pii/S1877050921000417).

SCHWARZ, Gideon. Estimating the dimension of a model. **Annals of Statistics**, v. 6, n. 2, p. 461–464, mar. 1978.

SHAPIRO, Samuel Sanford; WILK, Martin B. An Analysis of Variance Test for Normality (Complete Samples). **Biometrika**, Oxford University Press, v. 52, n. 3-4, p. 591–611, 1965. DOI: [10.1093/biomet/52.3-4.591](https://doi.org/10.1093/biomet/52.3-4.591).

SHUMWAY, Robert H.; STOFFER, David S. **Time Series Analysis Using the R Statistical Package**. [S.l.]: Free Dog Publishing, 2017.

SOUZA, A. L. de; SAMOHYL, M. F.; MIRANDA, J. A. C. A previsão por suavização ou alisamento exponencial. **Revista Brasileira de Estatística**, v. 71, n. 3, p. 45–56, 2009.

TAYLOR, S. J.; LETHAM, B. Forecasting at scale. **The American Statistician**, v. 72, n. 1, p. 37–45, 2018.

TEAM, R Core. **R: A language and environment for statistical computing**. [S.l.], 2023. R Foundation for Statistical Computing. Disponível em: <https://www.R-project.org/>.

THEIL, Henri. **Applied economic forecasting**. Amsterdam: North-Holland Publishing Company, 1966.

TIBULO, Cleiton; TIBULO, Vaneza De Carli. Previsão do preço do milho, através de séries temporais. **Scientia Plena**, v. 10, n. 10, 2014.

WERNER, Liane; RIBEIRO, José Luis Duarte. Previsão de demanda: uma aplicação dos modelos Box-Jenkins na área de assistência técnica de computadores pessoais. **Gestão & Produção**, SciELO Brasil, v. 10, p. 47–67, 2003.

WINTER, R. A. Forecasting with a Generalized Exponential Smoothing Method. **Management Science**, v. 6, n. 4, p. 324–330, 1960.